Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Mathematical models in probability and statistics
Exercise A, Question 1

Question:

Give one advantage and one disadvantage of using a mathematical model.

Solution:

Mathematical models are quicker and cheaper to produce.

But the model is only a simplification and does not include all aspects of the real problem.

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Mathematical models in probability and statistics
Exercise A, Question 2

Question:

Describe briefly the process of refining a mathematical model.

Solution:

Predictions based on the model are compared with observed data.

In the light of this comparison the model may be adjusted (refined).

The process of collecting observed data and comparing with revised prediction from the model is repeated.

Solutionbank S1

Edexcel AS and A Level Modular Mathematics

Representation and summary of data – location

Exercise A, Question 1

Question:

State whether each of the following variables is qualitative or quantitative.

a Height of a tree.

b Colour of car.

c Time waiting in a queue.

d Shoe size.

e Name of pupils in a class.

Solution:

a Quantitative as it is numerical.

b Qualitative as it is a descriptive word.

c Quantitative as it is numerical.

d Quantitative as it is numerical.

e Qualitative as they are descriptive words.

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Representation and summary of data – location
Exercise A, Question 2

Question:

State whether each of the following quantitative variables is continuous or discrete.

a Shoe size.

b Length of leaf.

c Number of people on a bus.

d Weight of sugar.

e Time required to run 100 m.

f Lifetime in hours of torch batteries.

Solution:

a Discrete – you do not get a size 4.78

b Continuous – it can be any size – not necessarily a whole number.

c Discrete – you do not get bits of people.

d Continuous – it can be any weight – not necessarily a whole number.

e Continuous – it can be any amount of time – not necessarily a whole number

f Continuous – it can be any amount of time – not necessarily a whole number

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Representation and summary of data – location
Exercise A, Question 3

Question:

Explain why

a 'Type of tree' is a qualitative variable.

b 'The number of pupils in a class' is a discrete quantitative variable.

c 'The weight of a collie dog' is a continuous quantitative variable.

Solution:

a It is a descriptive rather than numerical

b You can not have bits of pupils in a class – only whole numbers. It is quantitative as it is numerical.

c Weight can take any value in a given range.

Therefore it is continuous. It is quantitative as it is numerical.

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Representation and summary of data – location
Exercise A, Question 4

Question:

The nurse at a health centre records the heights, h cm, to the nearest cm, of a group of boys in the same school age group. The frequency table shows the results.

a Complete the table by putting in the cumulative frequency totals.

h	Frequency (f)	Cumulative frequency
165	8	
166	7	
167	9	
168	14	
169	18	
170	16	

b State the number of boys who are less than 168 cm tall.

c Write down the height that is the most common.

Solution:

a

h	Frequency (f)	Cumulative frequency
165	8	**8**
166	7	**15**
167	9	**24**
168	14	**38**
169	18	**56**
170	16	**72**

b 24 boys

c 169 cm

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Representation and summary of data – location
Exercise A, Question 5

Question:

The distribution of the lifetimes of torch batteries is shown in the grouped frequency table below.

Lifetime (nearest 0.1 of an hour)	**Frequency**	**Cumulative frequency**
5.0–5.9	5	
6.0–6.9	8	
7.0–7.9	10	
8.0–8.9	22	
9.0–9.9	10	
10.0–10.9	2	

a Complete the cumulative frequency column.

b Write down the class boundaries for the second group.

c Work out the mid-point of the fifth group.

Solution:

a

Lifetime (hours)	frequency	Cumulative frequency
5.0–5.9	5	**5**
6.0–6.9	8	**13**
7.0–7.9	10	**23**
8.0–8.9	22	**45**
9.0–9.9	10	**55**
10.0–10.9	2	**57**

b **5.95** and **6.95**

c $\frac{8.95+9.95}{2} = \mathbf{9.45}$

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Representation and summary of data – location
Exercise A, Question 6

Question:

The distribution of the weights of two-month-old piglets is shown in the grouped frequency table below.

Weight (kg)	Frequency	Cumulative frequency
1.2–1.3	8	
1.3–1.4	28	
1.4–1.5	32	
1.5–1.6	22	

a Write down the class boundaries for the third group.

b Work out the mid-point of the second group.

Solution:

a **1.4** and **1.5**. there are no gaps therefore the boundaries are the numbers of the class.

b $\frac{1.3+1.4}{2}=\mathbf{1.35}$

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Representation and summary of data – location
Exercise A, Question 7

Question:

Write down which of the following statements are true.

a The weight of apples is discrete data.

b The number of apples on the trees in an orchard is discrete data.

c The amount of time it takes a train to make a journey is continuous data.

d David collected data on car colours by standing at the end of his road and writing down the car colours. Is the data he collected qualitative?

Solution:

A is **not true**

B is **true**

C is **true**

D is **true**

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Representation and summary of data – location
Exercise B, Question 1

Question:

Meryl collected wild mushrooms every day for a week. When she got home each day she weighed them to the nearest 100 g. The weights are shown below.

500 700 400 300 900 700 700

a Write down the mode for these data.

b Calculate the mean for these data.

c Find the median for these data.

On the next day, Meryl collects 650 g of wild mushrooms.

d Write down the effect this will have on the mean, the mode and the median.

Solution:

a **700g** as this is the most often occurring.

b 500 + 700 + 400 + 300 + 900 + 700 + 700 = 4200

$\frac{4200}{7}$ = **600g**

c 300 400 500 **700** 700 700 900

700g is the median as it is the middle value.

d

It will **increase the mean** as 650>600.

The **mode will be unchanged**.

It will **decrease the median** as there will then be 8 values and 650 < median(700). It will lie mid way between 650 and 700.

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Representation and summary of data – location
Exercise B, Question 2

Question:

Joe collects six pieces of data x_1, x_2, x_3, x_4, x_5 and x_6. He works out that Σx is 256.2.

a Calculate the mean for these data.

He collects another piece of data. It is 52.

b Write down the effect this piece of data will have on the mean.

Solution:

a $\frac{256.2}{6} = \mathbf{42.7}$

b It will **increase the mean** as it is greater than the old mean of 42.7.

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Representation and summary of data – location
Exercise B, Question 3

Question:

A small workshop records how long it takes, in minutes, for each of their workers to make a certain item. The times are shown in the table.

Worker	*A*	*B*	*C*	*D*	*E*	*F*	*G*	*H*	*I*	*J*
Time in minutes	7	12	10	8	6	8	5	26	11	9

a Write down the mode for these data.

b Calculate the mean for these data.

c Find the median for these data.

d The manager wants to give the workers an idea of the average time they took. Write down, with a reason, which of the answers to **a**, **b** and **c** he should use.

Solution:

a 8 minutes as everything else occurs once but there are two 8's.

b $\frac{102}{10} = \mathbf{10.2}$ **minutes**

c 5 6 7 8 8 9 10 11 12 26

Median is **8.5 minutes**

d The median would be reasonable. The **mean** is **affected by** the extreme value of **26.**

In this case the mode is close to the median so would be acceptable but this would not always be the case.

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Representation and summary of data – location
Exercise B, Question 4

Question:

A farmer keeps a record of the weekly milk yield of his herd of seven cows over a period of six months. He finds that the mean yield is 24 litres. He buys another cow that he is told will produce 28 litres of milk a week. Work out the effect this will have on the mean milk yield of his cows.

Solution:

The new total milk yield will be $(7 \times 24) + 28 = 196$

New mean $= \frac{196}{8} = \mathbf{24.5}$ **litres**

The **new cow increases the mean milk yield** of the herd.

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Representation and summary of data – location
Exercise B, Question 5

Question:

A clothes retailer has two shops in the town of Field-gate. Shop A employs 15 people and shop B employs 22 people. The mean number of days of sickness in a year taken by the employees in shop A is 4.6 and the mean number of days of sickness taken by the employees in shop B is 6.5 days. Calculate the mean number of days of sickness taken per year by all 37 employees. Give your answer to one decimal place.

Solution:

Total number of days sickness for both shops = (4.6 × 15) + (6.5 × 22) = 212

Mean number of days for all employees $= \frac{212}{37} = \mathbf{5.73 \ days}$

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Representation and summary of data – location
Exercise B, Question 6

Question:

The rainfall in a certain seaside holiday resort was measured, in millimetres, every week for ten weeks. The hours of sunshine were also recorded. The data are shown in the table.

Rainfall (mm)	0	1	2	3	3	26	3	2	3	0
Sunshine (hours)	70	15	10	15	18	0	15	21	21	80

a Calculate the mean rainfall per week.

b Calculate the mean number of hours of sunshine per week.

c Write down the modal amount of rainfall and the modal amount of sunshine per week.

d Work out the median rainfall and the median amount of sunshine per week.

The council plans to produce a brochure and in it they wish to promote the resort as having lots of sunshine and little rain.

e Write down, with reasons, which of the mean, mode or median they should quote in their brochure as the average rainfall and hours of sunshine.

Solution:

a $\frac{43}{10} = \mathbf{4.3}$ **mm**

b $\frac{265}{10} = \mathbf{26.5}$ **hours**

c modal **rainfall** = **3 mm**

modal **sunshine** = **15 hours**

d Rainfall 0 0 1 2 **2** **3** 3 3 3 26

Median rainfall is 2.5 mm

Sunshine 0 10 15 15 **15** **18** 21 21 70 80

Median Sunshine is 16.5 hours

e They will probably want to quote the least rainfall and the highest amount of sunshine. They would use the **median rainfall** and the **mean amount of sunshine**.

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Representation and summary of data – location
Exercise B, Question 7

Question:

The mean marks for a statistics exam were worked out for three classes. Class 1 had 12 students with a mean mark of 78%. Class 2 had 16 students with a mean mark of 84%. Class 3 had 18 students with a mean mark of 54%. Work out the mean % mark to the nearest whole number for all 46 students.

Solution:

Total marks for all students = $(12 \times 78) + (16 \times 84) + (18 \times 54) = 3252$

Mean for all 46 students is $\frac{3252}{46}$ = **71 %** to nearest percent

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Representation and summary of data – location
Exercise C, Question 1

Question:

The marks scored in a multiple choice statistics test by a class of students are:

5	9	6	9	10	6	8	5	5	7	9	7
8	6	10	10	7	9	6	9	7	7	7	8
6	9	7	8	6	7	8	7	9	8	5	7

a Draw a frequency distribution table for these data.

b Calculate the mean mark for these data.

c Write down the number of students who got a mark greater than the mean mark.

d Write down whether or not the mean mark is greater than the modal mark.

Solution:

a

Mark	Frequency
5	**4**
6	**6**
7	**10**
8	**6**
9	**7**
10	**3**

b $(5 \times 4) + (6 \times 6) + (7 \times 10) + (8 \times 6) + (9 \times 7) + (10 \times 3) = 267$

Mean $= \frac{267}{36} = \mathbf{7.42}$ marks

c $6 + 7 + 3 =$ **16** students

d The modal mark is 7 so the **mean** mark **is greater** than the modal mark.

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Representation and summary of data – location
Exercise C, Question 2

Question:

The table shows the number of eggs laid in 25 blackbirds' nests.

Number of eggs	0	1	2	3	4	5	6	7
Number of nests	0	0	0	1	3	9	8	4

Using your knowledge of measures of location decide what number of eggs you could expect a blackbird's nest to contain. Give reasons for your answer.

Solution:

a Modal number of eggs is 5

Median number of eggs is 5

Mean number of eggs is 5.44

We would expect there to be **5 eggs in a nest** as both the **mode and median are 5** and the **mean rounded to the nearest whole number is also 5.**

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Representation and summary of data – location
Exercise C, Question 3

Question:

The table shows the frequency distribution for the number of petals in the flowers of a group of celandines.

Number of petals	Frequency (f)
5	8
6	57
7	29
8	3
9	1

a Work out how many celandines were in the group.

b Write down the modal number of petals.

c Calculate the mean number of petals.

d Calculate the median number of petals.

e If you saw a celandine, write down how many petals. you would expect it to have.

Solution:

a 8 + 57 + 29 + 3 + 1 = **98** celandines

b **6** petals

c (5 × 8) + (6 × 57) + (7 × 29) + (8 × 3) + (9 × 1) = 618 petals

Mean = $\frac{618}{98}$ = **6.31** petals

d median is the 49.5^{th} **value** = **6**

e **6** petals

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Representation and summary of data – location
Exercise C, Question 4

Question:

The frequency table shows the number of breakdowns, *b*, per month recorded by a road haulage firm over a certain period of time.

Breakdowns *b*	Frequency *f*	Cumulative frequency
0	8	8
1	11	19
2	12	31
3	3	34
4	1	35
5	1	36

a Write down the number of months for which the firm recorded the breakdowns.

b Write down the number of months in which there were two or fewer breakdowns.

c Write down the modal number of breakdowns.

d Find the median number of breakdowns.

e Calculate the mean number of breakdowns.

f In a brochure about how many loads reach their destination on time, the firm quotes one of the answers to **c**, **d** or **e** as the number of breakdowns per month for its vehicles. Write down which of the three answers the firm should quote in the brochure.

Solution:

a **36** months

b **31** months

c **2** breakdowns

d Median is 18.5^{th} value = **1**

e $(11 \times 1) + (12 \times 2) + (3 \times 3) + (1 \times 4) + (1 \times 5) = 53$

Mean = $\frac{53}{36} = \mathbf{1.47}$ breakdowns

f The **median since this is the lowest value**.

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Representation and summary of data – location
Exercise C, Question 5

Question:

A company makes school blazers in eight sizes. The first four sizes cost £48. The next three sizes cost £60 and the largest size costs £76.80. Write down, with a reason which of the mean, mode, or median cost the company is likely to use in advertising its average price.

Solution:

Mode is £48

Median is 4.5$^{\text{th}}$ value $= \frac{48+60}{2} = £\,54$

Mean is $\frac{(4\times 48)+(3\times 60)+(76.80)}{8} = \frac{448.8}{8} = £\,56.1$

The company would quote the mode as it is the lowest of the three averages.

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Representation and summary of data – location
Exercise D, Question 1

Question:

A hotel is worried about the reliability of its lift. It keeps a weekly record of the number of times it breaks down over a period of 26 weeks. The data collected are summarised in the table opposite.

a Estimate the mean number of breakdowns.

b Use interpolation to estimate the median number of breakdowns.

c The hotel considers that an average of more than one breakdown per week is not acceptable. Judging from your answers to **b** and **c** write down with a reason whether or not you think the hotel should consider getting a new lift.

Number of breakdowns	Frequency of breakdowns (*f*)
0–1	18
2–3	7
4–5	1

Solution:

a Mean is $\frac{(0.5 \times 18) + (2.5 \times 7) + (4.5 \times 1)}{26} = \mathbf{1.19}$ **breakdowns**

b Median value is 13th value. This is in the first class.

Let m be the median.

$\frac{m-0}{1.5-0} = \frac{13-0}{18-0}$ so $m \approx 0.722$ (3 sf)

c The median is less than 1 and the mean is only a little about 1 – it is 1 if rounded to the nearest whole number. The hotel need not consider getting a new lift yet, but should keep an eye on the situation.

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Representation and summary of data – location
Exercise D, Question 2

Question:

The weekly wages (to the nearest £) of the production line workers in a small factory is shown in the table.

Weekly wage £	**Number of workers, *f*,**
175–225	4
226–300	8
301–350	18
351–400	28
401–500	7

a Write down the modal class.

b Calculate an estimate of the mean wage.

c Use interpolation to find an estimate for the median wage.

Solution:

a 351 − 400

b $\frac{(200\times4)+(263\times8)+(325.5\times18)+(375.5\times28)+(450.5\times7)}{65}=\frac{800+2104+5859+10514+3153.5}{65}=\frac{22430.5}{65}=\mathbf{345.08}$

c Median is the 32.5th value. This is in the 351 − 400 class

$$\frac{m-350.5}{400.5-350.5}=\frac{32.5-30}{58-30}$$

$$\frac{m-350.5}{50}=\frac{2.5}{28}$$

$m - 350.5 = 4.46$

$m =$ £**354.96**

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Representation and summary of data – location
Exercise D, Question 3

Question:

The noise levels at 30 locations near an outdoor concert venue were measured to the nearest decibel. The data collected is shown in the grouped frequency table.

Noise (decibels)	65–69	70–74	75–79	80–84	85–89	90–94	95–99
Frequency (*f*)	1	4	6	6	8	4	1

a Calculate an estimate of the mean noise level.

b A noise level above 82 decibels was considered unacceptable. Estimate the number of locations that had unacceptable noise levels.

Solution:

$$\frac{(67\times 1)+(72\times 4)+(77\times 6)+(82\times 6)+(87\times 8)+(92\times 4)+(97\times 1)}{30}$$

$$= \frac{2470}{30} = \mathbf{82.3\ \ decibels}$$

b

82 is the middle of class 80 to 84.

So number of locations with unacceptable noise levels is 3 + 8 + 4 + 1 = **16**

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Representation and summary of data – location
Exercise D, Question 4

Question:

DIY store **A** considered that it was good at employing older workers. A rival store **B** disagreed and said that it was better. The two stores produced a frequency table of their workers' ages. The table is shown below

Age of workers (to the nearest year)	Frequency store A	Frequency store B
16–25	5	4
26–35	16	12
36–45	14	10
46–55	22	28
56–65	26	25
66–75	14	13

By comparing estimated means for each store decide which store employs more older workers.

Solution:

a Store A $\frac{(20.5\times5)+(30.5\times16)+(40.5\times14)+(50.5\times22)+(60.5\times26)+(70.5\times14)}{97}$

$= \frac{4828.5}{97} = \mathbf{50}$ **years**

Store B $\frac{(20.5\times4)+(30.5\times12)+(40.5\times10)+(50.5\times28)+(60.5\times25)+(70.5\times13)}{92}$

$= \frac{4696}{92} = \mathbf{51}$ **years**

Store B employs older workers but not by a great margin.

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Representation and summary of data – location
Exercise D, Question 5

Question:

The speeds of vehicles passing a checkpoint were measured over a period of one hour, to the nearest mph. The data collected is shown in the grouped frequency table.

Speed (mph)	21–30	31–40	41–50	51–60	61–65	66–70	71–75
No. of vehicles (f)	4	7	38	42	5	3	1

a Write down the modal class.

b Calculate the difference, to two decimal places, between the median and the mean estimated speeds.

c The speed limit on the road is 60 mph. Work out an estimate for the percentage of cars that exceeded the speed limit.

Solution:

a 51 − 60

b mean $= \frac{(25.5 \times 4) + (35.5 \times 7) + (45.5 \times 38) + (55.5 \times 42) + (63 \times 5) + (68 \times 3) + (73 \times 1)}{100}$

$= \frac{5002.5}{100} = 50.025$ mph

median = 50$^{\text{th}}$ value. This is in the 51 to 60 class

$\frac{m - 50.5}{60.5 - 50.5} = \frac{50 - 49}{91 - 49}$

$\frac{m - 50.5}{10} = \frac{1}{42}$

$42m = 10 + (49.5 \times 42)$

median = 50.738 mph

Difference is 50.738 − 50.025

Answer 0.713 mph

c Approximately 9% out of the 100 exceeded 60 mph

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Representation and summary of data – location
Exercise E, Question 1

Question:

Calculate the mean of the following data set (x) using the coding $y = \frac{x}{10}$.

110 90 50 80 30 70 60

Solution:

$11 + 9 + 5 + 8 + 3 + 7 + 6 = 49$

$\frac{49}{7} = 7$

mean = 70

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Representation and summary of data – location
Exercise E, Question 2

Question:

Find the mean of the following data set (x) using the coding $y = \frac{x-3}{7}$.

52 73 31 73 38 80 17 24

Solution:

7 + 10 + 4 + 10 + 5 + 11 + 2 + 3 = 52

$\frac{52}{8} = 6.5$

Mean is $(6.5 \times 7) + 3 =$ **48.5**

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Representation and summary of data – location
Exercise E, Question 3

Question:

a Calculate the mean of 1, 2, 3, 4, 5 and 6.

Using your answer to **a**:

b Write down the mean of:

i 2, 4, 6, 8, 10 and 12,

ii 10, 20, 30, 40, 50 and 60,

iii 12, 22, 32, 42, 52 and 62.

Solution:

a $1 + 2 + 3 + 4 + 5 + 6 = 21$ Mean $= \frac{21}{6} = \mathbf{3.5}$

b i) 7, coding is $\times 2$

ii) 35, coding is $\times 10$

iii) 37, coding is $\times 10$ and add 2

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Representation and summary of data – location
Exercise E, Question 4

Question:

The coded mean price of televisions in a shop was worked out. Using the coding $y = \frac{x - 65}{200}$ the mean price was 1.5. Find the true mean price of the televisions.

Solution:

$(1.5 \times 200) + 65 =$ **365**

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Representation and summary of data – location
Exercise E, Question 5

Question:

The grouped frequency table shows the age (a years) at which a sample of 100 women had their first child.

Age of Women (a years)	Frequency (f)	Mid-point (x)	$y = \frac{x-14}{2}$
11–21	11		
21–27	24		
27–31	27		
31–37	26		
37–43	12		

a Copy and complete the table

b Use the coding $y = \frac{x-14}{2}$ to calculate an estimate of the mean age at which women have their first child.

Solution:

a

Age of Women (a)	Frequency (f)	Mid Point (x)	$y = \frac{x-14}{2}$
11 – 21	11	**16**	**1**
21 – 27	24	**24**	**5**
27 – 31	27	**29**	**7.5**
31 – 37	26	**34**	**10**
37 – 43	12	**40**	**13**

b

$$\text{Coded mean} = \frac{(11\times 1)+(24\times 5)+(27\times 7.5)+(26\times 10)+(12\times 13)}{100} = \frac{749.5}{100} = 7.495$$

Actual Mean = (7.495 × 2) + 14 = **28.99** = **29 years** to nearest year

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Representation and summary of data – location
Exercise F, Question 1

Question:

The following figures give the number of children injured on English roads each month for a certain period of seven months.

55	72	50	66	50	47	38

a Write down the modal number of injuries.

b Find the median number of injuries.

c Calculate the mean number of injuries.

Solution:

a 50

b 50

c 54

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Representation and summary of data – location
Exercise F, Question 2

Question:

The mean Science mark for one group of eight students is 65. The mean mark for a second group of 12 students is 72. Calculate the mean mark for the combined group of 20 students.

Solution:

$(8 \times 65) + (12 \times 72) = 1384$

$\frac{1384}{20} = \mathbf{69.2\ \ marks}$

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Representation and summary of data – location
Exercise F, Question 3

Question:

A computer operator transfers an hourly wage list from a paper copy to her computer. The data transferred is given below:

£5.50 £6.10 £7.80 £6.10 £9.20 £91.00 £11.30

a Find the mean, mode and median of these data.

The office manager looks at the figures and decides that something must be wrong.

b Write down, with a reason, the mistake that has probably been made.

Solution:

a Mean $= \frac{137}{7} = 19.57$

Mode = £6.10

Median = £7.80

b A transfer error. Probably the value £91.00 is wrong

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Representation and summary of data – location
Exercise F, Question 4

Question:

A piece of data was collected from each of nine people. The data was found to have a mean value of 35.5. One of the nine people had given a value of 42 instead of a value of 32.

a Write down the effect this will have had on the mean value.

The correct data value of 32 is substituted for the incorrect value of 42.

b Calculate the new mean value.

Solution:

a The mean is higher than it should be.

b The mean will be $(9 \times 35.5 - 10) / 9$

$\frac{309.5}{9} = \mathbf{34.39}$

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Representation and summary of data – location
Exercise F, Question 5

Question:

On a particular day in the year 2007, the prices (x) of six shares were as follows:

807 967 727 167 207 767

Use the coding $y = \frac{x-7}{80}$ to work out the mean value of the shares.

Solution:

Coded values are: 10 12 9 2 2.5 9.5

Mean $= \frac{10+12+9+2+2.5+9.5}{6} = 7.5$

Uncoded this is $7.5 \times 80 + 7 = 607$

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Representation and summary of data – location
Exercise F, Question 6

Question:

The coded mean of employee's annual earnings (£x) for a store is 18. The code used was $y = \frac{x - 720}{1000}$. Work out the uncoded mean earnings.

Solution:

$18 \times 1000 + 720 =$ **£18720**

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Representation and summary of data – location
Exercise F, Question 7

Question:

Different teachers using different methods taught two groups of students. Both groups of students sat the same examination at the end of the course. The students' marks are shown in the grouped frequency table.

a Work out an estimate of the mean mark for Group *A* and an estimate of the mean mark for Group *B*.

b Write down whether or not the answer to **a** suggests that one method of teaching is better than the other. Give a reason for your answer.

Exam Mark	Frequency Group *A*	Frequency Group *B*
20–29	1	1
30–39	3	2
40–49	6	4
50–59	6	13
60–69	11	15
70–79	10	6
80–89	8	3

Solution:

a Group *A*:

$(1 \times 24.5) + (3 \times 34.5) + (6 \times 44.5) + (6 \times 54.5) + (11 \times 64.5) + (10 \times 74.5) + (8 \times 84.5)$

Mean $\frac{2852.5}{45} = \mathbf{63.39}$ **marks**

Group *B*:

$(1 \times 24.5) + (2 \times 34.5) + (4 \times 44.5) + (13 \times 54.5) + (15 \times 64.5) + (6 \times 74.5) + (3 \times 84.5)$

Mean $= \frac{2648}{44} = \mathbf{60.18}$ **marks**

b The method used to teach group A is best as the mean mark is higher.

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Representation and summary of data – location
Exercise F, Question 8

Question:

The table summarises the distances travelled by 150 students to college each day.

a Use interpolation to calculate the median distance for these data.

The mid-point is x and the corresponding frequency is f. Calculations give the following values $\Sigma fx = 1056$

b Calculate an estimate of the mean distance for these data.

Distance (nearest km)	Number of students
0–2	14
3–5	24
6–8	70
9–11	32
12–14	8
15–17	2

Solution:

a We need the 75th student

This is in the 6 – 8 class.

$$\frac{m-5.5}{8.5-5.5}=\frac{75-38}{108-38}$$

70m = 111 + 385 = 496

median = **7.09**

b Mean = $\frac{1056}{150}$ = **7.04**

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Representation and summary of data – location
Exercise F, Question 9

Question:

Chloe enjoys playing computer games.

Her parents think that she spends too much time playing games. Chloe decides to keep a record of the amount of time she plays computer games each day for 50 days. She draws up a grouped frequency table to show these data.

The mid-point of each class is represented by x and its corresponding frequency is f. (You may assume $\Sigma fx = 1275$)

a Calculate an estimate of the mean time Chloe spends on playing computer games each day. Chloe's parents thought that the mean was too high and suggested that she try to reduce the time spent on computer games. Chloe monitored the amount of time she spent on computer games. Chloe monitored the amount of time spent on games for another 40 days and found that at the end of 90 days her overall mean time was 26 minutes.

b Find an estimate for the mean amount of time Chloe spent on games for the 40 days.

c Comment on the two mean values.

Time to the nearest minute	Frequency
10–14	1
15–19	5
20–24	18
25–29	15
30–34	7
35–39	3
40–44	1

Solution:

a

Mean $= \frac{1275}{50} = \mathbf{25.5}$ **minutes**

b

Minutes for 90 days = 90 × 26 = 2340 minutes

Minutes for first 50 days = 1275 minutes

Minutes for last 40 days = 2340 − 1275 = 1065

Mean for last 40 days $= \frac{1065}{40} = \mathbf{26.63}$ **minutes**

c Chloe did not reduce the number of minutes spent playing computer games each day. She increased the time. Chloe spent over a minute more on games in the last 40 days.

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Representation and summary of data – location
Exercise F, Question 10

Question:

The lifetimes of 80 batteries, to the nearest hour, is shown in the table below.

a Write down the modal class for the lifetime of the batteries.

Lifetime (hours)	Number of batteries
6–10	2
11–15	10
16–20	18
21–25	45
26–30	5

b Use interpolation to find the median lifetime of the batteries.

The mid-point of each class is represented by x and its corresponding frequency by f, giving $\Sigma fx = 1645$.

c Calculate an estimate of the mean lifetime of the batteries.

Another batch of 12 batteries is found to have an estimated mean lifetime of 22.3 hours.

d Find the mean lifetime for all 92 batteries.

Solution:

a Modal Class is **21 – 25** hours

b We need the 40^{th} value. This is in the 21 – 25 class.

$$\frac{m-20.5}{25.5-20.5}=\frac{40-30}{75-30}$$

45m = 50 + 922.5 = 972.5

median = **21.6 hours**

c

mean = **20.6 hours**

d 12 × 22.3 = 267.6 hours

Total hours for all 92 batteries is 267.6 + 1645 = 1912.6

Mean life for 92 batteries is 20.8 hours

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Representation and summary of data – measures of dispersion
Exercise A, Question 1

Question:

15 students do a mathematics test. Their marks are shown opposite.

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

a Find the value of the median.

b Find Q_1 and Q_3.

c Work out the interquartile range.

Solution:

a 3,3,4,5,6,7,7,7,8,8,9,9,10,11,12

Median: $\frac{15}{2} = 7.5$ therefore 8th value is needed. Median = **7**

b $Q_1 : \frac{15}{4} = 3.75$ therefore the 4th value is needed Q_1 = **5**

$Q_3 : 3 \times \frac{15}{4} = 11.25$ therefore the 12th value is needed Q_3 = **9**

c IQR = 9 − 5 = **4**

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Representation and summary of data – measures of dispersion
Exercise A, Question 2

Question:

A group of workers were asked to write down their weekly wage. The wages were:

£550 £400 £260 £320 £500 £450 £460 £480 £510 £490 £505

a Work out the range for these wages.

b Find Q_1 and Q_3.

c Work out the interquartile range.

Solution:

a Range = £550 − £260 = **£290**

b

260 320 400 450 460 480 490 500 505 510 550

$Q_1 : \frac{11}{4} = 2.75$ therefore the 3rd value is needed $Q_1 = \mathbf{400}$

$Q_3 : 3 \times \frac{11}{4} = 8.25$ therefore the 9th value is needed $Q_3 = \mathbf{505}$

c IQR = 505 − 400 = **105**

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Representation and summary of data – measures of dispersion
Exercise A, Question 3

Question:

A superstore records the number of hours overtime worked by their employees in one particular week. The results are shown in the table.

a Fill in the cumulative frequency column and work out how many employees the superstore had in that week.

b Find Q_1 and Q_3.

c Work out the interquartile range.

Number of hours	Frequency	Cumulative frequency
0	25	
1	10	
2	20	
3	10	
4	25	
5	10	

Solution:

a

CF = 25 35 55 65 90 100

Superstore had 100 employees

b $Q_1 : \frac{100}{4} = 25$ therefore the average of the 25th and 26th values is $Q_1 = \mathbf{0.5}$

$Q_3 : 3 \times \frac{100}{4} = 75$ therefore the average of the 75th and 76th values is $Q_3 = \mathbf{4}$

c IQR = 4 − 0.5 = **3.5**

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Representation and summary of data – measures of dispersion
Exercise A, Question 4

Question:

A moth trap was set every night for five weeks. The number of moths caught in the trap was recorded. The results are shown in the table.

Number of moths	Frequency
7	2
8	5
9	9
10	14
11	5

Find the interquartile range.

Solution:

CF = 2 7 16 30 35

$Q_1 : \frac{35}{4} = 8.75$ therefore the 9th value is needed $Q_1 = 9$

$Q_3 : 3 \times \frac{35}{4} = 26.25$ therefore the 27th value is needed $Q_3 = 10$

IQR = 10 − 9 = **1**

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Representation and summary of data – measures of dispersion
Exercise A, Question 5

Question:

The weights of 31 Jersey cows were recorded to the nearest kilogram. The weights are shown in the table.

Weight of cattle (kg)	Frequency	Cumulative frequency
300–349	3	
350–399	6	
400–449	10	
450–499	7	
500–549	5	

a Complete the cumulative frequency column in the table.

b Find the lower quartile, Q_1.

c Find the upper quartile, Q_3.

d Find the interquartile range.

Solution:

a
CF = 3 9 19 26 31

b Data is continuous so:

$Q_1 : \frac{31}{4} = 7.75^{\text{th}}$ value so Q_1 is in class 350 − 399

$$\begin{aligned}
\frac{Q_1 - 349.5}{399.5 - 349.5} &= \frac{7.75 - 3}{9 - 3} \\
\frac{Q_1 - 349.5}{50} &= \frac{4.75}{6} \\
Q_1 &= 39.58 + 349.5 \\
Q_1 &= \mathbf{389.1}
\end{aligned}$$

c $Q_3 : 3 \times \frac{31}{4} = 23.25^{\text{th}}$ value so Q_3 is in class 450 − 499

$$\begin{aligned}
\frac{Q_3 - 449.5}{499.5 - 449.5} &= \frac{23.25 - 19}{26 - 19} \\
\frac{Q_3 - 449.5}{50} &= \frac{4.25}{7} \\
Q_3 &= 30.36 + 449.5 \\
&= \mathbf{479.9}
\end{aligned}$$

d IQR = 479.86 − 389.08 = **90.8**

Solutionbank S1

Edexcel AS and A Level Modular Mathematics

Representation and summary of data – measures of dispersion

Exercise A, Question 6

Question:

The number of visitors to a hospital in a week was recorded. The results are shown in the table.

Number of visitors	Frequency
500–1000	10
1000–1500	25
1500–2000	15
2000–2500	5
2500–3000	5

Giving your answers to the nearest whole number find:

a the lower quartile Q_1, **b** the upper quartile Q_3, **c** the interquartile range.

Solution:

Data is continuous so

a

CF = 10 35 50 55 60

$Q_1 : \frac{60}{4} = 15^{\text{th}}$ value so Q_1 is in class $1000 - 1500$

$$\begin{aligned} \frac{Q_1 - 1000}{1500 - 1000} &= \frac{15 - 10}{35 - 10} \\ Q_1 - 1000 &= 100 \\ Q_1 &= \mathbf{1100} \end{aligned}$$

b $Q_3 : 3 \times \frac{60}{4} = 45^{\text{th}}$ value so Q_3 is in class $1500 - 2000$

$$\begin{aligned} \frac{Q_3 - 1500}{2000 - 1500} &= \frac{45 - 35}{50 - 35} \\ Q_3 - 1500 &= 333.33 \\ Q_3 &= \mathbf{1833} \end{aligned}$$

c IQR = 1833 − 1100 = **733**

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Representation and summary of data – measures of dispersion
Exercise A, Question 7

Question:

The lengths of a number of slow worms were measured, to the nearest mm. The results are shown in the table.

a Work out how many slow worms were measured.

b Find the interquartile range for the lengths of the slow worms.

Lengths of slow worms (mm)	Frequency
125–139	4
140–154	4
155–169	2
170–184	7
185–199	20
200–214	24
215–229	10

Solution:

Data is continuous

a
CF = 4 8 10 17 37 61 71

71 slow worms

b $Q_1 : \frac{71}{4} = 17.75^{\text{th}}$ value so Q_1 is in class 185 − 199

$$\begin{aligned}\frac{Q_1 - 184.5}{199.5 - 184.5} &= \frac{17.75 - 17}{37 - 17}\\ Q_1 - 184.5 &= 0.5625\\ Q_1 &= 185.0625\end{aligned}$$

$Q_3 : 3 \times \frac{71}{4} = 53.25^{\text{th}}$ value so Q_3 is in class 200−214

$$\begin{aligned}\frac{Q_3 - 199.5}{214.5 - 199.5} &= \frac{53.25 - 37}{61 - 37}\\ Q_3 - 199.5 &= 243.75/24\\ Q_3 &= 209.656\end{aligned}$$

IQR = 209.656 − 185.0625 = **24.6 to 3SF**

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Representation and summary of data – measures of dispersion
Exercise B, Question 1

Question:

A gardener counted the peas in a number of pea pods. The results are shown in the table.

Number of peas	Frequency	Cumulative frequency
6	8	
7	12	
8	36	
9	18	
10	15	
11	10	

a Complete the cumulative frequency column.

b Calculate the 80th percentile.

c Calculate the 40th percentile

d Calculate the 65th percentile.

Solution:

a CF 8, 20, 56, 74, 89, 99

b 80^{th} : $\frac{80}{100} \times 99 = 79.2$ therefore 80^{th} term needed P_{80} = **10**

c 40^{th} : $\frac{40}{100} \times 99 = 39.6$ therefore 40^{th} term needed P_{40} = **8**

d 65^{th} : $\frac{65}{100} \times 99 = 64.35$ therefore 65^{th} term needed P_{65} = **9**

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Representation and summary of data – measures of dispersion
Exercise B, Question 2

Question:

A shopkeeper goes to a clothes fair. He records the costs of jeans. The costs are shown in the table.

Cost of jeans (£'s)	Frequency	Cumulative frequency
10–15	11	
16–20	35	
21–25	34	
26–30	16	
31–35	10	
36–40	5	

a Complete the cumulative frequency table.

b Calculate P_{20}.

c Calculate P_{80}.

d Calculate the 20% to 80% interpercentile range.

Solution:

a CF 11, 46, 80, 96, 106, 111

b

$$20^{\text{th}}: \frac{20}{100} \times 111 = 22.2$$

$$\frac{P_{20} - 15.5}{20.5 - 15.5} = \frac{22.2 - 11}{46 - 11}$$

$$P_{20} = \mathbf{17.1}$$

c

$$80^{\text{th}}: \frac{80}{100} \times 111 = 88.8$$

$$\frac{P_{80} - 25.5}{30.5 - 25.5} = \frac{88.8 - 80}{96 - 80}$$

$$P_{80} = \mathbf{28.25}$$

d 20% to 80% interpercentile = 28.25 − 17.1 = **11.15**

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Representation and summary of data – measures of dispersion
Exercise B, Question 3

Question:

The table shows the monthly income for a number of workers in a factory.

Calculate the 34% to 66% interpercentile range.

Monthly income (£'s)	Frequency
900–1000	3
1000–1100	24
1100–1200	28
1200–1300	15

Solution:

$34^{\text{th}} : \frac{34}{100} \times 70 = 23.8$

$$\frac{P_{34} - 1000}{1100 - 1000} = \frac{23.8 - 3}{27 - 3}$$

$$P_{34} = 1086.7$$

$66^{\text{th}} : \frac{66}{100} \times 70 = 46.2$

$$\frac{P_{66} - 1100}{1200 - 1100} = \frac{46.2 - 27}{55 - 27}$$

$$P_{66} = 1168.57$$

34% to 66% interpercentile = 1168.57 − 1086.7 = **81.9**

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Representation and summary of data – measures of dispersion
Exercise B, Question 4

Question:

A train travelled from Lancaster to Preston. The times, to the nearest minute, it took for the journey were recorded over a certain period. The times are shown in the table.

Time for journey (minutes)	15–16	17–18	19–20	21–22
Frequency	5	10	35	10

Calculate the 5% to 95% interpercentile range.

Solution:

$5^{\text{th}} : \frac{5}{100} \times 60 = 3$

$\frac{P_5 - 14.5}{16.5 - 14.5} = \frac{3 - 0}{5 - 0}$

$P_5 = 15.7$

$95^{\text{th}} : \frac{95}{100} \times 60 = 57$

$\frac{P_{95} - 20.5}{22.5 - 20.5} = \frac{57 - 50}{60 - 50}$

$P_{95} = 21.9$

5% to 95% interpercentile $= 21.9 - 15.7 =$ **6.2**

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Representation and summary of data – measures of dispersion
Exercise B, Question 5

Question:

A roadside assistance firm kept a record over a week of the amount of time, in minutes, people were kept waiting for assistance in a particular part of the country. The time taken was from the time the phone call was received to the arrival of the breakdown mechanic. The times are shown below.

Time Waiting (minutes)	20–30	30–40	40–50	50–60	60–70
Frequency	6	10	18	13	2

a Work out the number of people who called for assistance.

b Calculate the 30th percentile.

c Calculate the 65th percentile.

Solution:

a CF 6, **16**, **34**, **47**, 49

$30^{\text{th}} : \frac{30}{100} \times 49 \quad = 14.7$

b $\frac{P_{30} - 30}{40 - 30} = \frac{14.7 - 6}{16 - 6}$

$P_{30} = \mathbf{38.7}$

$65^{\text{th}} : \frac{65}{100} \times 49 \quad = 31.85$

c $\frac{P_{80} - 40}{50 - 40} = \frac{31.85 - 16}{34 - 16}$

$P_{65} = \mathbf{48.8}$

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Representation and summary of data – measures of dispersion
Exercise C, Question 1

Question:

Given that for a variable x:

$\Sigma x = 24$ $\Sigma x^2 = 78$ $n = 8$

Find:

a The mean. **b** The variance σ^2. **c** The standard deviation σ.

Solution:

a Mean $= \frac{24}{8} = \mathbf{3}$

b Variance $= \frac{78}{8} - 3^2 = \mathbf{0.75}$

c Standard deviation $= \sqrt{0.75} = \mathbf{0.866}$

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Representation and summary of data – measures of dispersion
Exercise C, Question 2

Question:

Ten collie dogs are weighed (w kg). The following summary data for the weights are shown below:

$\Sigma w = 241$ $\Sigma w^2 = 5905$

Use this summary data to find the standard deviation of the collies' weights.

Solution:

Standard deviation $= \sqrt{\frac{5905}{10} - \left(\frac{241}{10}\right)^2} = \mathbf{3.11}$

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Representation and summary of data – measures of dispersion
Exercise C, Question 3

Question:

Eight students' heights (h cm) are measured. They are as follows:

165 170 190 180 175 185 176 184

a Work out the mean height of the students.

b Given $\Sigma h^2 = 254\,307$ work out the variance. Show all your working.

c Work out the standard deviation.

Solution:

a $\Sigma h = 165 + 170 + 190 + 180 + 175 + 185 + 176 + 184 = 1425$

Mean $= \frac{1425}{8} = \mathbf{178.125 = 178}$

b Variance $= \frac{254307}{8} - 178.125^2 = \mathbf{59.9}$

c Standard deviation $= \sqrt{59.9} = \mathbf{7.74}$

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Representation and summary of data – measures of dispersion
Exercise C, Question 4

Question:

For a set of 10 numbers:

$\Sigma x = 50$ $\Sigma x^2 = 310$

For a set of 15 numbers:

$\Sigma x = 86$ $\Sigma x^2 = 568$

Find the mean and the standard deviation of the combined set of 25 numbers.

Solution:

$\Sigma x = 50 + 86 = 136$

$\Sigma x^2 = 310 + 568 = 878$

Mean $= \frac{136}{25} = \mathbf{5.44}$

Standard deviation $= \sqrt{\frac{878}{25} - \left(\frac{136}{25}\right)^2} = \mathbf{2.35}$

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Representation and summary of data – measures of dispersion
Exercise C, Question 5

Question:

The number of members (m) in six scout groups was recorded. The summary statistics for these data are:

$\Sigma m = 150$ $\Sigma m^2 = 3846$

a Work out the mean number of members in a scout group.

b Work out the standard deviation of the number of members in the scout groups.

Solution:

a Mean $= \frac{150}{6} = \mathbf{25}$

b Standard deviation $= \sqrt{\frac{3846}{6} - \left(\frac{150}{6}\right)^2} = \mathbf{4}$

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Representation and summary of data – measures of dispersion
Exercise C, Question 6

Question:

There are two routes for a worker to get to his office. Both the routes involve hold ups due to traffic lights. He records the time it takes over a series of six journeys for each route. The results are shown in the table.

Route 1	15	15	11	17	14	12
Route 2	11	14	17	15	16	11

a Work out the mean time taken for each route.

b Calculate the variance and the standard deviation of each of the two routes.

c Using your answers to **a** and **b** suggest which route you would recommend. State your reason clearly.

Solution:

a Route 1: Mean $= \frac{84}{6} = \mathbf{14}$

Route 2: Mean $= \frac{84}{6} = \mathbf{14}$

b Route 1: Standard deviation $= \sqrt{\frac{1200}{6} - \left(\frac{84}{6}\right)^2} = \mathbf{2}$

Route 2: Standard deviation $= \sqrt{\frac{1208}{6} - \left(\frac{84}{6}\right)^2} = \mathbf{2.31}$

c Route 1 as although the means are the same the standard deviation is less meaning the times are more consistent and the journey is less likely to take longer than expected.

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Representation and summary of data – measures of dispersion
Exercise D, Question 1

Question:

For a certain set of data:

$\Sigma fx = 1975$ $\Sigma fx^2 = 52\ 325$ $n = 100$

Work out the variance for these data.

Solution:

Variance $= \frac{52325}{100} - \left(\frac{1975}{100}\right)^2 = \mathbf{133.2 = 133}$

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Representation and summary of data – measures of dispersion
Exercise D, Question 2

Question:

For a certain set of data

$\Sigma fx = 264$ $\Sigma fx^2 = 6456$ $n = 12$

Work out the standard deviation for these data.

Solution:

Standard deviation $= \sqrt{\frac{6456}{12} - \left(\frac{264}{12}\right)^2} = \mathbf{7.35}$

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Representation and summary of data – measures of dispersion
Exercise D, Question 3

Question:

Nahab asks the students in his year group how much pocket money they get per week. The results, rounded to the nearest pound, are shown in the table.

Number of £'s (x)	Number of students f	fx	fx^2
8	14		
9	8		
10	28		
11	15		
12	20		
Totals			

a Complete the table.

b Using the formula $\frac{\Sigma fx^2}{\Sigma f} - \left(\frac{\Sigma fx}{\Sigma f}\right)^2$ work out the variance for these data.

c Work out the standard deviation for these data.

Solution:

a fx: 112, 72, 280, 165, 240, total 869

fx^2: 896, 648, 2800, 1815, 2880, total 9039

total f: 85

b Variance $= \frac{9039}{85} - \left(\frac{869}{85}\right)^2 = \mathbf{1.82}$

c Standard deviation $= \sqrt{1.82} = \mathbf{1.35}$

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Representation and summary of data – measures of dispersion
Exercise D, Question 4

Question:

In a student group, a record was kept of the number of days absence each student had over one particular term. The results are shown in the table.

Number days absent (x)	**Number of students f**	fx	fx^2
0	12		
1	20		
2	10		
3	7		
4	5		

a Complete the table.

b Calculate the variance for these data.

c Work out the standard deviation for these data.

Solution:

a fx**: 0, 20, 20, 21, 20**

fx^2**: 0, 20, 40, 63, 80**

b Variance $= \frac{203}{54} - \left(\frac{81}{54}\right)^2 = \mathbf{1.51}$

c Standard deviation $= \sqrt{1.51} = \mathbf{1.23}$

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Representation and summary of data – measures of dispersion
Exercise D, Question 5

Question:

A certain type of machine contained a part that tended to wear out after different amounts of time. The time it took for 50 of the parts to wear out was recorded. The results are shown in the table.

Lifetime in hours	Number of parts	Mid-point x	fx	fx^2
$5 < h \le 10$	5			
$10 < h \le 15$	14			
$15 < h \le 20$	23			
$20 < h \le 25$	6			
$25 < h \le 30$	2			

a Complete the table.

b Calculate an estimate for the variance and the standard deviation for these data.

Solution:

a midpoint: 7.5, 12.5, 17.5, 22.5, 27.5

***fx*: 37.5, 175, 402.5, 135, 55**

***fx*^{2}: 281.25, 2187.5, 7043.75, 3037.5, 1512.5**

b Variance $= \frac{14062.5}{50} - \left(\frac{805}{50}\right)^2 = \mathbf{22.0}$

Standard deviation $= \sqrt{22.04} = \mathbf{4.69}$

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Representation and summary of data – measures of dispersion
Exercise D, Question 6

Question:

The heights (x cm) of a group of 100 women were recorded.

The summary data is as follows:

$\Sigma fx = 17\,100$ $\Sigma fx^2 = 2\,926\,225$

Work out the variance and the standard deviation.

Solution:

Variance $= \frac{2926225}{100} - \left(\frac{17100}{100}\right)^2 = \mathbf{21.25}$

Standard deviation $= \sqrt{21.25} = \mathbf{4.61}$

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Representation and summary of data – measures of dispersion
Exercise E, Question 1

Question:

a Work out the standard deviation of the following data.

x	11	13	15	20	25

b Use the following coding to find the standard deviation of the data. Show all the working.

i $x - 10$ **ii** $\frac{x}{10}$ **iii** $\frac{x-3}{2}$

Solution:

a Standard deviation $= \sqrt{\frac{1540}{5} - \left(\frac{84}{5}\right)^2} = \mathbf{5.08}$

b i coded data 1, 3, 5, 10, 15

Standard deviation of coded data$= \sqrt{\frac{360}{5} - \left(\frac{34}{5}\right)^2} = 5.08$

Standard deviation = **5.08**

ii coded data 1.1, 1.3, 1.5, 2.0, 2.5

Standard deviation of coded data$= \sqrt{\frac{15.4}{5} - \left(\frac{8.4}{5}\right)^2} = 0.\mathbf{508}$

Standard deviation = 0.508 × 10 = **5.08**

iii coded data 4, 5, 6, 8.5, 11

Standard deviation of coded data $= \sqrt{\frac{270.25}{5} - \left(\frac{34.5}{5}\right)^2} = 2.5377$

Standard deviation = 2.5377 × 2 = **5.08**

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Representation and summary of data – measures of dispersion
Exercise E, Question 2

Question:

Use the following codings to find the standard deviation of the following weights, *w*.

w	210	260	310	360	410

i $\frac{w}{10}$

ii $w - 200$

iii $\frac{w-10}{200}$

Solution:

i coded data 21, 26, 31, 36, 41

Standard deviation of coded data = $\sqrt{\frac{5055}{5} - \left(\frac{155}{5}\right)^2} = 7.07$

Standard deviation = 7.07 × 10 = **70.7**

ii coded data 10, 60, 110, 160, 210

Standard deviation of coded data = $\sqrt{\frac{85500}{5} - \left(\frac{550}{5}\right)^2} = 70.7$

Standard deviation = **70.7**

iii coded data 1, 1.25, 1.5, 1.75, 2

Standard deviation of coded data = $\sqrt{\frac{11.875}{5} - \left(\frac{7.5}{5}\right)^2} = 0.354$

Standard deviation = 0.354 × 200 = **70.7**

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Representation and summary of data – measures of dispersion
Exercise E, Question 3

Question:

a The coding $y = \frac{x - 50}{28}$ gives a standard deviation of 0.01 for y. Work out the standard deviation for x.

b The coding $y = \frac{h}{15}$ produced a standard deviation for y of 0.045. What is the standard deviation of h?

c The coding $y = x - 14$ produced a standard deviation for y of 2.37. What is the standard deviation of x?

d The coding $y = \frac{s - 5}{10}$ produced a standard deviation for y of 0.65. What is the standard deviation of s?

Solution:

a Standard deviation = 0.01 × 28 = **0.28**

b Standard deviation = 0.045 × 15 = **0.675**

c Standard deviation = **2.37**

d Standard deviation = 0.65 × 10 = **6.5**

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Representation and summary of data – measures of dispersion
Exercise E, Question 4

Question:

The coding $y = x - 40$ gives a standard deviation for y of 2.34. Write down the standard deviation of x.

Solution:

Standard deviation = **2.34**

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Representation and summary of data – measures of dispersion
Exercise E, Question 5

Question:

The lifetime, x, in hours, of 70 light bulbs is shown in the table.

Lifetime in hours	Number of light bulbs
20–22	3
22–24	12
24–26	40
26–28	10
28–30	5
total	70

Using the coding $y = \frac{x-1}{20}$ estimate the standard deviation of the actual lifetime in hours of a light bulb.

Solution:

Lifetime in hours	Number of lightbulbs	Midpoint	$y = \frac{x-1}{20}$	fy	fy^2
20 – 22	3	21	1	3	3
22 – 24	12	23	1.1	13.2	14.52
24 – 26	40	25	1.2	48	57.6
26 – 28	10	27	1.3	13	16.9
28 – 30	5	29	1.4	7	9.8
total	70			84.2	101.82

Standard deviation of coded data = $\sqrt{\frac{101.82}{70} - \left(\frac{84.2}{70}\right)^2} = 0.0878$

Standard deviation = 0.0878 × 20 = **1.76**

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Representation and summary of data – measures of dispersion
Exercise E, Question 6

Question:

The weekly income, i, of 100 women workers was recorded.

The data were coded using $y = \frac{i - 90}{100}$ and the following summations were obtained.

$\Sigma y = 131$, $\Sigma y^2 = 176.84$

Work out an estimate for the standard deviation of the actual women workers' weekly income.

Solution:

Standard deviation of coded data $= \sqrt{\frac{176.84}{100} - \left(\frac{131}{100}\right)^2} = 0.229$

Standard deviation $= 0.229 \times 100 =$ **22.9**

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Representation and summary of data – measures of dispersion
Exercise E, Question 7

Question:

A meteorologist collected data on the annual rainfall, x mm, at six randomly selected places. The data were coded using $s = 0.01x - 10$ and the following summations were obtained.

$\Sigma s = 16.1$, $\Sigma s^2 = 147.03$

Work out an estimate for the standard deviation of the actual annual rainfall.

Solution:

Standard deviation of coded data $= \sqrt{\frac{147.03}{6} - \left(\frac{16.1}{6}\right)^2} = 4.16$

Standard deviation $= 4.16 \div 0.01 =$ **416**

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Representation and summary of data – measures of dispersion
Exercise F, Question 1

Question:

For the set of numbers:

2 5 3 8 9 10 3 2 15 10 5 7 6 7 5

a work out the median,

b work out the lower quartile,

c work out the upper quartile,

d work out the interquartile range.

Solution:

2, 2, 3, 3, 5, 5, 5, 6, 7, 7, 8, 9, 10, 10, 15

a Median: $\frac{15}{2} = 7.5$ therefore 8th value needed = **6**

b Q_1: $\frac{15}{4} = 3.75$ therefore 4th value needed = **3**

c Q_3: $3 \times \frac{15}{4} = 11.25$ therefore 12th value needed = **9**

d IQR = 9 − 3 = **6**

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Representation and summary of data – measures of dispersion
Exercise F, Question 2

Question:

A frequency distribution is shown below

Class interval	1–20	21–40	41–60	61–80	81–100
Frequency (*f*)	5	10	15	12	8

Work out the interquartile range.

Solution:

CF = 5 15 30 42 50

Data is continuous so:

$Q_1 : \frac{50}{4} = 12.5^{\text{th}}$ value so Q_1 is in class 21 – 40

$$\frac{Q_1 - 20.5}{40.5 - 20.5} = \frac{12.5 - 5}{15 - 5}$$

$$\frac{Q_1 - 20.5}{20} = \frac{7.5}{10}$$

$$Q_1 = 35.5$$

$Q_3 : 3 \times \frac{50}{4} = 37.5^{\text{th}}$ value so Q_3 is in class 61 – 80

$$\frac{Q_3 - 60.5}{80.5 - 60.5} = \frac{37.5 - 30}{42 - 30}$$

$$Q_3 = 73$$

IQR = 73 − 35.5 = **37.5**

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Representation and summary of data – measures of dispersion
Exercise F, Question 3

Question:

A frequency distribution is shown below.

Class interval	1–10	11–20	21–30	31–40	41–50
Frequency (*f*)	10	20	30	24	16

a Work out the value of the 30th percentile.

b Work out the value of the 70th percentile.

c Calculate the 30% to 70% interpercentile range.

Solution:

a $30^{\text{th}} : \frac{30}{100} \times 100 = 30$

$P_{30} = \mathbf{20.5}$

b $70^{\text{th}} : \frac{70}{100} \times 100 = 70$

$$\frac{P_{70} - 30.5}{40.5 - 30.5} = \frac{70 - 60}{84 - 60}$$

$P_{66} = \mathbf{34.7}$

c 30% to 70% interpercentile = 34.7 − 20.5 = **14.2**

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Representation and summary of data – measures of dispersion
Exercise F, Question 4

Question:

The heights (h) of 10 mountains were recorded to the nearest 10 m and are shown in the table below.

h	1010	1030	1050	1000	1020	1030	1030	1040	1020	1000

Use the coding $y = \frac{h - 1000}{10}$ to find the standard deviation of the heights.

Solution:

coded data 10, 10.2, 10.4, 9.9 10.1, 10.2, 10.2, 10.3, 10.1, 9.9

Standard deviation of coded data = $\sqrt{\frac{1026.41}{10} - \left(\frac{101.3}{10}\right)^2} = 0.155$

Standard deviation = 0.155 × 100 = **15.5**

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Representation and summary of data – measures of dispersion
Exercise F, Question 5

Question:

The times it took a random sample of runners to complete a race are summarised in the table.

Time taken (*t* minutes)	Frequency
20–29	5
30–39	10
40–49	36
50–59	20
60–69	9

a Work out the lower quartile.

b Work out the upper quartile.

c Work out the interquartile range.

The mid-point of each class was represented by x and its corresponding frequency by f giving:

$\Sigma fx = 3740$ $\Sigma fx^2 = 183\,040$

d Estimate the variance and standard deviation for these data.

Solution:

CF = 5 15 51 71 80

a Data is continuous so: $Q_1 : \frac{80}{4} = 20^{\text{th}}$ value so Q_1 is in class 40 – 49

$$\frac{Q_1 - 39.5}{49.5 - 39.5} = \frac{20 - 15}{51 - 15}$$
$$Q_1 = 40.9$$

b $Q_3 : 3 \times \frac{80}{4} = 60^{\text{th}}$ value so Q_3 is in class 50 – 59

$$\frac{Q_3 - 49.5}{59.5 - 49.5} = \frac{60 - 51}{71 - 51}$$
$$Q_3 = 54$$

c IQR = 54 − 40.9 = **13.1**

d Variance $= \frac{183040}{80} - \left(\frac{3740}{80}\right)^2 = \mathbf{102.4375 = 102}$

Standard deviation $= \sqrt{102.4375} = \mathbf{10.1}$

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Representation and summary of data – measures of dispersion
Exercise F, Question 6

Question:

20 birds were caught for ringing.

Their wing spans (x) were measured to the nearest centimetre.

The following summary data was worked out:

$\Sigma x = 316$ $\Sigma x^2 = 5078$

a Work out the mean and the standard deviation of the wing spans of the 20 birds. One more bird was caught. It had a wing span of 13 centimetres

b Without doing any further working say how you think this extra wing span will affect the mean wing span.

Solution:

a Mean $= \frac{316}{20} = \mathbf{15.8}$

Standard deviation $= \sqrt{\frac{5078}{20} - \left(\frac{316}{20}\right)^2} = \mathbf{2.06}$

b It will decrease the mean wing span since 13 < 15.8

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Representation and summary of data – measures of dispersion
Exercise F, Question 7

Question:

The heights of 50 clover flowers are summarised in the table.

Heights in mm (h)	Frequency
90–95	5
95–100	10
100–105	26
105–110	8
110–115	1

a Find Q_1.

b Find Q_2.

c Find the interquartile range.

d Use $\Sigma fx = 5075$ and $\Sigma fx^2 = 516\,112.5$ to find the standard deviation.

Solution:

CF = 5 15 41 49 50

a Data is continuous so:

$Q_1 : \frac{50}{4} = 12.5^{\text{th}}$ value so Q_1 is in class 95 – 100

$$\frac{Q_1 - 95}{100 - 95} = \frac{12.5 - 5}{15 - 5}$$
$$Q_1 = \mathbf{98.75}$$

b $Q_3 : 3 \times \frac{50}{4} = 37.5^{\text{th}}$ value so Q_3 is in class 100 – 105

$$\frac{Q_3 - 100}{105 - 100} = \frac{37.5 - 15}{41 - 15}$$
$$Q_3 = \mathbf{104.33}$$

c IQR = 104.33 − 98.75 = **5.58**

d Standard deviation $= \sqrt{\frac{516112.5}{50} - \left(\frac{5075}{50}\right)^2} = \mathbf{4.47}$

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Representation of data
Exercise A, Question 1

Question:

A group of thirty college students was asked how many DVDs they had in their collection. The results are as follows.

12	25	34	17	12	18	29	34	45	6
15	9	25	23	29	22	20	32	15	15
19	12	26	27	27	32	35	42	26	25

Draw a stem and leaf diagram to represent these data.

a Find the median.

b Find the lower quartile.

c Find the upper quartile.

Solution:

Unordered

key 1|2 means 12

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

Ordered

key 1|2 means 12

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

(a) $\frac{30}{2} = 15$ therefore 15.5^{th} term = 25

(b) $\frac{30}{4} = 7.5$ therefore 8^{th} term = 15

(c) $\frac{3(30)}{4} = 22.5$ therefore 23^{rd} term = 29

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Representation of data
Exercise A, Question 2

Question:

The following stem and leaf diagram shows some information about the marks gained by a group of students in a statistics test.

stem	leaf									Key: 2|3 means 23 marks
0	8	9								(2)
1	2	5	5	9						(4)
2	3	6	6	6	7					(5)
3	4	4	5	7	7	7	7	7	9	(9)
4	5	8	8	9						(4)

a Work out how many students there were in the group.

b Write down the highest mark.

c Write down the lowest mark.

d Write down how many students got 26 marks.

e Write down the modal mark.

f Find the median mark.

g Find the lower quartile.

h Find the upper quartile.

Solution:

(a) 24 (b) 49 (c) 8 (d) 3 (e) 37

(f) $\frac{24}{2} = 12$ therefore 12.5$^{\text{th}}$ term = 34

(g) $\frac{24}{4} = 6$ therefore 6.5$^{\text{th}}$ term $= \frac{19+23}{2} = 21$

(h) $\frac{3(24)}{4} = 18$ therefore 18.5$^{\text{th}}$ term = 37

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Representation of data
Exercise A, Question 3

Question:

3 The number of laptops sold by a store was recorded each month for a period of 26 months. The results are shown in the stem and leaf diagram.

stem	leaf									Key: 1|8 means 18 laptops
1	8									(1)
2	3	6	7	9	9					(5)
3	2	6	6	6	7	8	8			(7)
4	4	5	5	5	7	7	7	7	9	(9)
5	2	7	7	9						(4)

a Find the median.

b Find the lower quartile.

c Find the upper quartile.

d Work out the interquartile range.

e Write down the modal number of laptops sold.

Solution:

(a) $\frac{26}{2} = 13$ therefore 13.5^{th} term $= \frac{38 + 44}{2} = 41$

(b) $\frac{26}{4} = 6.5 = 7^{\text{th}}$ term $= 32$

(c) $\frac{3(26)}{4} = 19.5 = 20^{\text{st}}$ term $= 47$

(d) IQR $= 47 - 32 = 15$

(e) 47

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Representation of data
Exercise A, Question 4

Question:

A class of 16 boys and 13 girls did a Physics test. The test was marked out of 60. Their marks are shown below.

Boys				Girls			
45	54	32	60	26	54	47	32
28	34	54	56	34	34	45	46
32	29	47	48	39	52	24	28
44	45	56	57	33			

a Draw a back-to-back stem and leaf diagram to represent these data.

b Comment on your results.

Solution:

(a) Unordered

Boys						Girls Key 2|6 means 26				
			9	8	2	6	4	8		
		2	4	2	3	2	4	4	9	3
5	4	8	7	5	4	7	5	6		
7	6	6	4	4	5	4	2			
				0	6					

Ordered

Boys						Girls Key 2|6 means 26				
			9	8	2	4	6	8		
		4	2	2	3	2	3	4	4	9
8	7	5	5	4	4	5	6	7		
7	6	6	4	4	5	2	4			
				0	6					

(b) Girls gain lower marks than boys

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Representation of data
Exercise A, Question 5

Question:

The following stem and leaf diagram shows the weekend earnings of a group of college students.

Males		Females
8	0	6
7 6 5	1	0 5 5 5 8 8
9 9 9 8 6 6	2	5 5 8 8 9
8 8 5 5 5	3	5 5
8 5	4	0

Key: 5|1|0 means £15 for males and £10 for females

a Write down the number of male students and the number of female students.

b Write down the largest amount of money earned by the males.

c Comment on whether males or females earned the most in general.

Solution:

(a) 17 males and 15 females

(b) £48

(c) Males earned the most in general.

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Representation of data
Exercise B, Question 1

Question:

Some data are collected. The lower quartile is 46 and the upper quartile is 68.

An outlier is an observation that falls either 1.5 × (interquartile range) above the upper quartile or 1.5 × (interquartile range) below the lower quartile.

Work out whether the following are outliers using this rule.

a 7

b 88

c 105

Solution:

IQR = 68−46 = 22

46 − 1½ × 22 = 13

68 + 1½ × 22 = 101

(a) 7 is an outlier (b) 88 is not an outlier (c) 105 is an outlier.

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Representation of data
Exercise B, Question 2

Question:

Male and female turtles were weighed in grams. For males, the lower quartile was 400 g and the upper quartile was 580 g. For females, the lower quartile was 260 g and the upper quartile was 340 g.

An outlier is an observation that falls either 1 × (interquartile range) above the upper quartile or 1 × (interquartile range) below the lower quartile.

a Which of these male turtle weights would be outliers?

400 g   260 g   550 g   640 g

b Which of these female turtle weights would be outliers?

170 g   300 g   340 g   440 g

c What is the largest size a male turtle can be without being an outlier?

Solution:

a) no outliers

b) 170 g and 440 g are outliers

c) 760 g

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Representation of data
Exercise C, Question 1

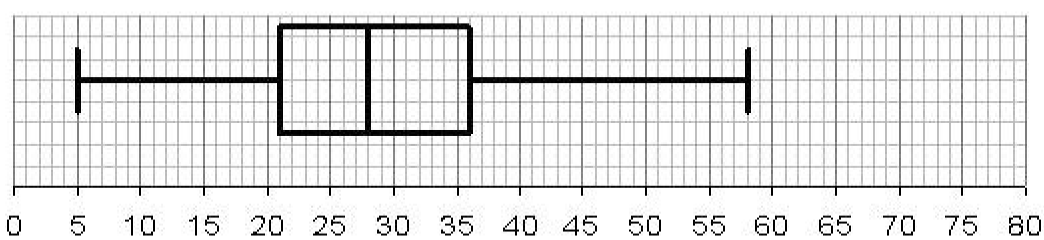
Question:

A group of students did a test. The summary data is shown in the table below.

Lowest value	Lower quartile	Median	Upper quartile	Highest value
5	21	28	36	58

Given that there were no outliers draw a box plot to illustrate these data.

Solution:

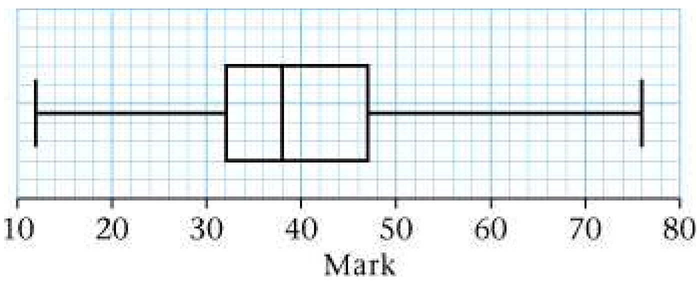


Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Representation of data
Exercise C, Question 2

Question:

Here is a box plot of marks in an examination.



a Write down the upper and lower quartiles.

b Write down the median.

c Work out the interquartile range.

d Work out the range.

Solution:

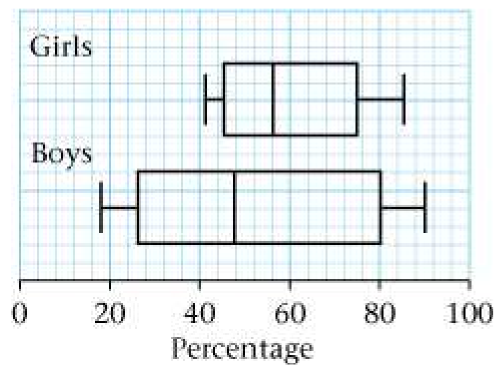
(a) 47 and 32 (b) 38 (c) IQR = 47 − 32 = 15 (d) Range = 76 − 12 = 64

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Representation of data
Exercise D, Question 1

Question:

A group of students took a statistics test. The summary data for the percentage mark gained by boys and by girls is shown in the box plots opposite.



a Write down the percentage mark which 75% of the girls scored more than.

b State the name given to this value.

c Compare and contrast the results of the boys and the girls.

Solution:

(a) 45

(b) Lower quartile

(c) Boys have a lower median and bigger IQR/range (or girls have a higher median and lower IQR/range)

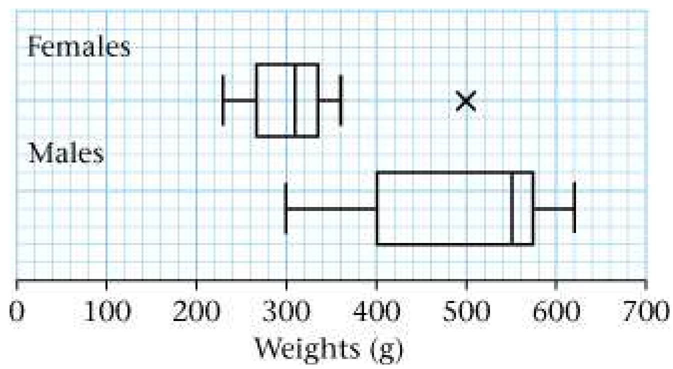
The person with the highest mark was a boy. The person with the lowest mark was a boy.

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Representation of data
Exercise D, Question 2

Question:

Male and female turtles were weighed in grams. Their weights are summarised in the box plots opposite.



a Compare and contrast the weights of the male and female turtles.

b A turtle was found to be 330 grams in weight. State whether it is likely to be a male or a female. Give a reason for your answer.

c Write down the size of the largest female turtle.

Solution:

(a) Male turtles have a higher median weight (or Females have a lower).

Males have a bigger range (or IQR) (females have a lower range IQR).

(b) It is more likely to be a female. Hardly any male turtles would weigh this little, but more than a quarter of female turtles would weigh more than this.

(c) 500 g

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Representation of data
Exercise E, Question 1

Question:

The heights of a year group of children were measured. The data are summarised in the group frequency table.

Height (cm)	Frequency	Class width	Frequency density
135–144	40	10	
145–149	40	5	
150–154	75		
155–159	65		
160–174	60		

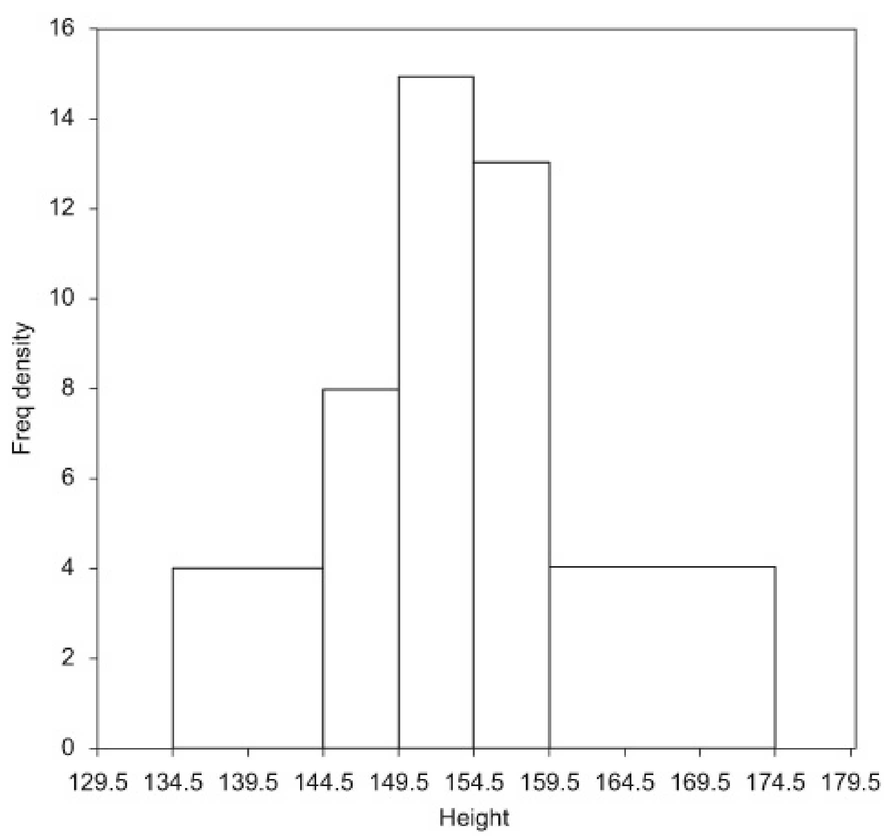
a Copy and complete the table.

b Draw a histogram for these data.

Solution:

a 5, 5, 15; 4, 8, 15, 13, 4

b



Solutionbank S1

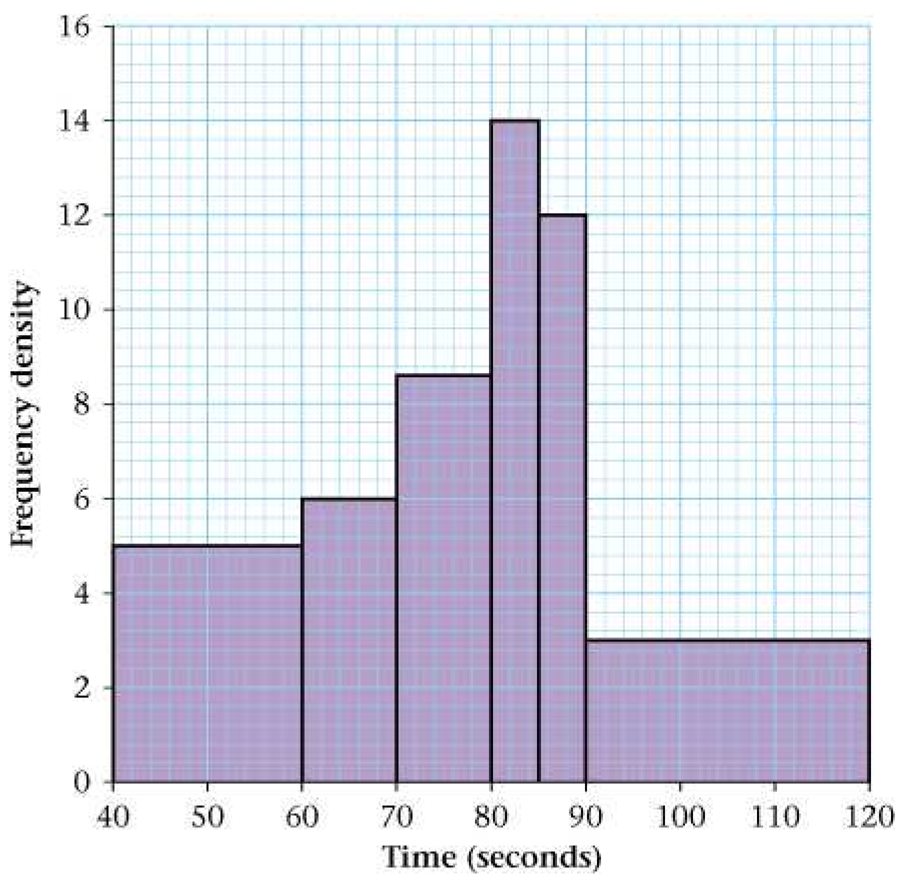
Edexcel AS and A Level Modular Mathematics

Representation of data

Exercise E, Question 2

Question:

Some students take part in an obstacle race. The time it took each student to complete the race was noted. The results are shown in the histogram.



a Give a reason to justify the use of a histogram to represent these data.

The number of students who took between 60 and 70 seconds is 90.

b Find the number of students who took between 40 and 60 seconds.

c Find the number of students who took 80 seconds or less.

d Calculate the total number of students who took part in the race.

Solution:

a Time is continuous data

b Area of 60 − 70 seconds bar is $10 \times 6 = 60$ units squared

1 unit squared = 90/60 = 1.5 students

Area of 40 − 60 seconds bar is $20 \times 5 = 100$ units squared

Number of students = 100 × 1.5 = **150**

c Area for 80 seconds or less = 20 × 5 + 10 × 6 + 10 × 8.6 = 246 units squared.

Number of students = 246 × 1.5 = **369**

d Total Area = 246 + 5 × 14 + 5 × 12 + 30 × 3 = 466 units squared.

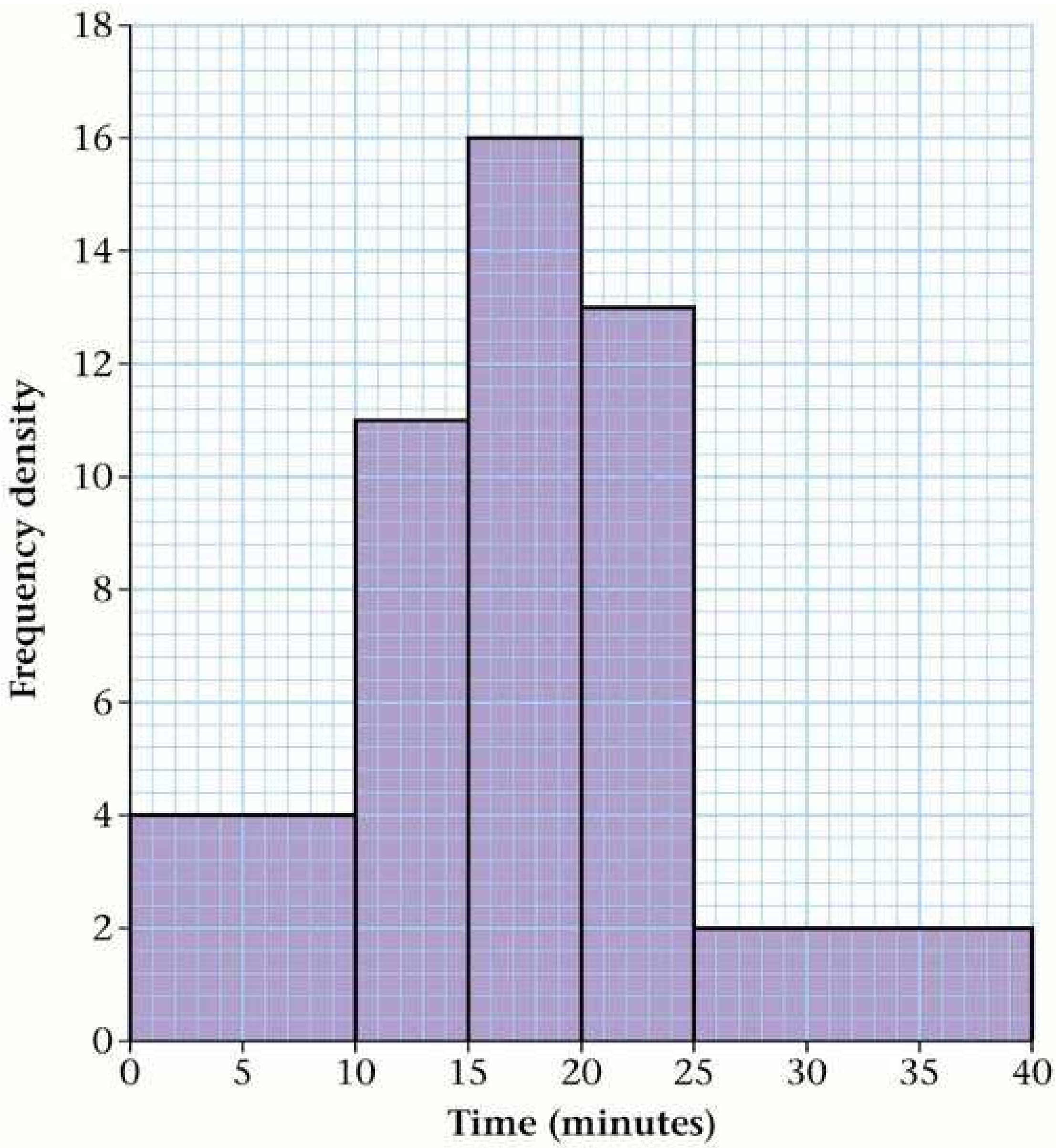
Number of employees = 466 × 1.5 = **699**

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Representation of data
Exercise E, Question 3

Question:

The time taken for each employee in a company to travel to work was recorded. The results are shown in the histogram.



The number of employees who took less than 10 minutes to travel to work is 48.

a Find how many employees took less than 15 minutes to travel to work.

b Estimate how many employees took between 20 and 30 minutes to travel to work.

c Estimate how many employees took more than 30 minutes to travel to work.

Solution:

a Area less than 10 minutes is 10 × 4 = 40 units squared

1 unit squared = 48 ÷ 40 = 1.2 students

Area less than 15 minutes is 40 + 5 × 11 = 95 units squared

Number of employees = 95 × 1.2 = **114**

b Area of 20 − 30 minutes bars is 5 × 13 + 5 × 2 = 75 units squared

Number of employees = 75 × 1.2 = **90**

c Area for more than 30 minutes seconds bars is 10 × 2 = 20 units squared

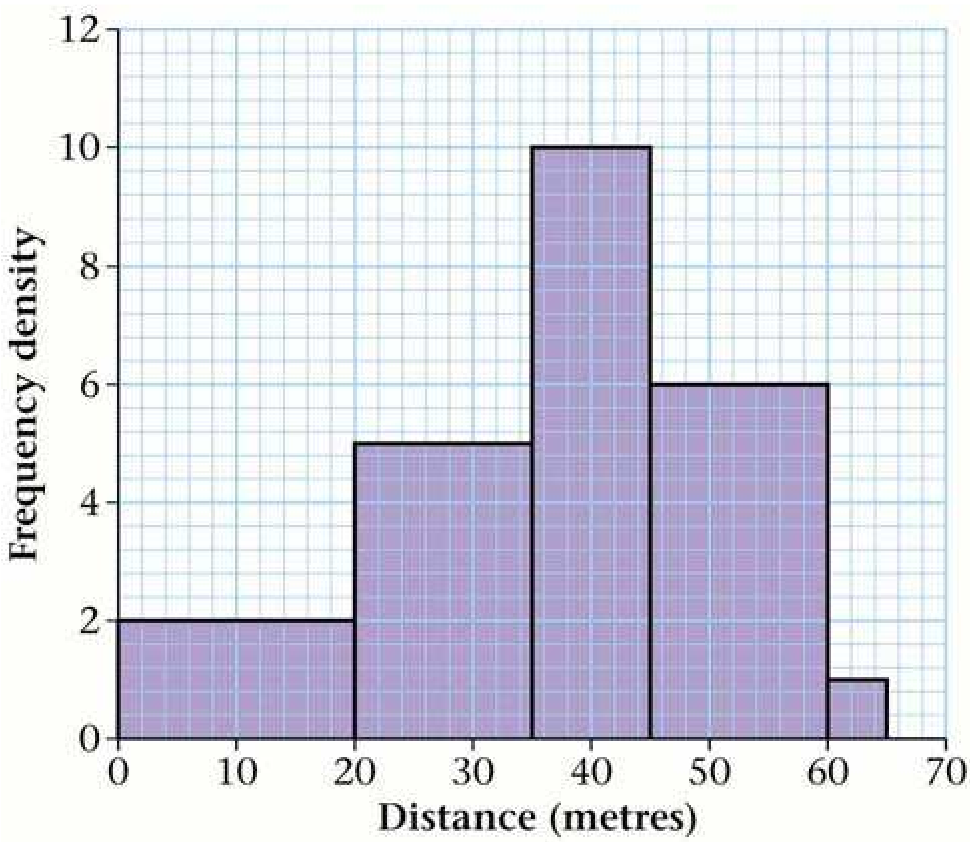
Number of employees = 20 × 1.2 = **24**

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Representation of data
Exercise E, Question 4

Question:

A Fun Day committee at a local sports centre organised a throwing the cricket ball competition. The distance thrown by every competitor was recorded. The data were collected and are shown in the histogram. The number of competitors who threw less than 10 m was 40.



a Why is a histogram a suitable diagram to represent these data?

b How many people entered the competition?

c Estimate how many people threw between 30 and 40 metres.

d How many people threw between 45 and 65 metres?

e Estimate how many people threw less than 25 metres.

Solution:

a Distance is continuous data.

b Area for less than 10m is $10 \times 2 = 20$ units squared

1 unit squared $= 40 \div 20 = 2$ people

Total Area $10 \times 2 + 10 \times 2 + 15 \times 5 + 10 \times 10 + 15 \times 6 + 5 \times 1 = 310$ units squared

Number of people = 310 × 2 = **620**

c Area for 30 − 40 m is 5 × 5 + 5 × 10 = 75 units squared.

Number of people = 75 × 2 = **150**

d Area for 45 − 65 m is 15 × 6 + 5 × 1 = 95 units squared.

Number of people = 95 × 2 = **190**

e Area for less than 25 m is 10 × 2 + 10 × 2 + 5 × 5 = 65 units squared.

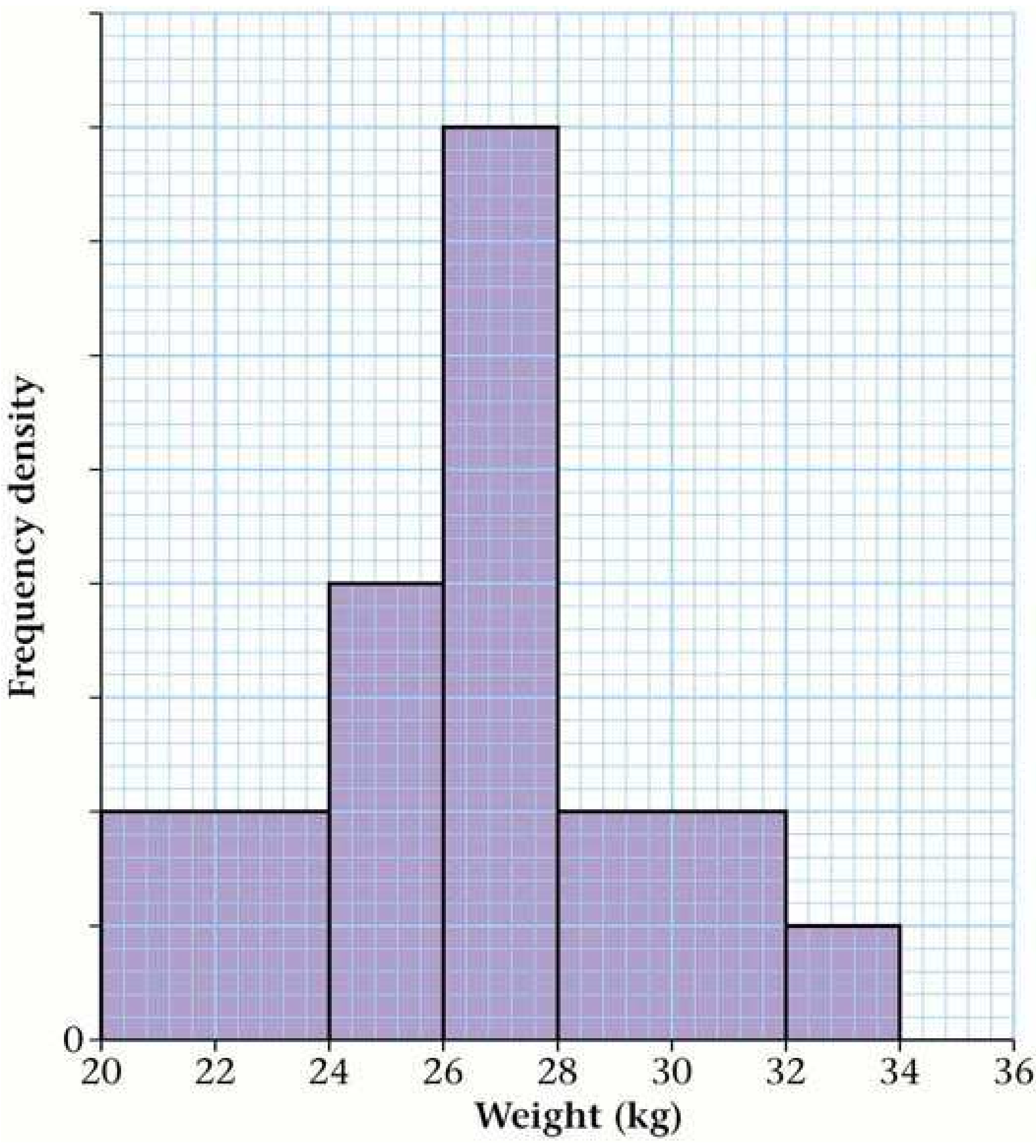
Number of people = 65 × 2 = **130**

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Representation of data
Exercise E, Question 5

Question:

A farmer weighed a random sample of pigs. The weights were summarised in a grouped frequency table and represented by a histogram.



One of the classes in the grouped frequency distribution was 28–32 and its associated frequency was 32. On the histogram the height of the rectangle representing that class was 2 cm and the width was 2 cm.

a Give a reason to justify the use of a histogram to represent these data.

b Write down the underlying feature associated with each of the bars in a histogram.

c Show that on this histogram each pig was represented by 0.125 cm^2.

d How many pigs did the farmer weigh altogether?

e Estimate the number of pigs that weighed between 25 and 29 kg.

Solution:

a Weight is continuous data.

b The area of the bar is proportional to the frequency.

c Area for 28 – 32 kg is 2 × 2 = 4 units squared

Area for 1 pig is 4 ÷ 32 = 0.125

d Total area in cm^2 is 2 × 2 + 1 × 4 + 1 × 8 + 2 × 2 + 1 × 1 = 21 units squared.

Number of pigs = 21 ÷ 0.125 = **168**

e Area for 25 − 29 m is ½ × 4 + 1 × 8 + ½ × 2 = 11 units squared.

Number of pigs = 11 ÷ 0.125 = **88**

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Representation of data
Exercise F, Question 1

Question:

In a survey of the earnings of some sixth form students who did Saturday jobs the median wage was £36.50. The 75th percentile was £45.75 and the interquartile range was £30.50.

Use the quartiles to describe the skewness of the distribution.

Solution:

1. $Q_1 = 45.75 - 30.5 = 15.25$

$Q_3 - Q_2 = 45.75 - 36.5 = 9.25$

$Q_2 - Q_1 = 36.5 - 15.25 = 21.25$

$Q_2 - Q_1 > Q_3 - Q_2$ therefore it is negatively skewed

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Representation of data
Exercise F, Question 2

Question:

A group of estate agents recorded the time spent on the first meeting with a random sample of 120 of their clients. The times, to the nearest minute, are summarised in the table.

Time	Number of clients
10–15	2
15–20	5
20–25	17
25–30	38
30–35	29
35–45	25
45–80	4
Total	120

a Calculate estimates of the mean and variance of the times.

b By interpolation obtain estimates of the median and quartiles of the times spent with customers.

One measure of skewness is found using $\frac{3(\text{mean} - \text{median})}{\text{standard deviation}}$.

c Evaluate this measure and describe the skewness of these data.

The estate agents are undecided whether to use the median and quartiles, or the mean and standard deviation to summarise these data.

d State, giving a reason, which you would recommend them to use.

Solution:

(a) Mean $= \frac{\sum fx}{n} = \frac{12.5 \times 2 + 17.5 \times 5 + 22.5 \times 17 + 27.5 \times 38 + 32.5 \times 29 + 40 \times 25 + 62.5 \times 4}{120}$

$= 31.104166 = 31.1$ minutes

Variance $= \frac{\sum fx^2}{\sum f} - \mu^2 = \frac{12.5^2 \times 2 + 17.5^2 \times 5 + 22.5^2 \times 17 + 27.5^2 \times 38 + 32.5^2 \times 29 + 40^2 \times 25 + 62.5^2 \times 4}{120} - 31.1^2$
$= 1045.36 - 967.31 = 78.05$

Standard deviation $= \sqrt{78.05} = 8.835$

(b) Median is the $\frac{120}{2} = 60$ th value so Q_2 is in class $25 - 30$

$$\frac{Q_2 - 25}{30 - 25} = \frac{60 - 24}{62 - 24}$$
Q_2 $=$ **29.7 minutes**

$Q_1 : \frac{120}{4} = 30^{\text{th}}$ value so Q_1 is in class 25 − 30

$$\frac{Q_1 - 25}{30 - 25} = \frac{30 - 24}{62 - 24}$$
$$Q_1 = \mathbf{25.8 \ minutes}$$

$Q_3 : 3 \times \frac{120}{4} = 90^{\text{th}}$ value so Q_3 is in class 30−35

$$\frac{Q_3 - 30}{35 - 30} = \frac{90 - 62}{91 - 62}$$
$$Q_3 = \mathbf{34.8 \ minutes}$$

(c) Skew $= \frac{3(31.10 - 29.74)}{8.835} = 0.46$

Positive skew.

(d) Median and quartiles because of the skew.

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Representation of data
Exercise F, Question 3

Question:

The following stem and leaf diagram summarises the wing length, to the nearest mm, of a random sample of 67 owl moths.

	Wing length											Key:5|0 means 50
5	0	0	0	1	1	2	2	3	3	3	4	4 (12)
5	5	5	6	6	6	7	8	8	9	9		(10)
6	0	1	1	1	3	3	4	4	4	4		(10)
6	5	5	6	7	8	9	9					(7)
7	1	1	2	2	3	3						(6)
7	5	7	9	9								(4)
8	1	1	1	2	2	3	3	4				(8)
8	7	8	9									(3)
9	0	1	1	2								(4)
9	5	7	9									(3)

a Write down the mode of these data.

b Find the median and quartiles of these data.

c On graph paper, construct a box plot to represent these data.

d Comment on the skewness of the distribution.

e Calculate the mean and standard deviation of these data.

f Use a further method to show that these data are skewed.

g State, giving a reason, which of **b** or **e** you would recommend using to summarise the data in the table.

Solution:

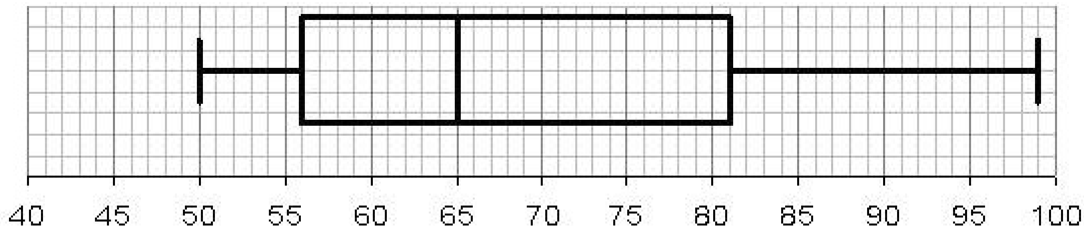
(a) 64

(b) $Q_2 = \frac{67}{2} = 33.5$ therefore 34^{th} term = 65

$Q_1 = \frac{67}{4} = 16.75 = 17^{th}$ term = 56

$Q_3 = 16.75 \times 3 = 50.25 = 51^{st}$ term = 81

(c)



(d) Positive skew

(e) mean 68.72 sd 13.73

(f) $Q_2 - Q_1 = 65 - 56 = 9$ and $Q_3 - Q_2 = 81 - 65 = 16$

$Q_2 - Q_1 < Q_3 - Q_2 \Rightarrow$ Positive skew

$\frac{3(Q_3 - Q_1)}{Q_2} = \frac{3(81 - 56)}{65} = 1.15 \Rightarrow$ Positive skew

(g) (b) because of the skew.

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Representation of data
Exercise F, Question 4

Question:

A TV company wishes to appeal to a wider range of viewers. They decide to purchase a programme from another channel. They have the option of buying one of two programmes. The company collects information from a sample of viewers for each programme. The results are summarised in the table. State which programme the company should buy to increase the range of their viewers. Give a reason for your answer.

	Mean age	**Standard deviation of age**
Programme 1	50	5
Programme 2	50	10

Solution:

Program 2 because it has a bigger standard deviation and hence bigger range in the age groups watching it.

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Representation of data
Exercise G, Question 1

Question:

Jason and Perdita decided to go for a touring holiday on the continent for the whole of July. They recorded the number of kilometres they travelled each day. The data are summarised in the stem and leaf diagram below.

Key: 15|5 means 155 kilometres

stem	leaf								
15	5								
16	4	8	9						
17	3	5	7	8	8	8	9	9	9
18	4	4	5	5	8				
19	2	3	4	5	5	6			
20	4	7	8	9					
21	1	2							
22	6								

a Find Q_1, Q_2, and Q_3

Outliers are values that lie outside $Q_1 - 1.5(Q_3 - Q_1)$ and $Q_3 + 1.5(Q_3 - Q_1)$.

b Find any outliers.

c Draw a box plot of these data.

d Comment on the skewness of the distribution.

Solution:

(a) $Q_2 = \frac{32}{2} = 16$ therefore 16.5^{th} term = 185

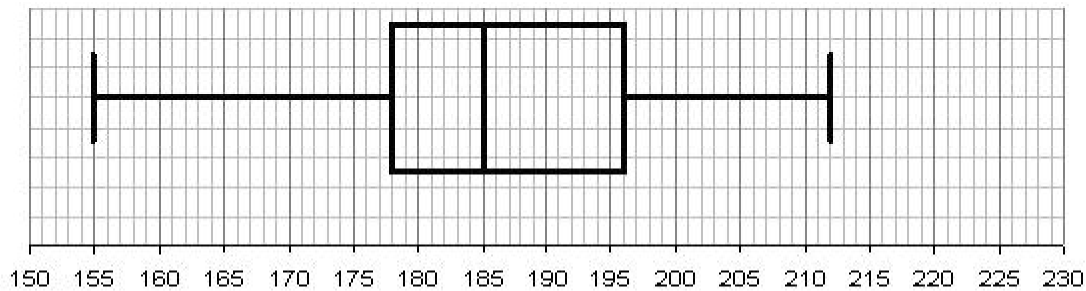
$Q_1 = \frac{31}{4} = 7\frac{3}{4} \Rightarrow 8^{\text{th}}$ term = 178

$Q_3 = \frac{3(31)}{4} = 23\frac{1}{4} \Rightarrow 24^{\text{th}}$ term = 196

(b) $Q_1 - 1.5(Q_3 - Q_1) = 178 - 1.5(196 - 178) = 151$

$Q_3 + 1.5(Q_3 - Q_1) = 196 + 1.5(196 - 178) = 223$ Outliers are 226.

(c)



(d) Positive skew

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Representation of data
Exercise G, Question 2

Question:

Sophie and Jack do a survey every day for three weeks. Sophie counts the number of pedal cycles using Market Street. Jack counts the number of pedal cycles using Strand Road. The data they collected are summarised in the back-to-back stem and leaf diagram.

Sophie	Stem	Jack
9 9 7 5	0	6 6
7 6 5 3 3 2 2 2 1 1	1	1 1 5
5 3 3 2 2	2	1 2 2 2 3 7 7 8 9
2 1	3	2 3 4 7 7 8
	4	2

Key: 5|0|6 means Sophie counts 5 cycles and Jack counts 6 cycles

a Write down the modal number of pedal cycles using Strand Road.

The quartiles for these data are summarised in the table below.

	Sophie	**Jack**
Lower quartile	X	21
Median	13	Y
Upper quartile	Z	33

b Find the values for X, Y and Z.

c Write down the road you think has the most pedal cycles travelling along it overall. Give a reason for your answer.

Solution:

(a) 22

(b) $X = \frac{21}{4} = 5\frac{1}{4} \Rightarrow 6^{\text{th}}$ item = 11

$Z = \frac{3(21)}{4} = 15\frac{3}{4} \Rightarrow 16^{\text{th}}$ item = 22

$Y = \frac{21}{2} = 10.5$ therefore 11^{th} item = 27

(c) Strand road has the highest median.

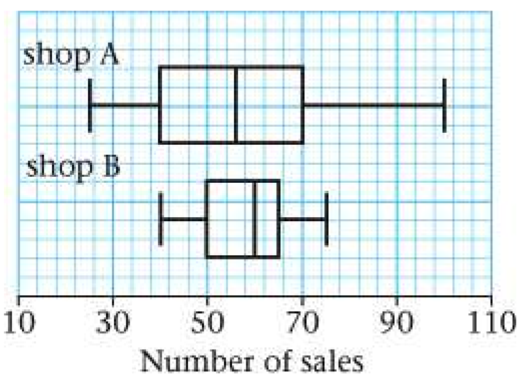
Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Representation of data
Exercise G, Question 3

Question:

Shop A and Shop B both sell mobile phones.

They recorded how many they sold each day over a long period of time. The data they collected are represented in the box plots.



a Shop B says that for 50% of the days they sold 60 or more phones a day. State whether or not this is a true statement. Give a reason for your answer.

b Shop A says that for 75% of the days they sold 40 or more phones a day. State whether or not this is a true statement. Give a reason for your answer.

c Compare and contrast the two box plots.

d Write down the shop you think had the most consistent sales per day. Explain the reason for your choice.

Solution:

(a) This is a true statement. The median is 60 phones a day so for half the days they sold 60 or more.

(b) True. The lower quartile is 40.

(c) Shop A has a lower median and a bigger IQR/range. Overall shop A sells less phones but the daily quantity sold is more variable. Shop A had both the highest and the lowest number of phones sold in a day.

(d) Shop B as the interquartile range smaller.

Solutionbank S1

Edexcel AS and A Level Modular Mathematics

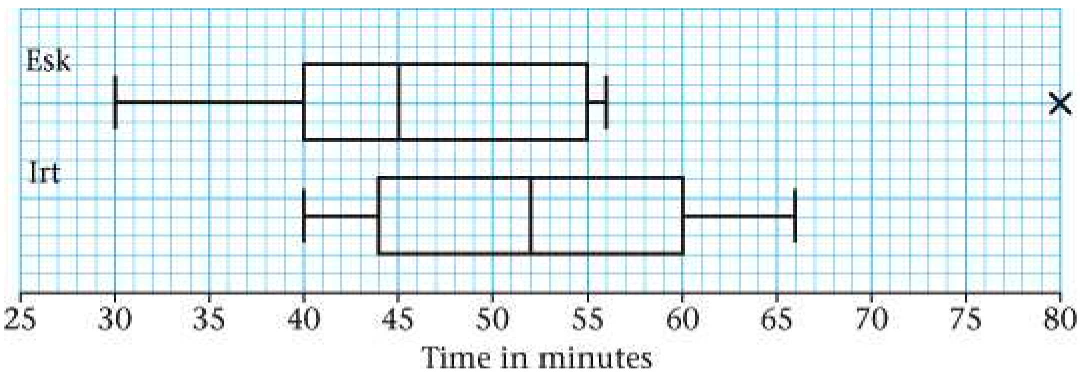
Representation of data

Exercise G, Question 4

Question:

4 Fell runners from the Esk Club and the Irt Club were keen to see which club had the fastest runners overall. They decided that all the members from both clubs would take part in a fell run. The time each runner took to complete the run was recorded.

The results are summarised in the box plot.



a Write down the time by which 50% of the Esk Club runners had completed the run.

b Write down the time by which 75% of the Irt Club runners had completed the run.

c Explain what is meant by the cross (×) on the Esk Club box plot.

d Compare and contrast these two box plots.

e Comment on the skewness of the two box plots.

f What conclusions can you draw from this information about which club has the fastest runners?

Solution:

(a) 45 minutes

(b) 60 minutes

(c) This is an outlier that does not fit the pattern.

(d) The Irt club had the highest median so overall they had the slowest runners.

The IQR ranges were about the same.

(e) Esk club times are positively skewed. Irt club times are symmetric.

(f) Esk had the fastest runners because they had the lower times.

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

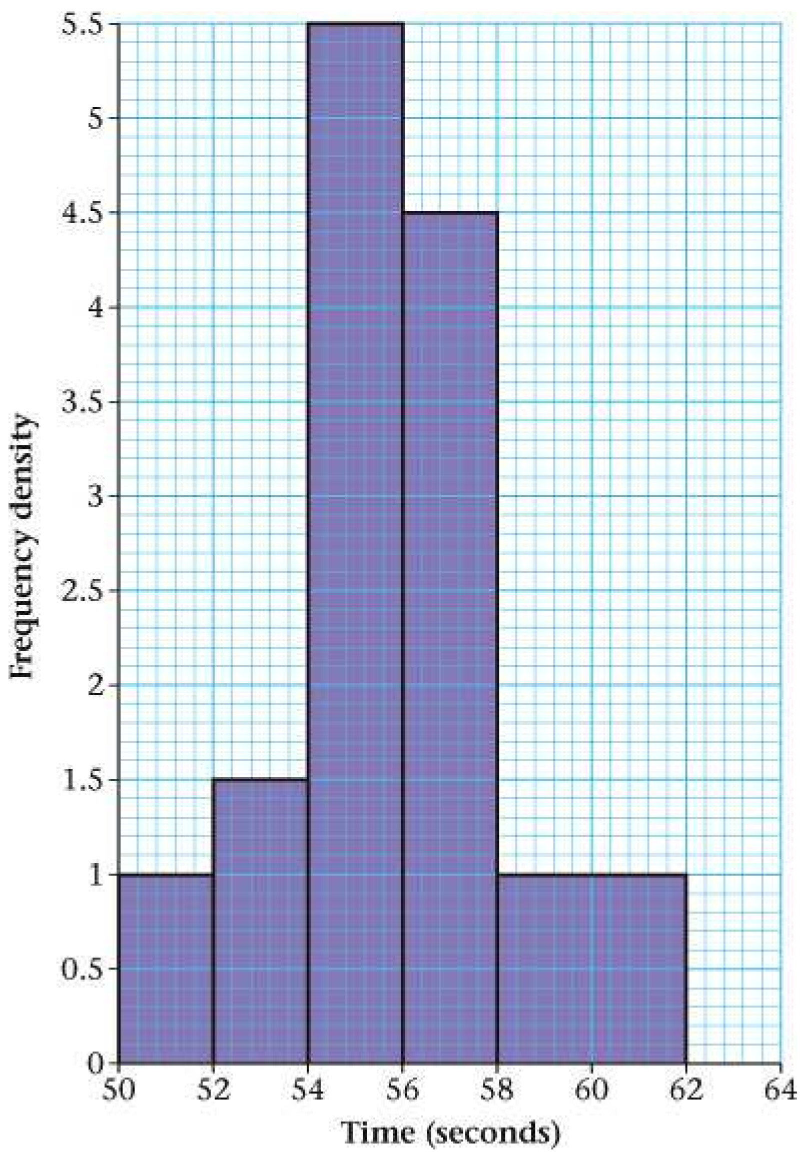
Representation of data
Exercise G, Question 5

Question:

The histogram shows the time taken by a group of 58 girls to run a measured distance.

a Work out the number of girls who took longer than 56 seconds.

b Estimate the number of girls who took between 52 and 55 seconds.



Solution:

(a) Area = frequency

2k(1 + 1.5 + 5.5 + 4.5) + 4k(1) = 58

29k = 58 so k = 2

Number of girls who took longer than 56 seconds = 2 $\{(4.5 \times 2) + (1 \times 4)\}$ = 26 girls

(b) Number of girls between 52 and 55 seconds = $2\{(1.5 \times 2) + (1 \times 5.5)\}$ = 17 girls

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Representation of data
Exercise G, Question 6

Question:

The table gives the distances travelled to school, in km, of the population of children in a particular region of the United Kingdom.

Distance, km	0–1	1–2	2–3	3–5	5–10	10 and over
Number	2565	1784	1170	756	630	135

A histogram of this data was drawn with distance along the horizontal axis. A bar of horizontal width 1.5 cm and height 5.7 cm represented the 0–1 km group.

Find the widths and heights, in cm to one decimal place, of the bars representing the following groups:

a 2–3,

b 5–10.

Solution:

$1.5 \times 5.7 \times k = 2565$, so $k = 300$

a width = 1.5 cm height = $\frac{1170}{2565} \times 5.7 = 2.6$ cm

b width = $5 \times 1.5 = 7.5$ cm height = $\frac{630}{300 \times 7.5} = 0.28$ cm

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Representation of data
Exercise G, Question 7

Question:

The labelling on bags of garden compost indicates that the bags weigh 20 kg.

The weights of a random sample of 50 bags are summarised in the table opposite.

Weight in kg	Frequency
14.6–14.8	1
14.8–18.0	0
18.0–18.5	5
18.5–20.0	6
20.0–20.2	22
20.2–20.4	15
20.4–21.0	1

a On graph paper, draw a histogram of these data.

b Estimate the mean and standard deviation of the weight of a bag of compost.

[You may use $\Sigma fy = 988.85$, $\Sigma fy^2 = 19\ 602.84$]

c Using linear interpolation, estimate the median.

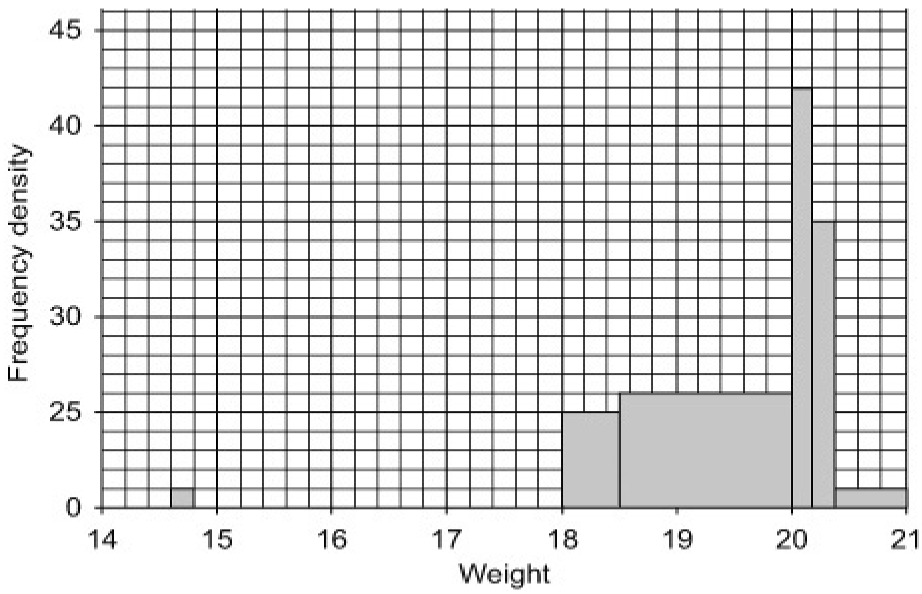
One coefficient of skewness is given by

$$\frac{3(\text{mean} - \text{median})}{\text{standard deviation}}.$$

d Evaluate this coefficient for the above data.

e Comment on the skewness of the distribution of the weights of bags of compost.

Solution:



(b) Mean $= \frac{\sum fy}{n} = \frac{988.85}{50} = 19.777$

sd $= \sqrt{\frac{\sum fy^2}{n} - \mu^2} = \sqrt{\frac{19602.84}{50} - 19.777^2} = \sqrt{0.927} = \sqrt{0.927} = 0.963$

(c) median $= 20 + \frac{13}{22} \times 0.2 = 20.118181818 = 20.118$

(d) $\frac{3(19.777 - 20.118)}{0.963} = -1.0623 = -1.06$

(e) The distribution of the weights of bags of compost is negatively skewed.

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Representation of data
Exercise G, Question 8

Question:

The number of bags of potato crisps sold per day in a bar was recorded over a two-week period. The results are shown below.

20 15 10 30 33 40 5 11 13 20 25 42 31 17

a Calculate the mean of these data.

b Draw a stem and leaf diagram to represent these data.

c Find the median and the quartiles of these data.

An outlier is an observation that falls either 1.5 × (interquartile range) above the upper quartile or 1.5 × (interquartile range) below the lower quartile.

d Determine whether or not any items of data are outliers.

e On graph paper draw a box plot to represent these data. Show your scale clearly.

f Comment on the skewness of the distribution of bags of crisps sold per day. Justify your answer.

Solution:

(a) 22.285714 bags

(b)

Number of bags

0	5				
1	0	1	3	5	7
2	0	0	5		
3	0	1	3		
4	0	2			

key 1|1 = 11 bags

(c) Median = $\frac{15}{2} = 7.5$ therefore 8^{th} item = 20

$Q_1 = \frac{14}{4} = 3.5 \Rightarrow 4^{th}$ item = 13

$Q_3 = \frac{3(14)}{4} = 10.5 \Rightarrow 11^{th}$ item = 31

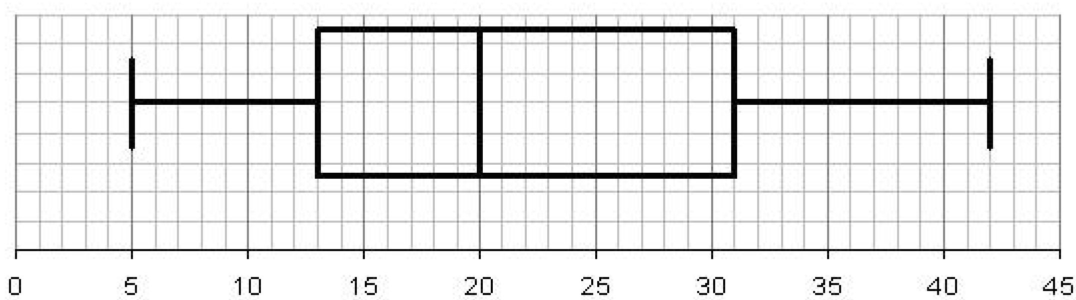
(d) IQR = 31 − 13 = 18 so 1.5 × IQR = 27

13 − 27 = −14

31 + 27 = 58

There are no outliers

(e)



(f) Positive skew. $Q_2 - Q_1 < Q_3 - Q_2$

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

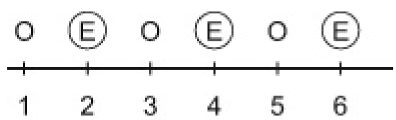
Probability
Exercise A, Question 1

Question:

For each of the following experiments, identify the sample space and find the probability of the event specified.

Throwing a six sided die once and recording if the number face up is odd or even. Find the probability of an even number landing face up.

Solution:



O = odd, E = Even Score on die

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

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

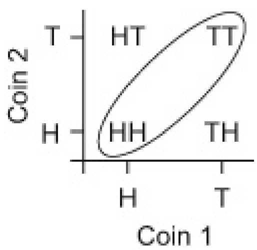
Probability
Exercise A, Question 2

Question:

For the following experiment, identify the sample space and find the probability of the event specified.

Tossing two coins. Find the probability of the same outcome on each coin.

Solution:



P(Same outcome) = $\frac{2}{4} = \frac{1}{2}$

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Probability
Exercise A, Question 3

Question:

For the following experiment, identify the sample space and find the probability of the event specified.

A card is drawn from a pack of 52 playing cards. Find the probability that the card is a heart.

Solution:

S = {Spades, Clubs, Diamonds, Hearts}

P(Heart) = $\frac{1}{4}$

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

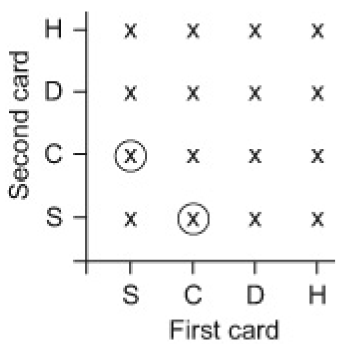
Probability
Exercise A, Question 4

Question:

For the following experiment, identify the sample space and find the probability of the event specified.

A card is drawn from a pack of 52 playing cards, its suit is recorded then it is replaced and another card is drawn. Find the probability of drawing a spade and a club in any order.

Solution:



$$\text{P(Spade } AND \text{ Club)} = \frac{2}{16} = \frac{1}{8}$$

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

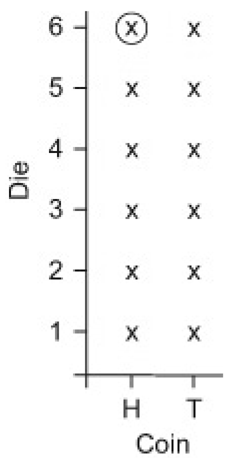
Probability
Exercise A, Question 5

Question:

For the following experiment, identify the sample space and find the probability of the event specified.

Throwing a die and tossing a coin. Find the probability of a head and a 6.

Solution:



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

Solutionbank S1

Edexcel AS and A Level Modular Mathematics

Probability

Exercise A, Question 6

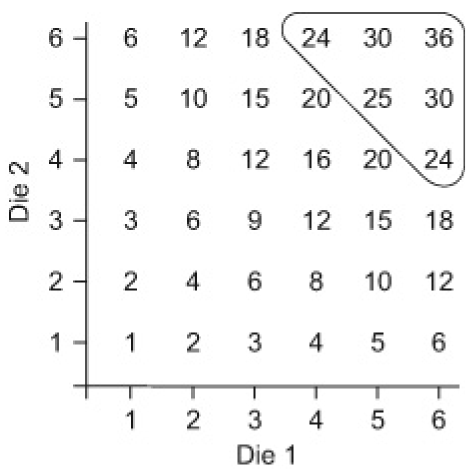
Question:

For each of the following experiments, identify the sample space and find the probability of the event specified.

Throwing two six-sided dice and recording the product of the values on the two sides that are uppermost.

Find the probability of the answer being greater than or equal to 24.

Solution:



$$P(\text{Product} \geq 24) = \frac{6}{36} = \frac{1}{6}.$$

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Probability
Exercise B, Question 1

Question:

A card is chosen at random from a pack of 52 playing cards. C is the event 'the card chosen is a club' and K is the event 'the card chosen is a King'. Find these.

a $\mathrm{P}(K)$ **b** $\mathrm{P}(C)$ **c** $\mathrm{P}(C \cap K)$
d $\mathrm{P}(C \cup K)$ **e** $\mathrm{P}(C')$ **f** $\mathrm{P}(K' \cap C)$

Solution:

a. $\mathrm{P}(K) = \frac{4}{52} = \frac{1}{13}$

b. $\mathrm{P}(C) = \frac{1}{4}$

c. $\mathrm{P}(C \cap K) = \frac{1}{52}$

d. $\mathrm{P}(C \cup K) = \frac{16}{52} = \frac{4}{13}$

e. $\mathrm{P}(C') = \frac{3}{4}$

f. $\mathrm{P}(K' \cap C) = \frac{12}{52} = \frac{3}{13}$

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Probability
Exercise B, Question 2

Question:

There are 25 students in a certain tutor group at Philips College. There are 16 students in the tutor group studying German, 14 studying French and six students studying both French and German.

Find the probability that a randomly chosen student in the tutor group

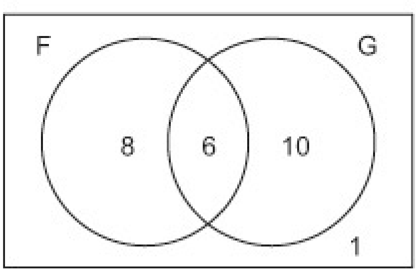
a studies French,

b studies French and German,

c studies French but not German,

d does not study French or German.

Solution:



a. $P(F) = \frac{14}{25}$

b. $P(F \cap G) = \frac{6}{25}$

c. $P(\text{French but not German}) = \frac{8}{25}$

d. $P(\text{N of French or German}) = \frac{1}{25}$

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Probability
Exercise B, Question 3

Question:

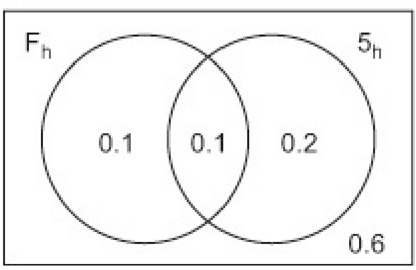
On a firing range, a rifleman has two attempts to hit a target. The probability of hitting the target with the first shot is 0.2 and the probability of hitting with the second shot is 0.3. The probability of hitting the target with both shots is 0.1.

Find the probability of

a missing the target with both shots,

b hitting with the first shot and missing with the second.

Solution:



a. P(Missing with both) = 0.6

b. P(Hitting then Missing) = 0.1

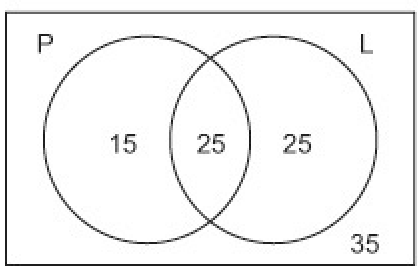
Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Probability
Exercise B, Question 4

Question:

Of all the households in the UK, 40% have a plasma TV and 50% have a laptop computer. There are 25% of households that have both a plasma TV and a laptop. Find the probability that a household chosen at random has either a plasma TV or a laptop computer but not both.

Solution:



$$\text{P(P or L but not both)} = \frac{15}{100} + \frac{25}{100}$$
$$= 0.4$$

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Probability
Exercise B, Question 5

Question:

There are 125 diners in a restaurant who were surveyed to find out if they had ordered garlic bread, beer or cheesecake.

- 15 diners had ordered all three items
- 43 diners had ordered garlic bread
- 40 diners had ordered beer
- 44 diners had ordered cheesecake
- 20 had ordered beer and cheesecake
- 26 had ordered garlic bread and cheesecake
- 25 had ordered garlic bread and beer.

A diner is chosen at random. Find the probability that the diner ordered

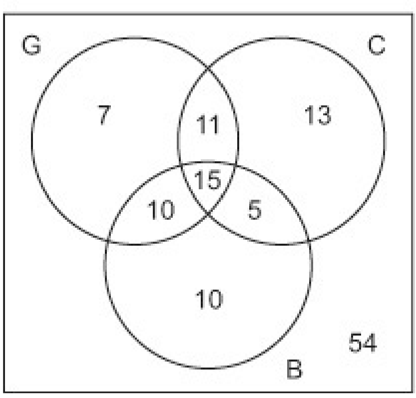
a all three items,

b beer but not cheesecake and not garlic bread,

c garlic bread and beer but not cheesecake,

d none of these items.

Solution:



a. P(all there) $= \frac{15}{125} = \frac{3}{25}$

b.
P(Beer but not cheesemake and not garlic bread) $= \frac{10}{125} = \frac{2}{25}$

c. P(Garlic bread and beer but not cheesemake) $= \frac{10}{125} = \frac{2}{25}$

d. P(None) $= \frac{54}{125}$

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Probability
Exercise B, Question 6

Question:

A group of 275 people at a music festival were asked if they play guitar, piano or drums.

- one person plays all three instruments
- 65 people play guitar and piano
- 10 people play piano and drums
- 30 people play guitar and drums
- 15 people play piano only
- 20 people play guitar only
- 35 people play drums only

a Draw a Venn diagram to represent this information.

A festival goer is chosen at random from the group.

Find the probability that the person chosen

b plays piano

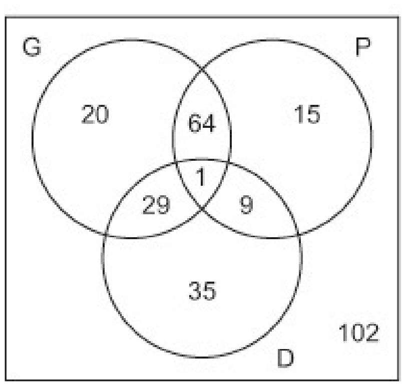
c plays at least two of guitar, piano or drums

d plays exactly one of the instruments

e plays none of the instruments.

Solution:

a.



b. P(Plays piano) $= \frac{89}{275}$

c. P(At least 2) $= \frac{103}{275}$

d. P(Plays one) $= \frac{70}{275} = \frac{14}{55}$

e. P(Plays none) $= \frac{102}{275}$

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Probability
Exercise C, Question 1

Question:

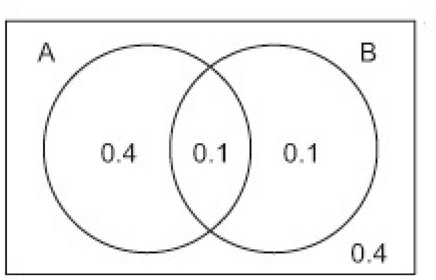
A and B are two events and P(A) = 0.5, P(B) = 0.2 and $P(A \cup B) = 0.1$.

Find

a $P(A \cup B)$, **b** $P(B')$,

c $P(A \cap B')$, **d** $P(A \cup B')$.

Solution:



a. $P(A \cup B) = 0.6$

b. $P(B') = 0.8$

c. $P(A \cap B') = 0.4$

d. $P(A \cup B') = 0.9$

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Probability
Exercise C, Question 2

Question:

A and C are two events and P(A) = 0.4, P(B) = 0.5 and $\mathrm{P}(A \cup B) = 0.6$.

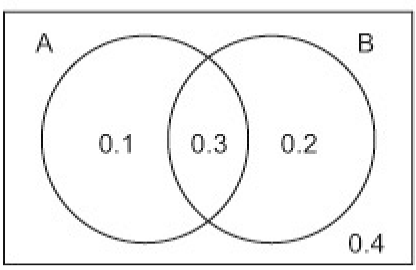
Find

a $\mathrm{P}(A \cap B)$, **b** $\mathrm{P}(A')$,

c $\mathrm{P}(A \cup B')$, **d** $\mathrm{P}(A' \cup B)$.

Solution:

a. $\mathrm{P}(A \cap B) = 0.4 + 0.5 - 0.6 = 0.3$



b. $\mathrm{P}(A') = 0.6$ c. $\mathrm{P}(A \cup B') = 0.8$ d. $\mathrm{P}(A' \cup B) = 0.9$

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Probability
Exercise C, Question 3

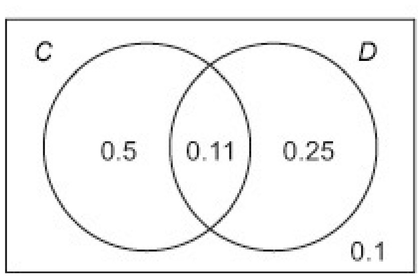
Question:

C and D are two events and P(D) = 0.4, P($C \cap D$) = 0.15 and P($C' \cap D'$) = 0.1.

Find

a $P(C' \cap D)$, **b** $P(C \cap D')$,
c $P(C)$, **d** $P(C' \cap D')$.

Solution:



a. $P(C' \cap D) = 0.25$

b. $P(C \cap D') = 0.5$

c. $P(C) = 0.65$

d. $P(C' \cap D') = 0.1$

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Probability
Exercise C, Question 4

Question:

There are two events T and Q where $\mathrm{P}(T) = \mathrm{P}(Q) = 3\mathrm{P}(T \cap Q)$ and $\mathrm{P}(T \cup Q) = 0.75$.

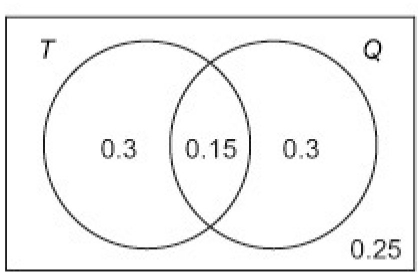
Find

a $\mathrm{P}(T \cup Q)$, **b** $\mathrm{P}(T)$, **c** $\mathrm{P}(Q')$,
d $\mathrm{P}(T' \cap Q')$, **e** $\mathrm{P}(T \cap Q')$.

Solution:

a.

$$\begin{aligned}
\mathrm{P}(T \cup Q) &= \mathrm{P}(T) + \mathrm{P}(Q) - \mathrm{P}(T \cap Q) \\
0.75 &= 3\mathrm{P}(T \cap Q) + 3\mathrm{P}(T \cap Q) - \mathrm{P}(T \cap Q) \\
5\mathrm{P}(T \cap Q) &= 0.75 \\
\mathrm{P}(T \cap Q) &= 0.15
\end{aligned}$$



b. $\mathrm{P}(T) = 0.45$

c. $\mathrm{P}(Q') = 0.55$

d. $\mathrm{P}(T' \cap Q') = 0.25$

e. $\mathrm{P}(T \cap Q') = 0.3$

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Probability
Exercise C, Question 5

Question:

A survey of all the households in the town of Bury was carried out. The survey showed that 70% have a freezer and 20% have a dishwasher and 80% have either a dishwasher or a freezer or both appliances. Find the probability that a randomly chosen household in Bury has both appliances.

Solution:

$$\begin{aligned} P(F \cap D) &= P(F) + P(D) - P(F \cup D) \\ &= 0.7 + 0.2 - 0.8 \\ &= 0.1 \end{aligned}$$

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Probability
Exercise C, Question 6

Question:

The probability that a child in a school has blue eyes is 0.27 and the probability they have blonde hair is 0.35. The probability that the child will have blonde hair or blue eyes or both is 0.45. A child is chosen at random from the school. Find the probability that the child has

a blonde hair and blue eyes,

b blonde hair but not blue eyes,

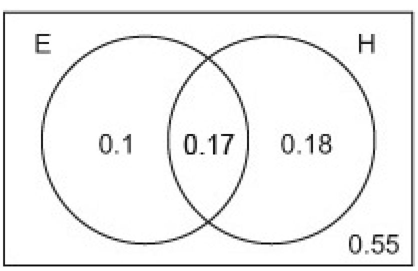
c neither feature.

Solution:

a. $P(E \cap H) = P(E) + P(H) - P(E \cup H)$

$= 0.27 + 0.35 - 0.45$

$= 0.17$



b. P(Blonde hair but not Blue eyes) = 0.18

c. P(Neither) = 0.55

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Probability
Exercise C, Question 7

Question:

A patient going in to a doctor's waiting room reads *Hiya* Magazine with probability 0.6 and *Dakor* Magazine with probability 0.4. The probability that the patient reads either one or both of the magazines is 0.7. Find the probability that the patient reads

a both magazines,

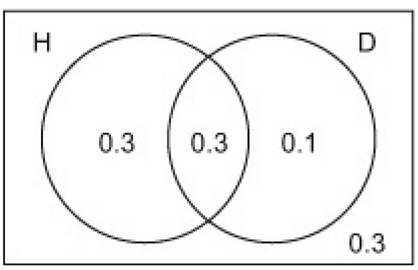
b *Hiya* Magazine only.

Solution:

a. $P(H \cap D) = P(H) + P(D) - P(H \cup D)$

$= 0.6 + 0.4 - 0.7$

$= 0.3$



b. P(Hiya only) = 0.3

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Probability
Exercise D, Question 1

Question:

A card is drawn at random from a pack of 52 playing cards. Given that the card is a diamond, find the probability that the card is an ace.

Solution:

$$\text{P(Ace | Diamond)} = \frac{1}{13} \text{ or } \text{P(Ace | Diamond)} = \frac{\text{P(Ace of Diamonds)}}{\text{P(Diamond)}} = \frac{\frac{1}{52}}{\frac{13}{52}} = \frac{1}{13}.$$

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

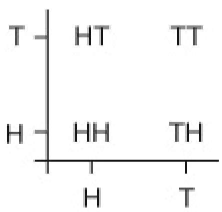
Probability
Exercise D, Question 2

Question:

Two coins are flipped and the results are recorded. Given that one coin lands on a head, find the probability of

a two heads, **b** a head and a tail.

Solution:



a.

$$P(HH\,|H) = \frac{P(\text{Head and Head})}{P(\text{Head})} = \frac{\frac{1}{4}}{\frac{3}{4}} = \frac{1}{3}.$$

b.

$$P(\text{Head and Tail}\mid\text{Head}) = \frac{P(\text{Head and Tail})}{P(\text{Head})} = \frac{\frac{2}{4}}{\frac{3}{4}} = \frac{2}{3}$$

Solutionbank S1

Edexcel AS and A Level Modular Mathematics

Probability

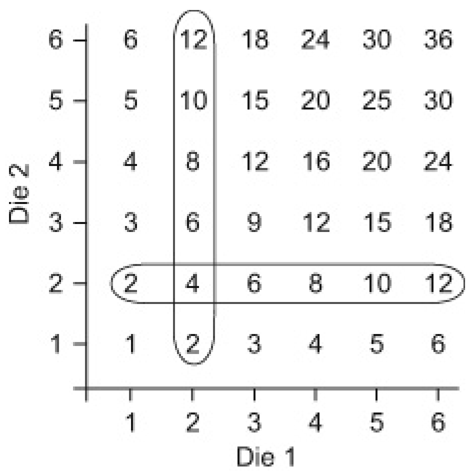
Exercise D, Question 3

Question:

Two fair dice are thrown and the product of the numbers on the dice is recorded. Given that one die lands on 2, find the probability that the product on the dice is

a exactly 6, **b** more than 5.

Solution:



a. $\text{P(Exactly } 6 \mid 2) = \dfrac{\text{P(Product of } 6 \text{ and } 2 \text{ on the die)}}{\text{P(}2 \text{ on one die)}}$

$$= \frac{\frac{2}{36}}{\frac{11}{36}} = \frac{2}{11}$$

b. $\text{P(More than } 5 \mid 2) = \dfrac{\frac{8}{36}}{\frac{11}{36}} = \dfrac{8}{11}$

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Probability
Exercise D, Question 4

Question:

A and B are two events such that $\mathrm{P}(A) = 0.6, \mathrm{P}(B) = 0.5$ and $\mathrm{P}(A \cap B) = 0.4$, find

a $\mathrm{P}(A \cup B)$, **b** $\mathrm{P}(B \mid A)$, **c** $\mathrm{P}(A \mid B)$, **d** $\mathrm{P}(A \mid B')$.

Solution:

a. $\mathrm{P}(A \cup B) = 0.6 + 0.5 - 0.4$

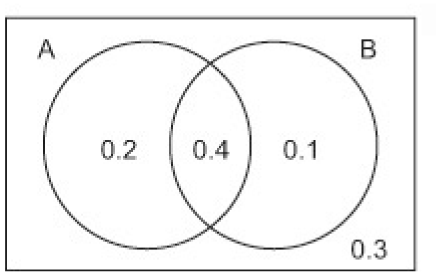
$= 0.7$

b. $\mathrm{P}(B \mid A) = \frac{\mathrm{P}(B \cap A)}{\mathrm{P}(A)} = \frac{0.4}{0.6} = \frac{2}{3}$

c. $\mathrm{P}(A \mid B) = \frac{\mathrm{P}(A \cap B)}{\mathrm{P}(B)} = \frac{0.4}{0.5} = 0.8$

d. $\mathrm{P}(A \mid B') = \frac{\mathrm{P}(A \cap B')}{\mathrm{P}(B')} = \frac{0.2}{0.5}$

$= 0.4.$



Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Probability
Exercise D, Question 5

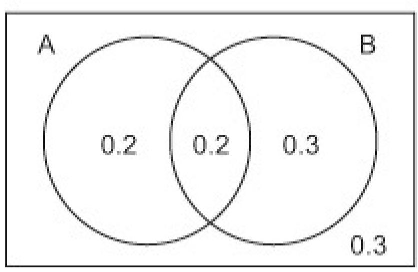
Question:

A and B are two events such that $\mathrm{P}(A) = 0.4, \mathrm{P}(B) = 0.5$ and $\mathrm{P}(A \mid B) = 0.4$, find

a $\mathrm{P}(B \mid A)$, **b** $\mathrm{P}(A' \cap B')$, **c** $\mathrm{P}(A' \cap B)$.

Solution:

$$\begin{aligned}\mathrm{P}(A \cap B) &= \mathrm{P}(A \mid B)\mathrm{P}(B)\\ &= 0.4 \times 0.5 = 0.2\end{aligned}$$



a. $\mathrm{P}(B \mid A) = \dfrac{\mathrm{P}(B \cap A)}{\mathrm{P}(A)} = \dfrac{0.2}{0.4} = \dfrac{1}{2}$

b. $\mathrm{P}(A' \cap B') = 0.3$

c. $\mathrm{P}(A' \cap B) = 0.3$

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Probability
Exercise D, Question 6

Question:

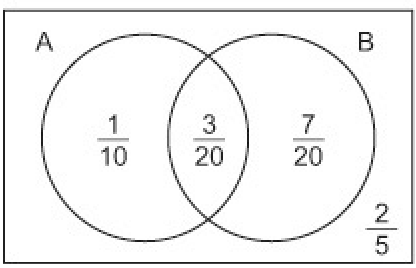
Let A and B be events such that $P(A) = \frac{1}{4}$, $P(B) = \frac{1}{2}$ and $P(A \cup B) = \frac{3}{5}$.

Find

a $P(A \mid B)$, **b** $P(A' \cap B)$, **c** $P(A' \cap B')$.

Solution:

$$P(A \cap B) = \frac{1}{4} + \frac{1}{2} - \frac{3}{5} = \frac{3}{20}$$



a. $P(A \mid B) = \frac{\frac{3}{20}}{\frac{1}{2}} = \frac{3}{10}$

b. $P(A' \cap B) = \frac{7}{20}$

c. $P(A' \cap B') = \frac{2}{5}$

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Probability
Exercise D, Question 7

Question:

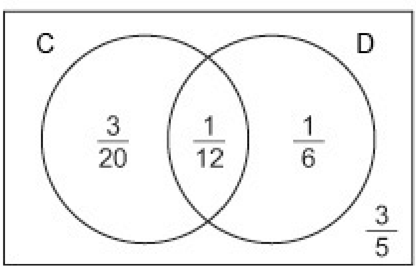
C and D are two events and $P(C \mid D) = \frac{1}{3}$, $P(C \mid D') = \frac{1}{5}$ and $P(D) = \frac{1}{4}$, find

a $P(C \cap D)$, **b** $P(C \cap D')$, **c** $P(C)$,
d $P(D \mid C)$, **e** $P(D' \mid C)$, **f** $P(C')$.

Solution:

a.

$$P(C \cap D) = P(C \mid P)P(D) = \frac{1}{3} \times \frac{1}{4} = \frac{1}{12}.$$



b. $$P(C \cap D') = P(C \mid D')P(D') = \frac{1}{5} \times \frac{3}{4} = \frac{3}{20}$$

c. $$P(C) = \frac{3}{20} + \frac{1}{12} = \frac{7}{30}$$

d. $$P(D \mid C) = \frac{P(D \cap C)}{P(C)} = \frac{\frac{1}{12}}{\frac{7}{30}} = \frac{5}{14}$$

e. $$P(D' \mid C) = 1 - \frac{5}{14} = \frac{9}{14}$$

f. $$P(C') = 1 - P(C) = \frac{7}{30} = \frac{23}{30}$$

$$\therefore P(D' \mid C') = P(C' \cap D') / P(C')$$

$$= \frac{\left(\frac{3}{5}\right)}{\left(\frac{23}{30}\right)} = \frac{18}{23}$$

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Probability
Exercise E, Question 1

Question:

A bag contains five red and four blue tokens. A token is chosen at random, the colour recorded and the token is not replaced. A second token is chosen and the colour recorded. Find the probability that

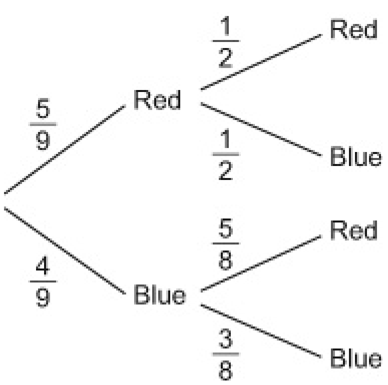
a the second token is red given the first token is blue,

b the second token is blue given the first token is red,

c both tokens chosen are blue,

d one red token and one blue token are chosen.

Solution:



a. P(Second Red | First Blue) $= \frac{5}{8}$

b. P(Second Blue | First Red) $= \frac{1}{2}$

c. P(Both Blue) $= \frac{4}{9} \times \frac{3}{8} = \frac{1}{6}$

d. P(One Red and One Blue)

= P(Red Blue) + P(Blue Red)

$= \frac{5}{9} \times \frac{1}{2} + \frac{4}{9} \times \frac{5}{8} = \frac{5}{9}$

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Probability
Exercise E, Question 2

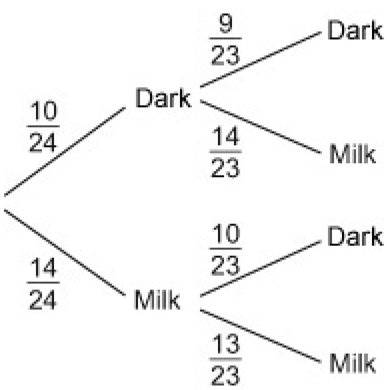
Question:

A box of 24 chocolates contains 10 dark and 14 milk chocolates. Linda chooses a chocolate at random and eats it, followed by another one. Find the probability that Linda eats

a two dark chocolates,

b one dark and one milk chocolate.

Solution:



a. P(Dark Dark) $= \frac{10}{24} \times \frac{9}{27} = \frac{15}{92}$

b. P(One Dark One Milk) $= \frac{10}{24} \times \frac{14}{23} + \frac{14}{24} \times \frac{10}{23}$

$= \frac{35}{69}$

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

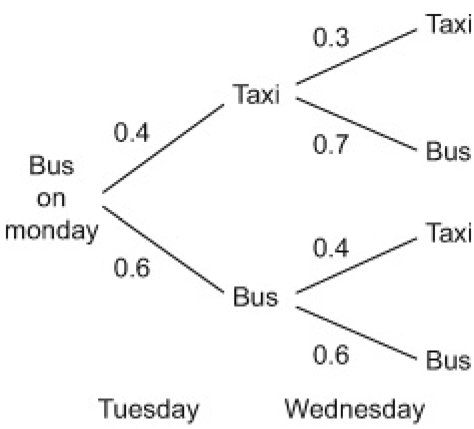
Probability
Exercise E, Question 3

Question:

Jean always goes to work by bus or takes a taxi. If one day she goes to work by bus, the probability she goes to work by taxi the next day is 0.4. If one day she goes to work by taxi, the probability she goes to work by bus the next day is 0.7.

Given that Jean takes the bus to work on Monday, find the probability that she takes a taxi to work on Wednesday.

Solution:



$$\begin{aligned} \text{P(Taxi on Weds)} &= 0.4 \times 0.3 + 0.6 \times 0.4 \\ &= 0.36 \end{aligned}$$

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Probability
Exercise E, Question 4

Question:

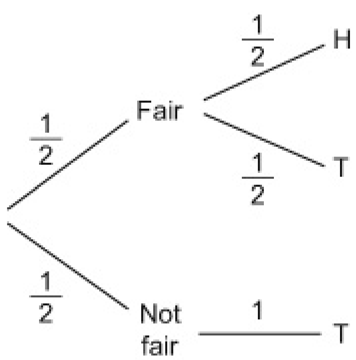
Sue has two coins. One is fair, with a head on one side and a tail on the other.

The second is a trick coin and has a tail on both sides. Sue picks up one of the coins at random and flips it.

a Find the probability that it lands heads up.

b Given that it lands tails up, find the probability that she picked up the fair coin.

Solution:



a. $\text{P(Head)} = \frac{1}{2} \times \frac{1}{2} = \frac{1}{4}$

b. $\text{P(Fair | Tail)} = \frac{\text{P(Fair and Tail)}}{\text{P(Tail)}}$

$$= \frac{\frac{1}{2} \times \frac{1}{2}}{\frac{1}{2} \times \frac{1}{2} + \frac{1}{2}} = \frac{1}{3}$$

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Probability
Exercise E, Question 5

Question:

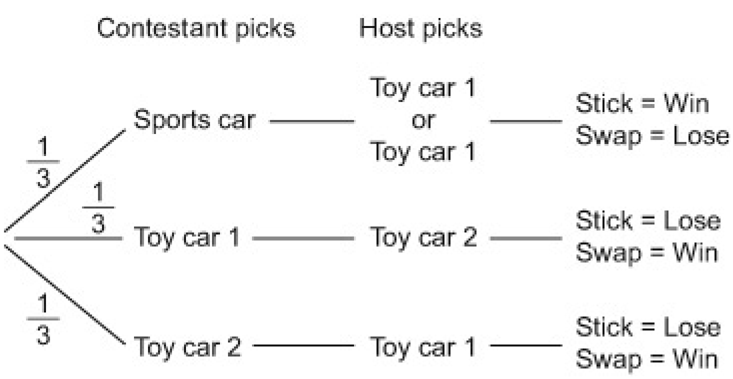
A contestant on a quiz show is asked to choose one of three doors. Behind one of the doors is the star prize of a sports car, but behind each of the other two doors there is a toy car.

The contestant chooses one of the three doors.

The host then opens one of the remaining two doors and reveals a toy car. The host then asks the contestant if they want to stick with their first choice or switch to the other unopened door.

State what you would recommend the contestant to do in order to have the greatest probability of winning the sports car. Show your working clearly.

Solution:



P(Winning if Contestant Sticks) = $\frac{1}{3}$

P(Winning if Contestant Swaps) = $\frac{2}{3}$

Contestant should *Swap* doors.

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Probability
Exercise F, Question 1

Question:

Event A and event B are mutually exclusive and $P(A) = 0.2, P(B) = 0.5$.

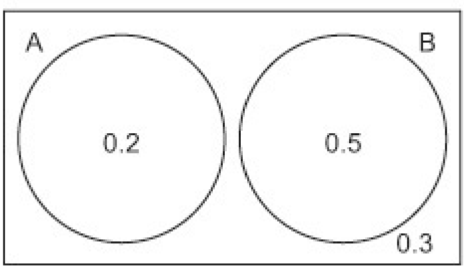
a Draw a Venn diagram to represent these two events.

b Find $P(A \cup B)$.

c Find $P(A' \cap B')$.

Solution:

a.



b. $P(A \cup B) = 0.7$

c. $P(A' \cap B') = 0.3$

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Probability
Exercise F, Question 2

Question:

Two events A and B are independent and $P(A) = \frac{1}{4}$ and $P(B) = \frac{1}{5}$.

Find

a $P(A \cap B)$, **b** $P(A \cap B')$, **c** $P(A' \cap B')$.

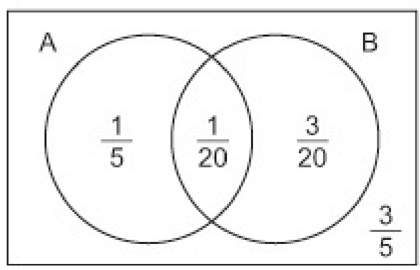
Solution:

a. $P(A \cap B) = \frac{1}{4} \times \frac{1}{5} = \frac{1}{20}$

b. $P(A \cap B') = \frac{1}{4} - \frac{1}{20} = \frac{1}{5}$

$or = \frac{1}{4} \times \frac{4}{5} = \frac{1}{5}$

c. $P(A' \cap B') = \frac{3}{5}$.



Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Probability
Exercise F, Question 3

Question:

Q and R are two events such that $P(Q) = 0.2$, $P(R) = 0.4$ and $P(Q' \cap R) = 0.4$.

Find

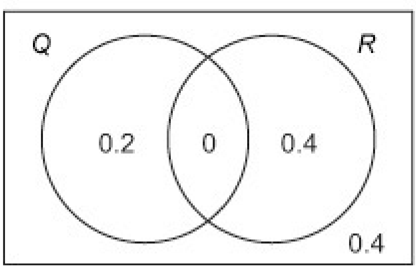
a the relationship between Q and R,

b $P(Q \cup R)$,

c $P(Q' \cap R')$.

Solution:

a. $P(Q \cap R) = 0$ so Q and R are mutually exclusive.



b. $P(Q \cup R) = 0.4 + 0.2 = 0.6$

c. $P(Q' \cap R') = 0.4$

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Probability
Exercise F, Question 4

Question:

Two fair dice are rolled and the result on each die is recorded. Show that the event ‘the sum of the scores on the dice is 4′ and ‘both dice land on the same number’ are *not* mutually exclusive.

Solution:

$\text{P(Sum of 4)} = \frac{3}{36} = \frac{1}{12}$

$\text{P(Same number)} = \frac{6}{36} = \frac{1}{6}$

$\text{P(Sum of 4)} + \text{P(Same number)} = \frac{1}{6} + \frac{1}{12} = \frac{1}{4}$

$\text{P(Sum of 4 or same number)} = \frac{8}{36} = \frac{2}{9}$

P(Sum of 4) + P(Same number) ≠ P(Sum of 4 or same number), so the events are not mutually exclusive

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Probability
Exercise F, Question 5

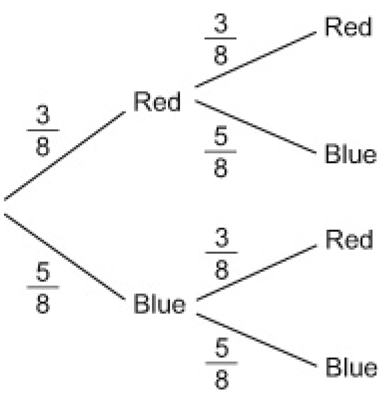
Question:

A bag contains three red beads and five blue beads. A bead is chosen at random from the bag, the colour is recorded and the bead is replaced. A second bead is chosen and the colour recorded.

a Find the probability that both beads are blue.

b Find the probability that the second bead is blue.

Solution:



a. P(Blue Blue) $= \frac{5}{8} \times \frac{5}{8} = \frac{25}{64}$

b. P(Second Blue) $= \frac{3}{8} \times \frac{5}{8} + \frac{5}{8} \times \frac{5}{8}$

$= \frac{5}{8}$

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Probability
Exercise F, Question 6

Question:

A box contains 24 electrical components of which four are known to be defective. Two items are taken at random from the box. Find the probability of selecting

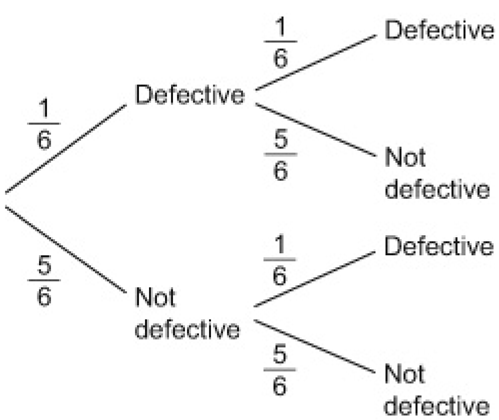
a two defective components if the first item is replaced before choosing the second item,

b two defective components if the first item is not replaced,

c one defective component and one fully functioning component if the first item is not replaced.

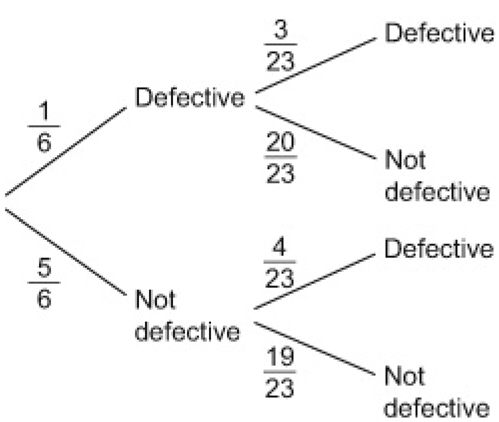
Solution:

a. *Replaced*



$$\text{P(2 defective)} = \frac{1}{6} \times \frac{1}{6} = \frac{1}{36}$$

b. Not replaced



$$\text{P(2 defective)} = \frac{1}{6} \times \frac{3}{23} = \frac{1}{46}$$

c. P(one defective and one not defective) $= \frac{1}{6} \times \frac{20}{23} + \frac{5}{6} \times \frac{4}{23} = \frac{20}{69}$

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Probability
Exercise F, Question 7

Question:

A bag contains one red, two blue and three green tokens. One token is chosen at random, the colour is recorded and the token replaced. A second token is then chosen and the colour recorded.

a Draw a tree diagram showing the possible outcomes.

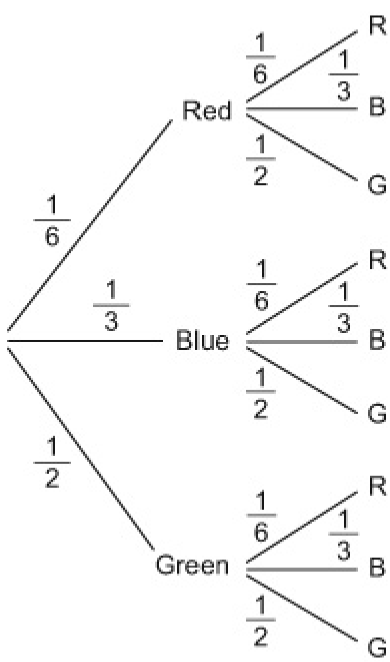
Find the probability of choosing

b two tokens of the same colour,

c two tokens that are different colours.

Solution:

a.



b. P(Same Colour)

$$= \frac{1}{6} \times \frac{1}{6} + \frac{1}{3} \times \frac{1}{3} + \frac{1}{2} \times \frac{1}{2}$$

$$= \frac{7}{18}$$

c. P(Different colours)

$$= \frac{1}{6} \times \frac{1}{3} + \frac{1}{6} \times \frac{1}{2} + \frac{1}{3} \times \frac{1}{6} + \frac{1}{3} \times \frac{1}{2} + \frac{1}{2} \times \frac{1}{6} + \frac{1}{2} \times \frac{1}{3}$$

$$= \frac{11}{18} \quad or \quad 1 - \frac{7}{18} = \frac{11}{18}$$

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

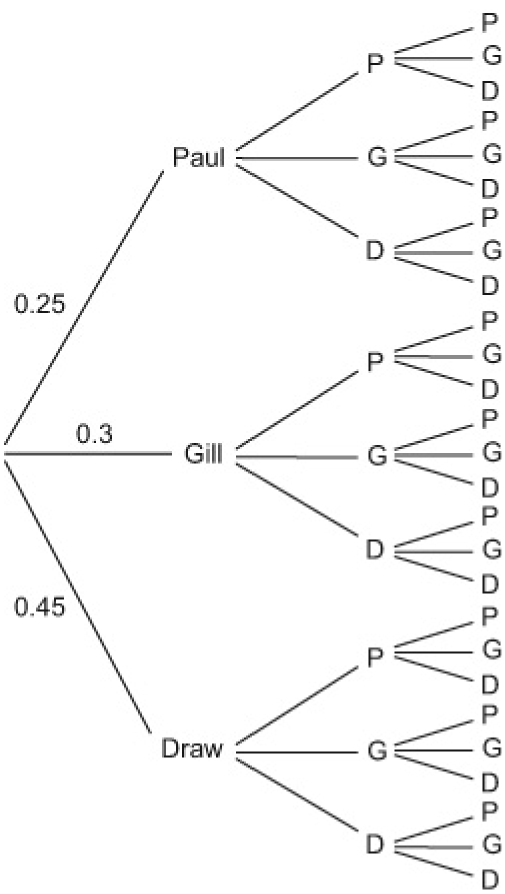
Probability
Exercise F, Question 8

Question:

Paul and Gill decide to play a board game. The probability that Paul wins the game is 0.25 and the probability that Gill wins is 0.3. They decide to play three games. Given that the results of successive games are independent, find the probability that

a Paul wins three games in a row,
b all games are drawn,
c Gill wins two games and Paul wins one game,
d each player wins just one game each.

Solution:



a.

$$\text{P(Paul Wins 3 games)} = 0.25 \times 0.25 \times 0.25 = 0.015625$$

b.

$$\text{P(All games drown)} = 0.45 \times 0.45 \times 0.45 = 0.091125$$

c.

P(Gill wins 2 and Paul wins 1) $= 3 \times 0.3 \times 0.3 \times 0.25$
$= 0.0675$

d.

P(Each player wins one game) $= 6 \times 0.25 \times 0.3 \times 0.45$
$= 0.2025$

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Probability
Exercise G, Question 1

Question:

The events A and B are such that $\mathrm{P}(A) = \frac{1}{3}$, $\mathrm{P}(B) = \frac{1}{4}$ and $\mathrm{P}(A \cup B) = \frac{1}{2}$.

a Show that A and B are independent.

b Represent these probabilities in a Venn diagram.

c Find $\mathrm{P}(A \mid B')$.

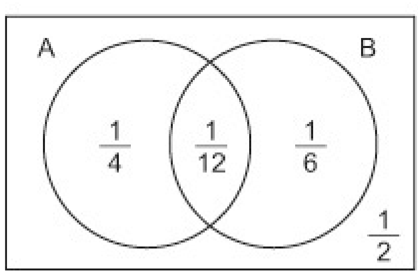
Solution:

a. $\mathrm{P}(A \cap B) = \frac{1}{3} + \frac{1}{4} - \frac{1}{2} = \frac{1}{12}$

$\mathrm{P}(A)\mathrm{P}(B) = \frac{1}{3} \times \frac{1}{4} = \frac{1}{12}$

Hence A and B are independent as $\mathrm{P}(A \cap B) = \mathrm{P}(A)\mathrm{P}(B) = \frac{1}{12}$.

b.



c. $\mathrm{P}(A \mid B') = \dfrac{\mathrm{P}(A \cap B')}{\mathrm{P}(B')}$

$= \dfrac{\frac{1}{4}}{\frac{3}{4}} = \dfrac{1}{3}$

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Probability
Exercise G, Question 2

Question:

A computer game has three levels and one of the objectives of every level is to collect a diamond. The probability of a randomly chosen player collecting a diamond on the first level is $\frac{4}{5}$, the second level is $\frac{2}{3}$ and the third level is $\frac{1}{2}$. The events are independent.

a Draw a tree diagram to represent collecting diamonds on the three levels of the game.

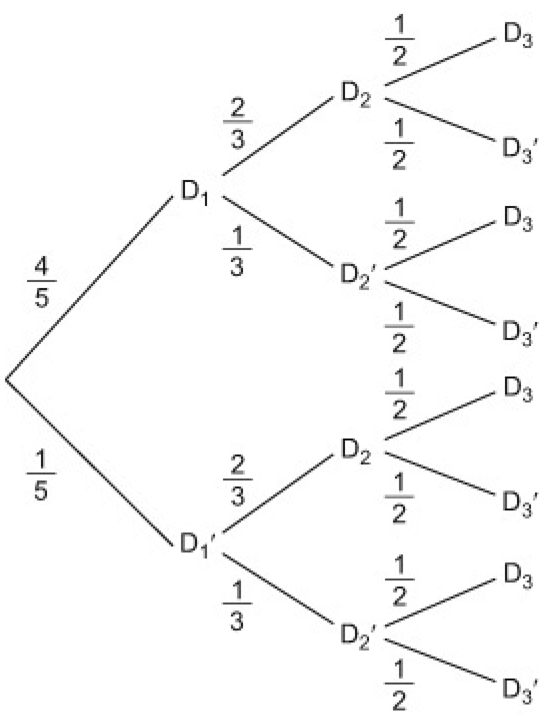
Find the probability that a randomly chosen player

b collects all three diamonds,

c collects only one diamond.

Solution:

a.



b. $P(D_1D_2D_3) = \frac{4}{5} \times \frac{2}{3} \times \frac{1}{2} = \frac{4}{15}$

c. P(only 1 diamond)

$= P(D_1D_2'D_3') + P(D_1'D_2D_3') + P(D_1'D_2'D_3)$

$= \frac{4}{5} \times \frac{1}{3} \times \frac{1}{2} + \frac{1}{5} \times \frac{2}{3} \times \frac{1}{2} + \frac{1}{5} \times \frac{1}{3} \times \frac{1}{2} = \frac{7}{30}$

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Probability
Exercise G, Question 3

Question:

An online readers' club has 50 members. Glasses are worn by 15 members, 18 are left handed and 21 are female. There are four females who are left handed, three females who wear glasses and five members who wear glasses and are left handed. Only one member wears glasses, is left handed and female.

a Draw a Venn diagram to represent these data.

A member is selected at random. Find the probability that the member

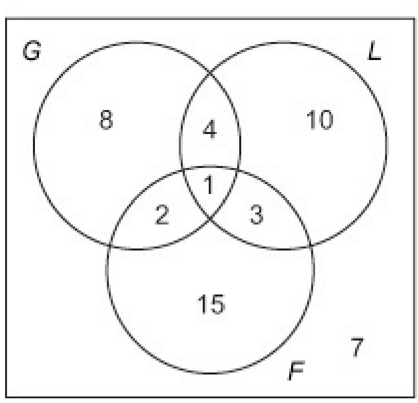
b is female, does not wear glasses and is not left handed,

c is male, does not wear glasses and is not left handed,

d wears glasses given that she is left handed and female.

Solution:

a.



b. $P(F \cap G' \cap L') = \frac{15}{50} = \frac{3}{10}$

c. $P(F' \cap G' \cap L') = \frac{7}{50}$

d. $P(G \mid L \cap F) = \frac{P(G \cap L \cap F)}{P(L \cap F)}$

$= \frac{1}{4}$

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Probability
Exercise G, Question 4

Question:

For the events J and K,

$P(J \cup K) = 0.5, P(J' \cap K) = 0.2, P(J \cap K') = 0.25.$

a Draw a Venn diagram to represent the events J and K and the sample space S.

Find

b $P(J)$,

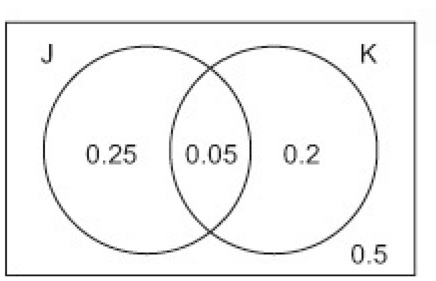
c $P(K)$,

d $P(J \mid K)$.

e Determine whether or not J and K are independent.

Solution:

a.



b. $P(J) = 0.3$

c. $P(K) = 0.25$

d. $P(J \mid K) = \frac{P(J \cap K)}{P(K)}$

$= \frac{0.05}{0.25} = 0.2$

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Probability
Exercise G, Question 5

Question:

There are 15 coloured beads in a bag; seven beads are red, three are blue and five are green. Three beads are selected at random from the bag and not replaced. Find the probability that

a the first and second beads chosen are red and the third bead is blue or green,

b one red, one blue and one green bead are chosen.

Solution:

a. P(2 red and third blue) + P(2 red and third green)

$$= \frac{7}{15} \times \frac{6}{14} \times \frac{3}{13} + \frac{7}{15} \times \frac{6}{14} \times \frac{5}{13} = \frac{8}{65}$$

b. P(one red, one blue and one green bead) $= 6 \times \frac{7}{15} \times \frac{3}{14} \times \frac{5}{13} = \frac{3}{13}$

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Probability
Exercise G, Question 6

Question:

A survey of a group of students revealed that 85% have a mobile phone, 60% have an MP3 player and 5% have neither phone nor MP3 player.

a Find the proportion of students who have both gadgets.

b Draw a Venn diagram to represent this information.

Given that a randomly selected student has a phone or an MP3 player,

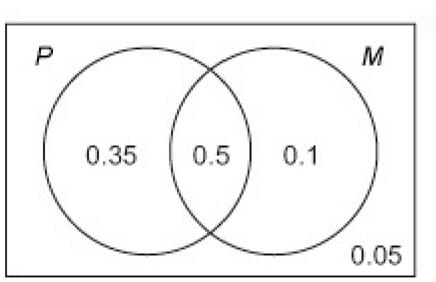
c find the probability that the student has a mobile phone.

Solution:

a. P(Phone and MP3) $= 0.85 + 0.6 - 0.95$

$= 0.5$

b.



c. $\mathrm{P}(P \mid P \cup M) = \frac{0.85}{0.95} = \frac{17}{19}$

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Probability
Exercise G, Question 7

Question:

In a factory, machines *A*, *B* and *C* produce electronic components. Machine *A* produces 16% of the components, machine *B* produces 50% of the components and machine *C* produces the rest. Some of the components are defective. Machine *A* produces 4%, machine *B* 3% and machine *C* 7% defective components.

a Draw a tree diagram to represent this information.

Find the probability that a randomly selected component is

b produced by machine *B* and is defective,

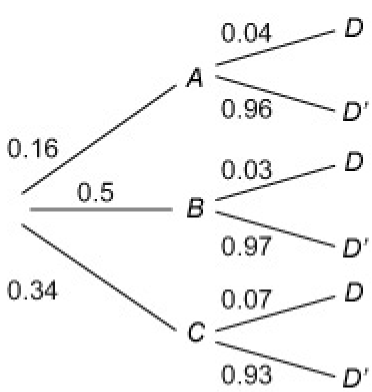
c defective.

Given that a randomly selected component is defective,

d find the probability that it was produced by machine *B*.

Solution:

a.



b. $P(B \text{ and Defective}) = 0.5 \times 0.03$

$= 0.015$

c. $P(\text{Defective}) = 0.16 \times 0.04 + 0.5 \times 0.03 + 0.34 \times 0.07$

$= 0.0452$

d. $P(B \mid \text{Defective}) = \frac{0.015}{0.0452} = 0.332$

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Probability
Exercise G, Question 8

Question:

A garage sells three types of fuel; U95, U98 and diesel. In a survey of 200 motorists buying fuel at the garage, 80 are female and the rest are male. Of the 90 motorists buying 'U95' fuel, 50 were female and of the 70 motorists buying diesel, 60 were male. A motorist does not buy more than one type of fuel.

Find the probability that a randomly chosen motorist

a buys U98 fuel,

b is male, given that the motorist buys U98 fuel.

Garage records indicate that 10% of the motorists buying U95 fuel, 30% of the motorists buying U98 fuel and 40% of the motorists buying diesel have their car serviced by the garage.

A motorist is chosen at random.

c Find the probability that this motorist has his or her car serviced by the garage.

d Given the motorist has his or her car serviced by the garage, find the probability that the motorist buys diesel fuel.

Solution:

a. $\mathrm{P(U98)} = \frac{200 - 90 - 70}{200} = \frac{1}{5}$

b. $\mathrm{P(Male \mid U98)} = \frac{\mathrm{P(Male\ and\ buys\ U98)}}{\mathrm{P(U98)}} = \frac{20}{40} = \frac{1}{2}$

c. $\mathrm{P(Serviced)} = \frac{9}{20} \times \frac{1}{10} + \frac{1}{5} \times \frac{3}{10} + \frac{7}{20} \times \frac{4}{10} = \frac{49}{200}$

d. $\mathrm{P(Diesel \mid Serviced)} = \frac{\frac{7}{20} \times \frac{4}{10}}{\frac{49}{200}} = \frac{4}{7}.$

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Probability
Exercise G, Question 9

Question:

A study was made of a group of 150 children to determine which of three cartoons they watch on television. The following results were obtained:

- 35 watch Toontime
- 54 watch Porky
- 62 watch Skellingtons
- 9 watch Toontime and Porky
- 14 watch Porky and Skellingtons
- 12 watch Toontime and Skellingtons
- 4 watch Toontime, Porky and Skellingtons

a Draw a Venn diagram to represent these data.

Find the probability that a randomly selected child from the study watches

b none of the three cartoons,

c no more than one of the cartoons.

A child selected at random from the study only watches one of the cartoons.

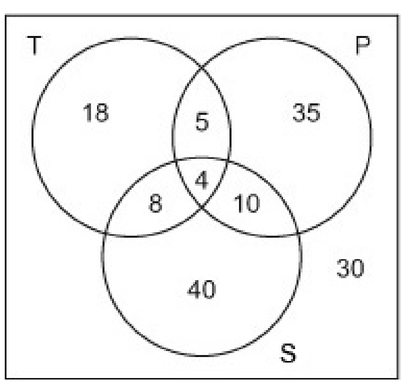
d Find the probability that it was Skellingtons.

Two different children are selected at random from the study.

e Find the probability that they both watch Skellingtons.

Solution:

a.



b. $P(\text{None}) = \frac{30}{150} = \frac{1}{5}$

c. $P(\text{No more than one}) = \frac{30+40+18+35}{150}$

$= \frac{41}{50}$

d. $P(S|\text{ only one}) = \frac{40}{93}$

e. $P(\text{Both watch K}) = \frac{62}{150} \times \frac{61}{149} = 0.169$

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Probability
Exercise G, Question 10

Question:

The members of a wine tasting club are married couples. For any married couple in the club, the probability that the husband is retired is 0.7 and the probability that the wife is retired 0.4. Given that the wife is retired, the probability that the husband is retired is 0.8.

For a randomly chosen married couple who are members of the club, find the probability that

a both of them are retired,

b only one of them is retired,

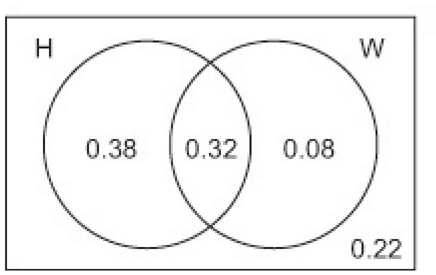
c neither of them is retired.

Two married couples are chosen at random.

d Find the probability that only one of the two husbands and only one of the two wives is retired.

Solution:

a. P(Both retired) $= 0.8 \times 0.4 = 0.32$



b. P(Only one) $= 0.38 + 0.08$

$= 0.46$

c. P(Neither) $= 0.22$

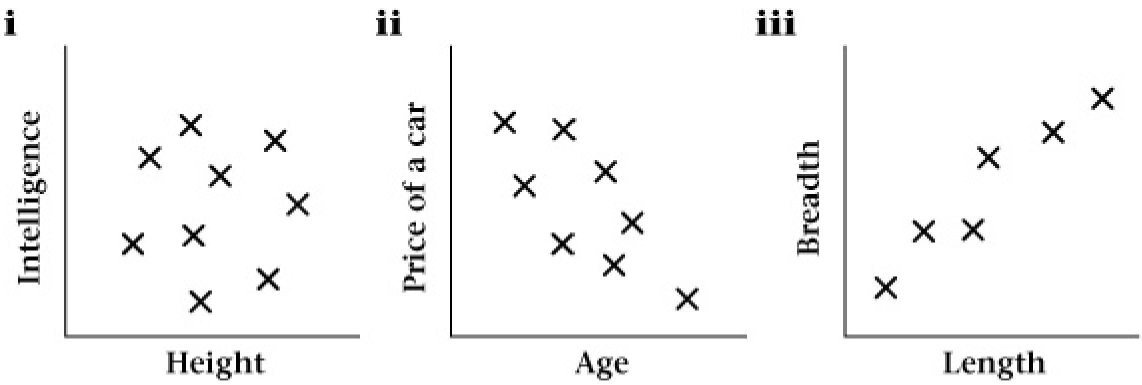
d. P(only one husband and only one wife) $= (0.38 \times 0.08 + 0.32 \times 0.22) \times 2 = 0.2016$

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Correlation
Exercise A, Question 1

Question:

The following scatter diagrams were drawn.



a Describe the type of correlation shown by each scatter diagram.

b Interpret each correlation.

Solution:

a i) no correlation – points in all four quadrants

ii) negative correlation – most points in second and fourth quadrant

iii) positive correlation – most points in first and third quadrant.

b i) There is no association between height and intelligence

ii) As age increases price decreases

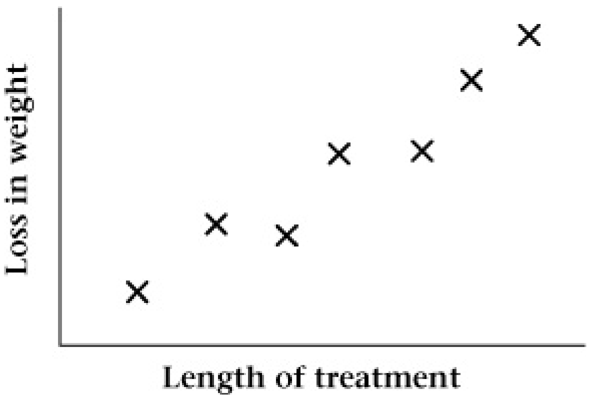
iii) As length increases breadth increases

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Correlation
Exercise A, Question 2

Question:

Some research was done into the effectiveness of a weight reducing drug. Seven people recorded their weight loss and this was compared with the length of time for which they had been treated. A scatter diagram was drawn to represent these data.



a Describe the type of correlation shown by the scatter diagram.

b Interpret the correlation in context.

Solution:

a Positive correlation.

b The longer the treatment the greater the loss of weight.

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Correlation
Exercise A, Question 3

Question:

Eight metal ingots were chosen at random and measurements were made of their breaking strength (x) and their hardness (y). The results are shown in the table below.

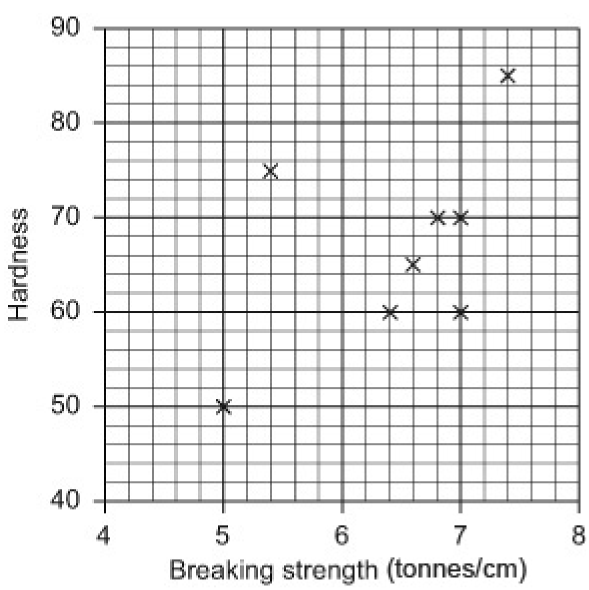
x (tonne/cm)	5	7	7.4	6.8	5.4	7	6.6	6.4
y (hardness units)	50	70	85	70	75	60	65	60

a Draw a scatter diagram to represent these data.

b Describe and interpret the correlation between the variables 'hardness' and 'breaking strength'.

Solution:

a



b There is positive correlation between hardness and breaking strength, but it is not very strong. There is some reason to believe that as breaking strength increases so does hardness.

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Correlation
Exercise A, Question 4

Question:

For each of the following data sets plot a scatter diagram, and then describe the correlation.

a

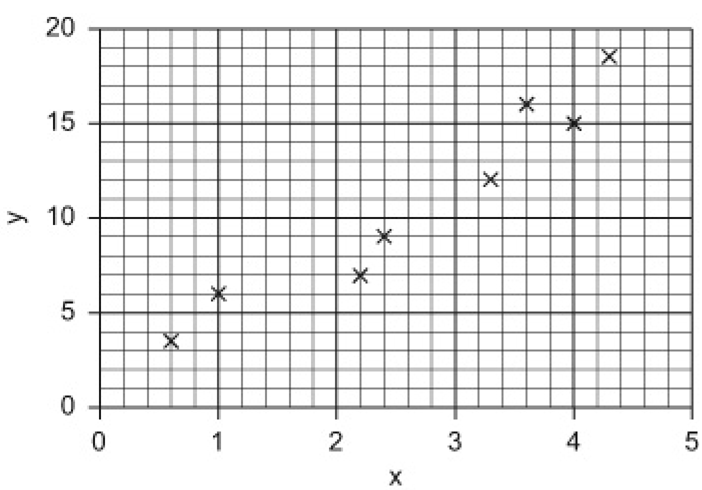
x	1	2.4	3.6	2.2	4.3	3.3	4.0	0.6
y	6.0	9.0	15.8	7.1	18.6	12.1	15.0	3.7

b

x	123	160	285	210	150	240	280	115	180
y	75	70	50	65	70	55	50	80	70

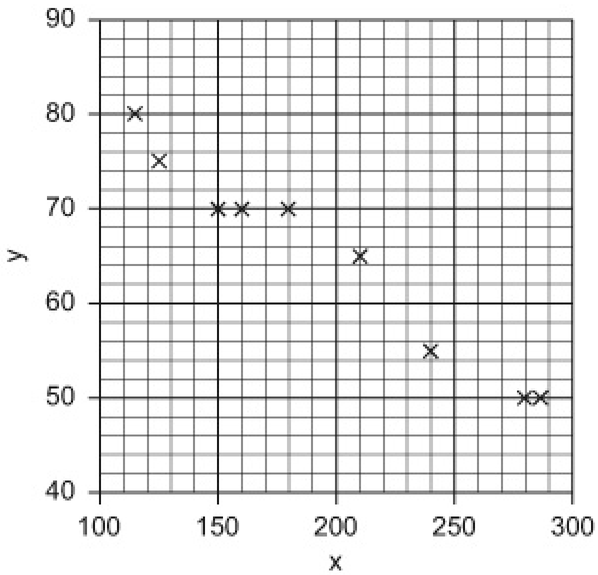
Solution:

a



The correlation is positive

b



The correlation is negative

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Correlation
Exercise A, Question 5

Question:

The table shows the armspan, in cm, and the height, in cm, of 10 adult men.

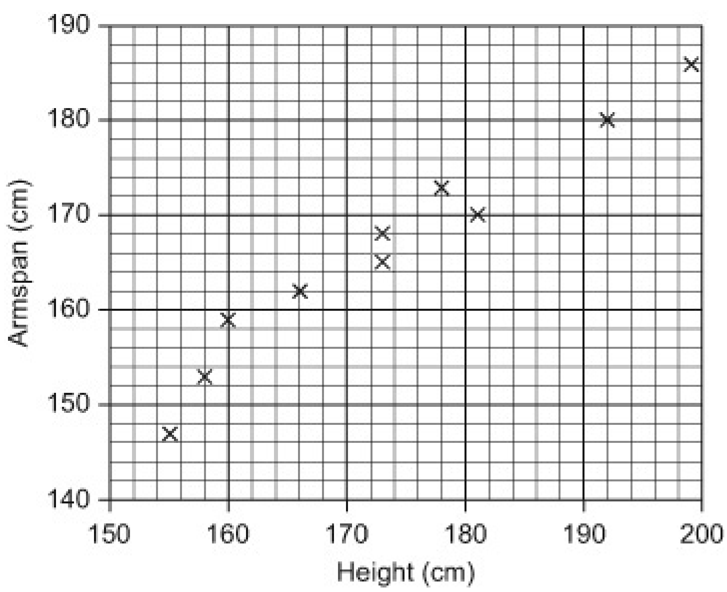
Height x (cm)	155	160	173	192	181	178	199	166	158	173
Armspan y (cm)	147	159	168	180	170	173	186	162	153	168

a Draw a scatter diagram to represent these data.

b Describe and interpret the correlation between the two variables ‘height’ and ‘armspan’.

Solution:

a



b It is positive correlation.

As height increases arm-span increases.

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Correlation
Exercise A, Question 6

Question:

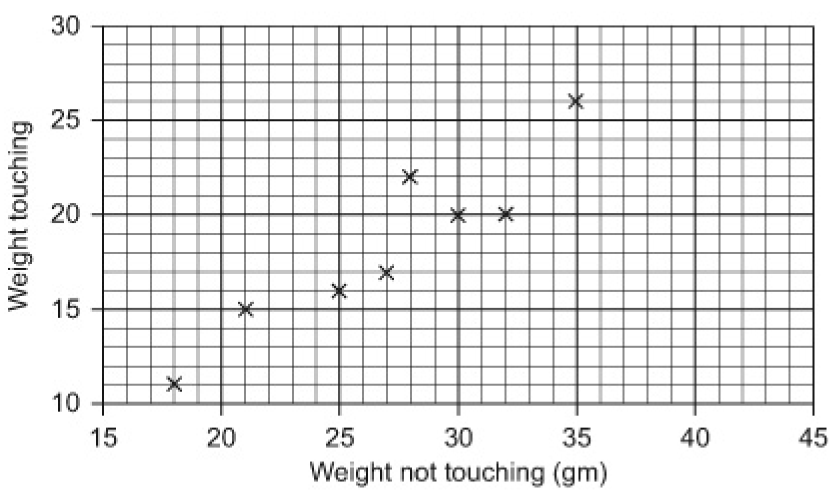
Eight students were asked to estimate the mass of a bag of sweets in grams. First they were asked to estimate the mass without touching the bag and then they were told to pick the bag up and estimate the mass again. The results are shown in the table below.

Student	A	B	C	D	E	F	G	H
Estimate of mass not touching bag (g)	25	18	32	27	21	35	28	30
Estimate of mass holding bag (g)	16	11	20	17	15	26	22	20

a Draw a scatter diagram to represent these data.

b Describe and interpret the correlation between the two variables.

Solution:



b. It shows positive correlation. As the weight not touching the bag increased so did the weight touching it. OR Students who guessed a heavy weight not touching the bag also did touching it and vice versa.

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Correlation
Exercise B, Question 1

Question:

Given $\Sigma x = 18.5$ $\Sigma x^2 = 36$ $n = 10$ find the value of S_{xx}.

Solution:

$S_{xx} = 36 - \frac{18.5 \times 18.5}{10} = 36 - 34.225 = 1.775$

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Correlation
Exercise B, Question 2

Question:

Given $\Sigma y = 25.7$ $\Sigma y^2 = 140$ $n = 5$ find the value of S_{yy}.

Solution:

$$S_{yy} = 140 - \frac{25.7 \times 25.7}{5} = 140 - 132.098 = \mathbf{7.90}$$

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Correlation
Exercise B, Question 3

Question:

Given $\Sigma x = 15$ $\Sigma y = 35$ $\Sigma xy = 91$ $n = 5$ find the value of S_{xy}.

Solution:

$S_{xy} = 91 - \frac{15 \times 35}{5} = 91 - 105 = \mathbf{-14}$

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Correlation
Exercise B, Question 4

Question:

Given that $S_{xx} = 92$, $S_{yy} = 112$ and $S_{xy} = 100$ find the value of the product moment correlation coefficient.

Solution:

$$\frac{100}{\sqrt{92 \times 112}} = \frac{100}{101.50862} = \mathbf{0.985} \ldots$$

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Correlation
Exercise B, Question 5

Question:

Given the following summary data,

$\Sigma x = 367$ $\Sigma y = 270$ $\Sigma x^2 = 33\,845$ $\Sigma y^2 = 12976$ $\Sigma xy = 17\,135$ $n = 6$

calculate the product moment correlation coefficient (r) using the formula:

$$r = \frac{S_{xy}}{\sqrt{S_{xx}S_{yy}}}$$

Solution:

$$S_{xx} = 33845 - \frac{367 \times 367}{6} = 33845 - 22448.166.. = 11396.833..$$

$$S_{yy} = 12976 - \frac{270 \times 270}{6} = 12976 - 12150 = 826$$

$$S_{xy} = 17135 - \frac{367 \times 270}{6} = 17135 - 16515 = 620$$

$$r = \frac{620}{\sqrt{11396.833 \times 826}} = \frac{620}{3068.189} = \mathbf{0.202}$$

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Correlation
Exercise B, Question 6

Question:

The ages, a years, and heights, h cm, of seven members of a team were recorded. The data were summarised as follows:

$\Sigma a = 115$ $\Sigma a^2 = 1899$ $S_{hh} = 571.4$ $S_{ah} = 72.1$

a Find S_{aa}.

b Find the value of the product moment correlation coefficient between a and h.

c Describe and interpret the correlation between the age and height of these seven people based on these data.

Solution:

a $S_{aa} = 1899 - \frac{115 \times 115}{7} = \mathbf{9.7142} \ldots ..$

b $r = \frac{72.1}{\sqrt{9.7142 \ldots \times 571.4}} = \frac{72.1}{74.50 \ldots} = 0.9677 \ldots = \mathbf{0.968}$

c This is positive correlation. The older the age the taller the person.

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Correlation
Exercise B, Question 7

Question:

In research on the quality of bacon produced by different breeds of pig, data were obtained about the leanness (l) and taste (t) of the bacon. The data is shown in the table.

Leanness l	1.5	2.6	3.4	5.0	6.1	8.2
Taste t	5.5	5.0	7.7	9.0	10.0	10.2

a Find S_{ll}, S_{tt} and S_{lt}.

b Calculate the product moment correlation coefficient between l and t using the values found in **a.** If you have a calculator that will work out r use it to check your answer.

Solution:

a $\sum l = 26.8 \quad \sum_{n=6} l^2 = 150.02 \quad \sum t = 47.4 \quad \sum t^2 = 399.58 \sum lt = 237.07$

$S_{ll} = 150.02 - \frac{26.8 \times 26.8}{6} = 150.02 - 119.7066\ldots. = \mathbf{30.3133}\ldots$

$S_{tt} = 399.58 - \frac{47.4 \times 47.4}{6} = 399.58 - 374.46 = \mathbf{25.12}$

$S_{lt} = 237.07 - \frac{26.8 \times 47.4}{6} = 237.07 - 211.72 = \mathbf{25.35}$

b $r = \frac{25.35}{\sqrt{30.3133 \times 25.12}} = \frac{25.35}{27.5947\ldots} = 0.9186\ldots. = \mathbf{0.919}$

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Correlation
Exercise B, Question 8

Question:

Eight children had their IQ measured and then took a general knowledge test. Their IQ, (x), and their marks, (y), for the test were summarised as follows:

$\Sigma x = 973$ $\Sigma x^2 = 120\ 123$ $\Sigma y = 490$ $\Sigma y^2 = 33\ 000$ $\Sigma xy = 61\ 595.$

a Calculate the product moment correlation coefficient.

b Describe and interpret the correlation coefficient between IQ and general knowledge.

Solution:

a $S_{xx} = 120123 - \frac{973 \times 973}{8} = 120123 - 118341.125 = \mathbf{1781.875}$

$S_{yy} = 33000 - \frac{490 \times 490}{8} = 33000 - 30012.5 = \mathbf{2987.5}$

$S_{xy} = 61595 - \frac{973 \times 490}{8} = 61595 - 59596.25 = \mathbf{1998.75}$

$r = \frac{1998.75}{\sqrt{1781.875 \times 2987.5}} = \frac{1998.75}{2307.2389} = 0.8662\ldots = \mathbf{0.866}$

b The correlation is positive. The higher the IQ, the higher the mark gained in the general knowledge test. (OR The higher the mark gained in the intelligence test the higher the IQ.)

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Correlation
Exercise B, Question 9

Question:

In a training scheme for young people, the average time taken for each age group to reach a certain level of proficiency was measured. The data are shown in the table.

Age *x* (years)	16	17	18	19	20	21	22	23	24	25
Average time *y* (hours)	12	11	10	9	11	8	9	7	6	8

a Find S_{xx}, S_{yy} and S_{xy}.

b Use your answers to calculate the product moment correlation coefficient (r).

c Describe and interpret the relationship between average time and age.

Solution:

a $\sum x = 205 \quad \sum x^2 = 4285 \quad \sum y = 91 \quad \sum y^2 = 861 \quad \sum xy = 1821$

$S_{xx} = 4285 - \frac{205 \times 205}{10} = 4285 - 4202.5 = \mathbf{82.5}$

$S_{yy} = 861 - \frac{91 \times 91}{10} = 861 - 828.1 = \mathbf{32.9}$

$S_{xy} = 1821 - \frac{205 \times 91}{10} = 1821 - 1865.5 = \mathbf{-44.5}$

b $r = \frac{-44.5}{\sqrt{82.5 \times 32.9}} = \frac{-44.5}{52.09846\ldots} = -0.8541\ldots = \mathbf{-0.854}$

c The correlation is negative. The greater the age the less time taken to reach the required level of proficiency.

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Correlation
Exercise C, Question 1

Question:

The following product moment correlation coefficients were calculated

i −0.96 **ii** −0.35 **iii** 0 **iv** 0.72

Write down the coefficient that

a shows the least correlation, **b** shows the most correlation.

Solution:

a (iii) The value 0 shows no correlation.

b (i) −0.96 is high negative correlation.

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Correlation
Exercise C, Question 2

Question:

Here are some product moment correlation coefficients.

i −1, **ii** −0.5, **iii** 0 **iv** 0.5, **v** 1.

Write down which one shows

a perfect negative correlation, **b** zero correlation.

Solution:

a (i)

b (iii)

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Correlation
Exercise C, Question 3

Question:

Ahmed works out the product moment correlation coefficient between the heights of a group of fathers and the heights of their sons to be 0.954. Write down what this tells you about the relationship between their heights.

Solution:

There is a strong positive correlation between the heights of fathers and their sons.

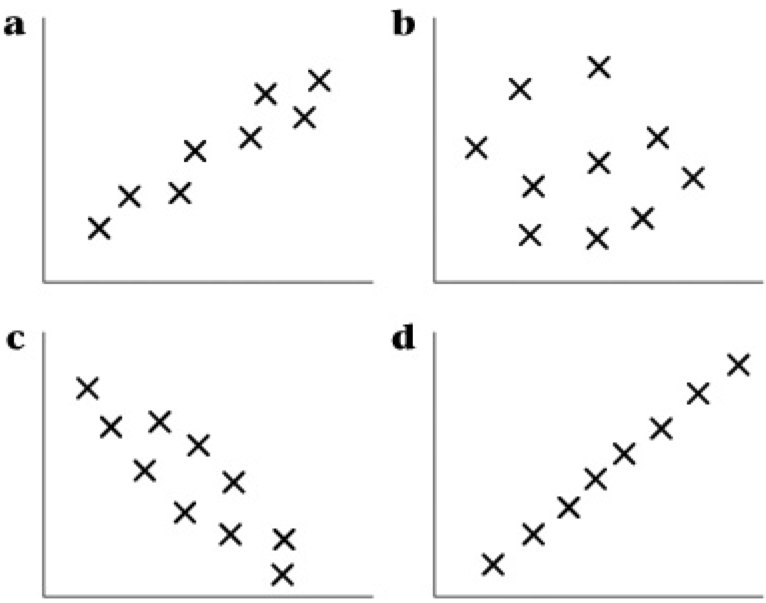
The taller the father the taller the son will be.

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Correlation
Exercise C, Question 4

Question:

Maria draws some scatter diagrams. They are shown below.



Write down which scatter diagram shows:

i a correlation of +1,

ii a correlation that could be described as strong positive correlation,

iii a correlation of about −0.97,

iv a correlation that shows almost no correlation.

Solution:

a goes with (ii)

b goes with (iv)

c goes with (iii)

d goes with (i).

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Correlation
Exercise C, Question 5

Question:

Jake finds that the product moment correlation coefficients between two variables x and y is 0.95. The product moment correlation coefficient between two other variables s and t was −0.95. Discuss how these two coefficients differ.

Solution:

x and y have a positive correlation that is close to 1. As one increases so does the other.

s and t have a negative correlation that is close to −1. As one rises the other falls.

The rate of rise in one pair of variables is the same as the rate of fall of the other pair.

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Correlation
Exercise C, Question 6

Question:

Patsy collects some data to find out if there is any relationship between the numbers of car accidents and computer ownership. She calculates the product moment correlation coefficient between the two variables. There is a strong positive correlation. She says as car accidents increase so does computer ownership. Write down whether or not this is sensible. Give reasons for your answer.

Solution:

This is not sensible as there is no way that one is directly dependent on the other. It could be that you are more likely to drive a car if you own a computer.

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Correlation
Exercise C, Question 7

Question:

Raj collects some data to find out whether there is any relationship between the height of students in his year group and the pass rate in driving tests. He finds that there is a strong positive correlation. He says that as height increases, so does your chance of passing your driving test. Is this sensible? Give reasons for your answer.

Solution:

This is not sensible. Pass rates in driving tests do not depend on height. There will be some other reason for his results. Possibly the ages of the students are different or it could just be accidental.

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Correlation
Exercise D, Question 1

Question:

Coding is to be used to work out the value of the product moment correlation coefficient for the following sets of data. Suggest a suitable coding for each.

a

x	2000	2010	2015	2005	2003	2006
y	3	6	21	6	9	18

b

s	100	300	200	400	300	700
t	2	0	1	3	3	6

Solution:

a

$x - 2000$ and $\frac{y}{3}$(OR x − any number beginning 20--)

b $\frac{s}{100}$ and leave t as it is.

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Correlation
Exercise D, Question 2

Question:

For the two variables x and y the coding of $A = x - 7$ and $B = y - 100$ is to be used.

The product moment correlation coefficient for A and B is found to be 0.973.

What is the product moment correlation coefficient for x and y?

Solution:

0.973

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Correlation
Exercise D, Question 3

Question:

Use the coding: $p = x$ and $q = y - 100$ to work out the product moment correlation coefficient for the following data.

x	0	5	3	2	1
y	100	117	112	110	106

Solution:

p	0	5	3	2	1
q	0	17	12	10	6

$\Sigma p = 11$ $\Sigma p^2 = 39$ $\Sigma q = 45$ $\Sigma q^2 = 569$ $\Sigma pq = 147$

$$S_{pp} = 39 - \frac{11 \times 11}{5} = 14.8$$

$$S_{qq} = 569 - \frac{45 \times 45}{5} = 164$$

$$S_{pq} = 147 - \frac{11 \times 45}{5} = 48$$

$$r = \frac{48}{\sqrt{14.8 \times 164}} = \frac{48}{49.2666\ldots} = 0.9742\ldots = \mathbf{0.974}$$

Coding does not affect the value of the product moment correlation coefficient.

So for x and y we have $\boldsymbol{r} = \mathbf{0.974}$

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Correlation
Exercise D, Question 4

Question:

The product moment correlation is to be worked out for the following data set using coding.

x	50	40	55	45	60
y	4	3	5	4	6

a Using the coding $p = \frac{x}{5}$ and $t = y$ find the values of S_{pp}, S_{tt} and S_{pt}.

b Calculate the product moment correlation between p and t.

c Write down the product moment correlation between x and y.

Solution:

a

p	10	8	11	9	12
t	4	3	5	4	6

$\Sigma p = 50$ $\Sigma p^2 = 510$ $\Sigma t = 22$ $\Sigma t^2 = 102$ $\Sigma pt = 227$

$S_{pp} = 510 - \frac{50 \times 50}{5} = \mathbf{10}$

$S_{tt} = 102 - \frac{22 \times 22}{5} = \mathbf{5.2}$

$S_{pt} = 227 - \frac{50 \times 22}{5} = \mathbf{7}$

b

$r = \frac{7}{\sqrt{10 \times 5.2}} = \frac{7}{7.2111\ldots} = 0.9707\ldots = \mathbf{0.971}$

c

$\boldsymbol{r} = \mathbf{0.971}$ (Coding has no effect on the value of r)

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Correlation
Exercise D, Question 5

Question:

The tail length (t cm) and the mass (m grams) for each of eight woodmice were measured. The data is shown in the table.

***t* (cm)**	8.5	7.5	8.6	7.3	8.1	7.5	8.0	7.8
***m* (g)**	28	22	26	21	25	20	20	22

a Using the coding $x = t - 7.3$ and $y = m - 20$ complete the following table

x	1.2			0				0.5
y	8			1				

b Find S_{xx}, S_{yy} and S_{xy}.

c Calculate the value of the product moment correlation coefficient between x and y.

d Write down the product moment correlation coefficient between t and m.

e Write down the conclusion that can be drawn about the relationship between tail length and mass of woodmice.

Solution:

a

x	1.2	0.2	1.3	0	0.8	0.2	0.7	0.5
y	8	2	6	1	5	0	0	2

b

$\Sigma x = 4.9$ $\Sigma x^2 = 4.59$ $\Sigma y = 24$ $\Sigma y^2 = 134$ $\Sigma xy = 22.8$

$S_{xx} = 4.59 - \frac{4.9 \times 4.9}{8} = \mathbf{1.58875}$

$S_{yy} = 134 - \frac{24 \times 24}{8} = \mathbf{62}$

$S_{xy} = 22.8 - \frac{4.9 \times 24}{8} = \mathbf{8.1}$

c

$r = \frac{8.1}{\sqrt{1.58875 \times 62}} = \frac{8.1}{9.9248\ldots} = 0.8161\ldots = \mathbf{0.816}$

d

***r* = 0.816**

e Positive correlation.

The greater the mass of a wood mouse the longer the tail length.

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Correlation
Exercise D, Question 6

Question:

A shopkeeper thinks that the more newspapers he sells in a week the more sweets he sells. He records the amount of money (m pounds) that he takes in newspaper sales and also the amount of money he takes in sweet sales (s pounds) each week for seven weeks. The data are shown are the following table.

Newspaper sales (*m* pounds)	380	402	370	365	410	392	385
Sweet sales (*s* pounds)	560	543	564	573	550	544	530

a Use the coding $x = m - 365$ and $y = s - 530$ to find S_{xx}, S_{yy} and S_{xy}.

b Calculate the product moment correlation coefficient for m and s.

c State, with a reason, whether or not what the shopkeeper thinks is correct.

Solution:

a

x	15	37	5	0	45	27	20
y	30	13	34	43	20	14	0

$\Sigma x = 149$ $\Sigma x^2 = 4773$ $\Sigma y = 154$ $\Sigma y^2 = 4670$ $\Sigma xy = 2379$

$S_{xx} = 4773 - \frac{149 \times 149}{7} = \mathbf{1601.4285}$

$S_{yy} = 4670 - \frac{154 \times 154}{7} = \mathbf{1282}$

$S_{xy} = 2379 - \frac{149 \times 154}{7} = \mathbf{-899}$

b

$\boldsymbol{r} = \frac{-899}{\sqrt{1601.4285 \times 1282}} = \frac{-899}{1432.84 \ldots ..} = -0.6274 \ldots = \mathbf{-0.627}$

c The shopkeeper is not correct. This is negative correlation so as the newspaper sales go up the sweet sales go down

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Correlation
Exercise E, Question 1

Question:

The following table shows the distance (x) in miles and the cost (y) in pounds of each of 10 taxi journeys.

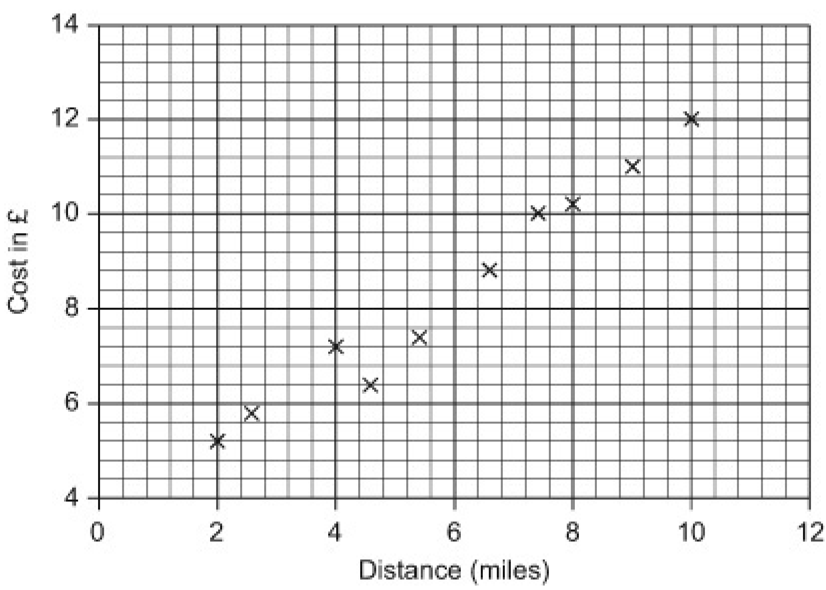
x (miles)	8	6.5	4	2.5	5.5	9	2	10	4.5	7.5
y (pounds)	10.2	8.8	7.2	5.7	7.4	11.0	5.2	12.0	6.4	10.0

a Draw a scatter diagram to represent these data.

b Describe and interpret the correlation between the two variables.

Solution:

a



b

The correlation is positive. The further the taxi travels the more it costs.

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Correlation
Exercise E, Question 2

Question:

The following scatter diagrams were drawn.

a State whether each shows positive, negative or no correlation.

b Interpret each scatter diagram in context.

Solution:

a i is positive correlation. ii is negative correlation. iii is no correlation.

b i – The older the snake the longer it is likely to be.

ii – The higher the unemployment the lower the drop in wages.

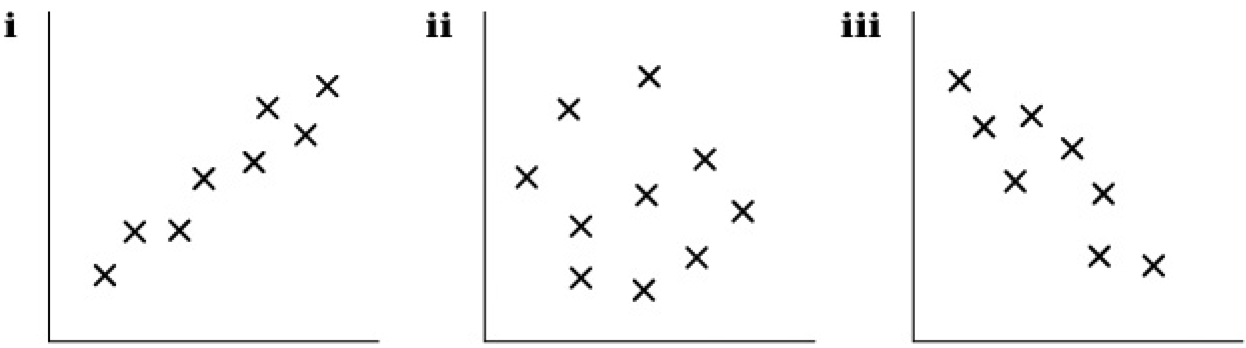
iii – There is no correlation between the age and the height of men.

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Correlation
Exercise E, Question 3

Question:

The following scatter diagrams were drawn by a student.



The student calculated the product moment correlation coefficient for each set of data. The values were:

a −0.12 **b** 0.87 **c** −0.81

Write down which value corresponds to each scatter diagram. Give a reason for your answer.

Solution:

(i) is 0.87 (ii) is − 0.12 (iii) is − 0.81.

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Correlation
Exercise E, Question 4

Question:

The product moment correlation coefficient for a person's age and his score on a memory test is −0.86. Interpret this value.

Solution:

As a persons age increases their score on a memory test decreases.

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Correlation
Exercise E, Question 5

Question:

Wai wants to know whether the 10 people in her group are as good at Science as they are at Art. She collected the end of term test marks for Science (s), and Art (a), and coded them using $x = \frac{s}{10}$ and $y = \frac{a}{10}$.

The data she collected can be summarised as follows,

$\Sigma x = 67 \quad \Sigma x^2 = 465 \quad \Sigma y = 65 \quad \Sigma y^2 = 429 \quad \Sigma xy = 434.$

a Work out the product moment correlation coefficient for x and y.

b Write down the product moment correlation coefficient for s and a.

c Write down whether or not it is it true to say that the people in Wai's group who are good at Science are also good at Art. Give a reason for your answer.

Solution:

a

$S_{xx} = 465 - \frac{67 \times 67}{10} = 16.1$

$S_{yy} = 429 - \frac{65 \times 65}{10} = 6.5$

$S_{xy} = 434 - \frac{67 \times 65}{10} = -1.5$

$r = \frac{-1.5}{\sqrt{16.1 \times 6.5}} = \frac{-1.5}{10.2298 \ldots} = -0.1466 \ldots = \mathbf{-0.147}$

b r = − 0.147

c This is negative correlation that is close to 0. There is little evidence to suggest that students in the group who are good at science will also be good at art.

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Correlation
Exercise E, Question 6

Question:

Nimer thinks that oranges that are very juicy cost more than those that are not very juicy. He buys 20 oranges from different places, and measures the amount of juice (j ml), that each orange produces. He also notes the price (p) of each orange.

The data can be summarised as follows,

$\Sigma j = 979 \quad \Sigma p = 735 \quad \Sigma j^2 = 52\,335 \quad \Sigma p^2 = 32\,156 \quad \Sigma jp = 39\,950.$

a Find S_{jj}, S_{pp} and S_{jp}.

b Using your answers to **a** calculate the product moment correlation coefficient.

c Describe the type of correlation between the amount of juice and the cost and state, with a reason, whether or not Nimer is correct.

Solution:

a

$$S_{jj} = 52335 - \frac{979 \times 979}{20} = \mathbf{4412.95}$$

$$S_{pp} = 32156 - \frac{735 \times 735}{20} = \mathbf{5144.75}$$

$$S_{jp} = 39950 - \frac{979 \times 735}{20} = \mathbf{3971.75}$$

b

$$r = \frac{3971.75}{\sqrt{4412.95 \times 5144.75}} = \frac{3971.75}{4764.8215} = 0.8335 \ldots = \mathbf{0.834}$$

c This is a positive correlation that is close to 1 so Nimer is correct.

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Correlation
Exercise E, Question 7

Question:

The following table shows the values of two variables v and m.

v	50	70	60	82	45	35	110	70	35	30
m	140	200	180	210	120	100	200	180	120	60

The results were coded using $x = v - 30$ and $y = \frac{m}{20}$.

a Complete the table for x and y.

x	20	40			15			40		0
y	7	10		10.5	6					3

b Calculate S_{xx}, S_{yy} and S_{xy}.

(You may use $\Sigma x = 287$, $\Sigma x^2 = 13\,879$, $\Sigma y = 75.5$, $\Sigma y^2 = 627.25$, $\Sigma xy = 2661$.)

c Using your answers to **b** calculate the product moment correlation coefficient for x and y.

d Write down the product moment correlation coefficient for v and m.

e Describe and interpret your product moment correlation coefficient for v and m.

Solution:

a

x	20	40	30	52	15	5	80	40	5	0
y	7	10	9	10.5	6	5	10	9	6	3

b

$S_{xx} = 13879 - \frac{287 \times 287}{10} = \mathbf{5642.1}$

$S_{yy} = 627.25 - \frac{75.5 \times 75.5}{10} = \mathbf{57.225}$

$S_{xy} = 2661 - \frac{287 \times 75.5}{10} = \mathbf{494.15}$

c $r = \frac{494.15}{\sqrt{5642.1 \times 57.225}} = \frac{494.15}{568.2} = 0.8696\ldots = \mathbf{0.870}$

d $r = \mathbf{0.870}$

e This is positive correlation that is close to 1. As v increases so m increases.

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Correlation
Exercise E, Question 8

Question:

Each of 10 cows was given an additive (x) every day for four weeks to see if it would improve their milk yield (y). At the beginning the average milk yield per day was 4 gallons. The milk yield of each cow was measured on the last day of the four weeks. The data collected is shown in the table.

Cow	A	B	C	D	E	F	G	H	I	J
Additive, x (25 gm units)	1	2	3	4	5	6	7	8	9	10
Yield, y (gallons)	4.0	4.2	4.3	4.5	4.5	4.7	5.2	5.2	5.1	5.1

a Draw a scatter diagram of these data.

b Write down any conclusions you can draw from the scatter diagram.

c From the diagram write down, with a reason, the amount of additive that could be given to each cow to maximise yield and minimise cost.

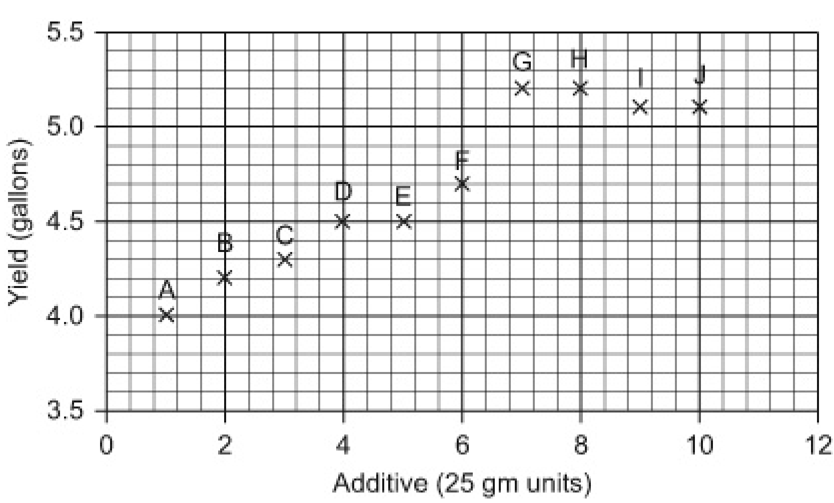
d The product moment correlation coefficient is to be calculated for the first seven cows. Write down why you think cows H, I and J are being left out for this calculation.

e Use the values $S_{xx} = 28$, $S_{yy} = 0.90857$ and $S_{xy} = 4.8$ to calculate the product moment correlation coefficient for the seven cows.

f Write down, with a reason, how the product moment correlation coefficient for all 10 cows would differ from your answer to **e**.

Solution:

a



b

The additive seems to improve milk yield as the scatter diagram shows positive correlation. Generally as the additive is increased so the yield increases.

There could however be some other reasons for some of the increase in yield.

c Seven units. (The yield levels off at this point.)

d The scatter diagram stops rising after cow seven (G).

e $r = \frac{4.8}{\sqrt{28 \times 0.90857}} = \frac{4.8}{5.0438\ldots} = 0.9516\ldots = \mathbf{0.952}$

f It would go down as the scatter diagram ceases to rise for the last three cows – it goes down for the last two so the correlation would not be as strong.

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Correlation
Exercise E, Question 9

Question:

The following table shows the engine size (c), in cubic centimetres, and the fuel consumption (f), in miles per gallon to the nearest mile, for 10 car models.

c (cm^3)	1000	1200	1400	1500	1600	1800	2000	2200	2500	3000
f (mpg)	46	42	43	39	41	37	35	29	28	25

a On graph paper draw a scatter diagram to represent these data.

b Write down whether the correlation coefficient between c and f is positive or negative. Give a reason for your answer.

The data can be summarised as follows:

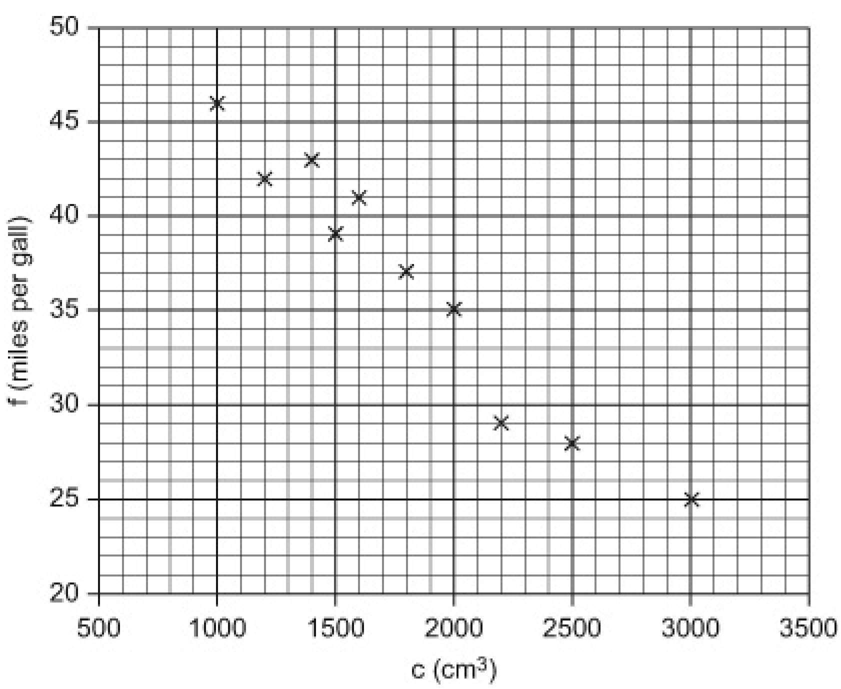
$\Sigma cf = 626\ 100$, $\Sigma c = 18\ 200$, $\Sigma f = 365$

c Calculate S_{cf}

d The product moment correlation coefficient could be found by using coding. Suggest suitable coding.

Solution:

a



b The correlation is negative. As the number of cc's goes up the petrol consumption goes down. If the axes were moved to go through the mean point most values would be in the second and fourth quadrant.

c $S_{cf} = 626100 - \frac{18200 \times 365}{10} = \mathbf{-38200}$

d $\frac{c}{100}$ or $\frac{c}{200}$ and $f - 25$ ($\frac{c-1000}{100}$ is another alternative)

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Correlation
Exercise E, Question 10

Question:

In a study on health, a clinic measured the age, (*a* years), and the diastolic blood pressure, (*d* in mm of mercury), of eight patients. The table shows the results.

***a* (years)**	20	35	50	25	60	45	25	70
***d* (mm)**	55	60	80	85	75	85	70	85

a Using the coding $x = \frac{a}{5}$ and $y = \frac{d}{5} - 11$ calculate S_{xx}, S_{yy} and S_{xy}.

b Using your answers to **a** work out the product moment correlation coefficient for x and y.

c Write down the product moment correlation coefficient between a and d.

d Interpret your answer to **c**.

Solution:

a

x	4	7	10	5	12	9	5	14
y	0	1	5	6	4	6	3	6

$\Sigma x = 66$ $\Sigma x^2 = 636$ $\Sigma y = 31$ $\Sigma y^2 = 159$ $\Sigma xy = 288$

$$S_{xx} = 636 - \frac{66 \times 66}{8} = \mathbf{91.5}$$

$$S_{yy} = 159 - \frac{31 \times 31}{8} = \mathbf{38.875}$$

$$S_{xy} = 288 - \frac{66 \times 31}{8} = \mathbf{32.25}$$

b $r = \frac{32.25}{\sqrt{91.5 \times 38.875}} = \frac{32.25}{59.6411\ldots} = 0.5407\ldots = \mathbf{0.541}$

c $\boldsymbol{r} = \mathbf{0.541}$

d This is a positive correlation that is midway between 0 and 1. There is some evidence to suggest that as age increases so does blood pressure.

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Regression
Exercise A, Question 1

Question:

An NHS trust has the following data on operations.

Number of operating theatres	5	6	7	8
Number of operations carried out per day	25	30	35	40

Which is the independent and which is the dependent variable?

Solution:

number of operating theatres – independent variable

number of operations – dependent variable

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Regression
Exercise A, Question 2

Question:

A park ranger collects data on the number of species of bats in a particular area.

Number of suitable habitats	10	24	28
Number of species	1	2	3

Which is the independent and which is the dependent variable?

Solution:

number of suitable habitats – independent variable

number of species – dependent variable

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Regression
Exercise A, Question 3

Question:

The equation of a regression line in the form $y = a + bx$ is to be found. Given that $S_{xx} = 15$, $S_{xy} = 90$, $\bar{x} = 3$ and $\bar{y} = 15$ work out the values of a and b.

Solution:

$b = \frac{90}{15} = \underline{\mathbf{6}}$

$a = 15 - (6 \times 3) = \underline{\mathbf{-3}}$

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Regression
Exercise A, Question 4

Question:

Given $S_{xx} = 30$, $S_{xy} = 165$, $\bar{x} = 4$ and $\bar{y} = 8$ find the equation of the regression line of y on x.

Solution:

$b = \frac{165}{30} = 5.5$

$a = 8 - (5.5 \times 4) = 8 - 22 = -14$

Equation is: $\underline{\mathbf{y = -14 + 5.5x}}$

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Regression
Exercise A, Question 5

Question:

The equation of a regression line is to be found. The following summary data is given.

$S_{xx} = 40,\qquad S_{xy} = 80,\qquad \bar{x} = 6,\ \ \bar{y} = 12.$

Find the equation of the regression line in the form $y = a + bx$.

Solution:

$b = \frac{80}{40} = 2$

$a = 12 - (2 \times 6) = 0$

Equation is: **$y = 2x$ or $y = 0 + 2x$**

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Regression
Exercise A, Question 6

Question:

Data is collected and summarised as follows.

$\Sigma x = 10$ $\Sigma x^2 = 30$ $\Sigma y = 48$ $\Sigma xy = 140$ $n = 4.$

a Work out $\bar{x}, \bar{y}$, S_{xx} and S_{xy}.

b Find the equation of the regression line of y on x in the form $y = a + bx$.

Solution:

a $\bar{x} = \frac{10}{4} = 2.5$

$\bar{y} = \frac{48}{4} = 12$

$S_{xx} = \Sigma x^2 - \frac{\left(\sum x\right)^2}{n}$

$S_{xx} = 30 - \frac{10 \times 10}{4} = 30 - 25 = \underline{\mathbf{5}}$

$S_{xy} = \Sigma xy - \frac{\sum x \sum y}{n}$

$S_{xy} = 140 - \frac{10 \times 48}{4} = 140 - 120 = \underline{\mathbf{20}}$

b

$b = \frac{20}{5} = 4$

$a = 12 - (4 \times 2.5) = 12 - 10 = 2$

Equation is: $\underline{\boldsymbol{y = 2 + 4x}}$

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Regression
Exercise A, Question 7

Question:

For the data in the table:

x	2	4	5	8	10
y	3	7	8	13	17

a calculate S_{xx} and S_{xy},

b find the equation of the regression line of y on x in the form $y = a + bx$.

Solution:

a $\Sigma x = 29$ $\quad \Sigma x^2 = 209$ $\quad \Sigma y = 48$ $\quad \Sigma xy = 348$

$\bar{x} = 5.8 \quad \bar{y} = 9.6 \quad n = 5$

$$S_{xx} = \Sigma x^2 - \frac{\left(\sum x\right)^2}{n}$$

$$S_{xx} = 209 - \frac{29 \times 29}{5} = 209 - 168.2 = \underline{\mathbf{40.8}}$$

$$S_{xy} = \Sigma xy - \frac{\sum x \sum y}{n}$$

$$S_{xy} = 348 - \frac{29 \times 48}{5} = 348 - 278.4 = \underline{\mathbf{69.6}}$$

b $\bar{x} = \frac{29}{5} = 5.8 \quad \bar{y} = \frac{48}{5} = 9.6$

$$b = \frac{69.6}{40.8} = 1.70(58823)$$

$$a = \bar{y} - b\bar{x} = 9.6 - (1.7058823 \times 5.8) = -0.29(41173)$$

Equation is: $\underline{\boldsymbol{y = -0.294 + 1.71x}}$

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Regression
Exercise A, Question 8

Question:

A field was divided into 12 plots of equal area. Each plot was fertilised with a different amount of fertilizer (h). The yield of grain (g) was measured for each plot. Find the equation of the regression line of g on h in the form $g = a + bh$ given the following summary data.

$\Sigma h = 22.09 \quad \Sigma g = 49.7 \quad \Sigma h^2 = 45.04 \quad \Sigma g^2 = 244.83 \quad \Sigma hg = 97.778 \quad n = 12$

Solution:

$$S_{hh} = 45.04 - \frac{22.09 \times 22.09}{12} = 45.04 - 40.66(4008) = 4.37(5992)$$

$$S_{hg} = 97.778 - \frac{22.09 \times 49.7}{12} = 97.778 - 91.48(9416) = 6.28(8583)$$

$$\bar{h} = \frac{22.09}{12} = 1.84(08333) \quad \bar{g} = \frac{49.7}{12} = 4.14(16666)$$

$$b = \frac{6.288583}{4.375992} = 1.43(70647)$$

$$a = 4.1416666 - (1.4370647 \times 1.8408333) = 1.49(627)$$

Equation is: $\underline{g = 1.50 + 1.44h}$

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Regression
Exercise A, Question 9

Question:

An accountant monitors the number of items produced per month by a company (n) together with the total production costs (p). The table shows these data.

Number of items, *n*, (1000s)	21	39	48	24	72	75	15	35	62	81	12	56
Production costs, *p*, (£1000s)	40	58	67	45	89	96	37	53	83	102	35	75

(You may use $\Sigma n = 540$ $\Sigma n^2 = 30\,786$ $\Sigma p = 780$ $\Sigma p^2 = 56\,936$ $\Sigma np = 41\,444$)

a Calculate S_{nn} and S_{np}.

b Find the equation of the regression line of p on n in the form $p = a + bn$.

Solution:

a) $\Sigma n = 540$ $\Sigma n^2 = 30786$ $\Sigma p = 780$ $\Sigma np = 41444$

$S_{nn} = 30786 - \frac{540 \times 540}{12} = 30786 - 24300 = \underline{\mathbf{6486}}$

$S_{np} = 41444 - \frac{540 \times 780}{12} = 41444 - 35100 = \underline{\mathbf{6344}}$

b) $\bar{n} = 45$ $\bar{p} = 65$

$b = \frac{6344}{6486} = 0.97(81066)$

$a = 65 - (0.9781066 \times 45) = 65 - 44.0148 = 20.98(52)$

Equation is: $\underline{\mathbf{y = 20.98 + 0.978x}}$

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Regression
Exercise A, Question 10

Question:

The relationship between the number of coats of paint applied to a boat and the resulting weather resistance was tested in a laboratory. The data collected are shown in the table.

Coats of paint (x)	1	2	3	4	5
Protection (years) (y)	1.4	2.9	4.1	5.8	7.2

a Calculate S_{xx} and S_{xy}.

b Find the equation of the regression line of y on x in the form $y = a + bx$.

Solution:

a) $\Sigma x = 15$ $\Sigma x^2 = 55$ $\Sigma y = 21.4$ $\Sigma xy = 78.7$ $n = 5$

$$S_{xx} = \Sigma x^2 - \frac{(\sum x)^2}{n}$$

$$S_{xx} = 55 - \frac{15 \times 15}{5} = \underline{\mathbf{10}}$$

$$S_{xy} = \Sigma xy - \frac{\sum x \sum y}{n}$$

$$S_{xy} = 78.7 - \frac{15 \times 21.4}{5} = 78.7 - 64.2 = \underline{\mathbf{14.5}}$$

b

$\bar{x} = 3$ $\bar{y} = 4.28$

$$b = \frac{14.5}{10} = 1.45$$

$a = 4.28 - (1.45 \times 3) = 4.28 - 4.35 = -0.07$

Equation is: $\underline{\boldsymbol{y = -0.07 + 1.45x}}$

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Regression
Exercise B, Question 1

Question:

Given that the coding $p = x + 2$ and $q = y - 3$ has been used to get the regression equation $p + q = 5$ find the equation of the regression line of y on x in the form $y = a + bx$.

Solution:

$(x + 2) + (y - 3) = 5$

$x + y - 1 = 5$

$\boldsymbol{y = 6 - x}$

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Regression
Exercise B, Question 2

Question:

Given the coding $x = p - 10$ and $y = s - 100$ and the regression equation $x = y + 2$ work out the equation of the regression line of s on p.

Solution:

$p - 10 = s - 100 + 2$

$\boldsymbol{s = p + 88}$

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Regression
Exercise B, Question 3

Question:

Given that the coding $g = \frac{x}{3}$ and $h = \frac{y}{4} - 2$ has been used to get the regression equation $h = 6 - 4g$ find the equation of the regression line of y on x.

Solution:

$$\frac{y}{4} - 2 = 6 - 4\left(\frac{x}{3}\right)$$

$3y - 24 = 72 - 16x$ (multiply through by 12)

$\mathbf{3y = 96 - 16x}$ **so** $\mathbf{y = 32 - \frac{16}{3}x}$

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Regression
Exercise B, Question 4

Question:

The regression line of t on s is found by using the coding $x = s - 5$ and $y = t - 10$.

The regression equation of y on x is $y = 14 + 3x$.

Work out the regression line of t on s.

Solution:

$t - 10 = 14 + 3(s - 5)$

$t = 24 + 3s - 15$

$\mathbf{t = 9 + 3s}$

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Regression
Exercise B, Question 5

Question:

A regression line of c on d is worked out using the coding $x = \frac{c}{2}$ and $y = \frac{d}{10}$.

a Given $S_{xy} = 120$, $S_{xx} = 240$, the mean of $x(\bar{x})$ is 5 and the mean of $y(\bar{y})$ is 6, calculate the regression line of y on x.

b Find the regression line of d on c.

Solution:

a $b = \frac{120}{240} = 0.5$

$a = 6 - 0.5 \times 5 = 3.5$

$\mathbf{y = 3.5 + 0.5x}$

b $\frac{d}{10} = 3.5 + 0.5 \times \frac{c}{2}$ (multiply by 10)

$\mathbf{d = 35 + 2.5c}$

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Regression
Exercise B, Question 6

Question:

Some data on heights (h) and weights (w) were collected. The results were coded such that $x = \frac{h-8}{2}$ and $y = \frac{w}{5}$. The coded results are shown in the table.

x	1	5	10	16	17
y	9	12	16	21	23

a Calculate S_{xy} and S_{xx} and use them to find the equation of the regression line of y on x.

b Find the equation of the regression line of w on h.

Solution:

a $\Sigma x = 49$ $\Sigma x^2 = 671$ $\Sigma y = 81$ $\Sigma xy = 956$

$S_{xy} = 956 - \frac{49 \times 81}{5} = 956 - 793.8 = \mathbf{162.2}$

$S_{xx} = 671 - \frac{\left(\sum x\right)^2}{5} = 671 - \frac{2401}{5} = 671 - 480.2 = \mathbf{190.8}$

$\bar{y} = 16.2$ $\bar{x} = 9.8$

$b = \frac{162.2}{190.8} = 0.85 \ldots \ldots ..$

$a = 16.2 - (0.85 \times 9.8) = 16.2 - 8.33 \ldots.. = 7.87$

Equation of y on x is: $\mathbf{y = 7.87 + 0.85x}$

b $\frac{w}{5} = 7.87 + 0.85\left(\frac{h-8}{2}\right)$ (multiply by 5)

$w = 39.35 + 2.125(h - 8)$

$w = 39.35 + 2.125h - 17$

$\mathbf{w = 22.35 + 2.125h}$

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Regression
Exercise C, Question 1

Question:

Given the regression line $y = 24 - 3x$ find the value of y when x is 6.

Solution:

$y = 24 - (3 \times 6) = 6$

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Regression
Exercise C, Question 2

Question:

The regression line for the weight (w) in grams on the volume (v) in cm^3 for a sample of small marbles is $w = 300 + 12v$.

Calculate the weight when the volume is 7 cm^3.

Solution:

$w = 300 + (12 \times 7) = 384$

Weight is **384** grams

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Regression
Exercise C, Question 3

Question:

a State what is meant by extrapolation.

b State what is meant by interpolation.

Solution:

a Extrapolation means using the regression line to estimate outside the range of the data collected. It can be unreliable.

b Interpolation means using the regression line to estimate within the range of the data collected. It is usually reasonably reliable.

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Regression
Exercise C, Question 4

Question:

12 children between the ages (x) of five and 11 years were asked how much pocket money (y) they were given each week. The equation for the regression line of y on x was found to be $y = 2x - 8$.

a Use the equation to estimate the amount of money a seven year old would get. State, with a reason, whether or not this is a reliable estimate.

b One of the children suggested that this equation must be wrong since it showed that a three year old would get a negative amount of pocket money. Explain why this has happened.

Solution:

a $y = (2 \times 7) - 8 = 6$

A 7 year old would get **£6.00**. This is a reasonable estimate as 7 years is within the range of ages asked. It is interpolation.

b This would involve extrapolation, which may not be reliable. Three years old is outside the range of ages asked. A three year old is probably not given pocket money.

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Regression
Exercise C, Question 5

Question:

The pulse rates (y) of 10 people were measured after doing different amounts of exercise (x) for between two and 10 minutes. The regression equation $y = 0.75x + 72$ refers to these data. The equation seems to suggest that someone doing 30 minutes of exercise would have a pulse rate of 94.5. State whether or not this is sensible. Give a reason for your answer.

Solution:

This is not a sensible estimate since the data collected only covers 2 to 10 minutes. The answer 94.5 involves extrapolation. In fact the pulse rate can not keep rising. 30 minutes is a long way outside the limits of the data.

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Regression
Exercise C, Question 6

Question:

Over a period of time the sales, (y) in thousands, of 10 similar text books and the amount, (x) in £ thousands, spent on advertising each book were recorded. The greatest amount spent on advertising was £4.4 thousand, and the least amount was £0.75 thousand.

An equation of the regression line for y on x was worked out for the data.

The equation was $y = 0.93 + 1.1x$.

a Use the equation to estimate the sales of a text book if the amount spent on advertising is to be set at £2.65 thousand. State, with a reason, whether or not this is a reliable estimate.

b Use the equation to estimate the sales of the book if the amount to be spent on advertising is £8000. State, with a reason, whether or not this is a reliable estimate.

c Explain what the value 1.1 tells you about the relationship between the sales of books and the amount spent on advertising.

d Interpret the meaning of the figure 0.93.

Solution:

a $y = 0.93 + (1.1 \times 2.65) =$ **3.845** thousands. This is reasonably reliable since £2.65 thousand is within the range of the collected data. It involves interpolation.

b $y = 0.93 + (1.1 \times 8) = 9.73$ thousands. This may be unreliable since £8 thousand is outside the range of the data collected. It involves extrapolation.

c 1.1 thousand is the number of extra books sold for each £1 thousand spent on advertising.

d 0.93 thousand is the number of books likely to be sold if there is no money spent on advertising. This is only just outside the range of values so it is a reasonably reliable estimate.

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Regression
Exercise C, Question 7

Question:

Research was done to see if there is a relationship between finger dexterity and the ability to do work on a production line. The data is shown in the table.

Dexterity Score (x)	2.5	3	3.5	4	5	5	5.5	6.5	7	8
Productivity (y)	80	130	100	220	190	210	270	290	350	400

The equation of the regression line for these data is $y = -59 + 57x$

a Use the equation to estimate the productivity of someone with a dexterity of 6.

b State the contextual meaning of the figure 57.

c State, giving in each case a reason, whether or not it would be reasonable to use this equation to work out the productivity of someone with dexterity of:
i 2 **ii** 14.

Solution:

a $y = -59 + (57 \times 6) = 283$

b 57 is the gradient. For every one rise in the dexterity production rises by 57.

c i) 2 may be a little unreliable since it lies just outside the values in the table. It would involve extrapolation.

ii) 14 could be very unreliable as it lies well outside the range of the values in the table. It would involve extrapolation.

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Regression
Exercise C, Question 8

Question:

A regression line of $y = 5 + 3x$ is found using 10 data sets. Another piece of data is recorded which when put on a scatter diagram with the original data proves to be well above the regression line for all the other data. Write down whether or not the regression equation would change if this piece of data were included in the calculation.

Solution:

The equation of the regression line would change. The line is likely to tilt.

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Regression
Exercise D, Question 1

Question:

A metal rod was found to increase in length as it was heated. The temperature (t) and the increase in length (l mm) were measured at intervals between 30°C and 400°C degrees. The regression line of l on t was found to be $l = 0.009t - 0.25$.

a Find the increase in length for a temperature of 300°C.

b Find the increase in length for a temperature of 530°C.

c Write down, with reasons, why the answer to **a** might be reliable and the answer **b** unreliable.

Solution:

a $l = (0.009 \times 300) - 0.25.$ = **2.45mm**

b $l = (0.009 \times 530) - 0.25.$ = **4.52mm**

c a is likely to be a reasonable estimate since it involves interpolation. 300°C is within the range of the data covered.

b involves extrapolation so is likely to be unreliable. 530°C is outside the range of the data used. The metal might well melt before this temperature is reached.

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Regression
Exercise D, Question 2

Question:

Two variables s and t are thought to be connected by a law of the form $t = a + bs$, where a and b are constants.

a Use the summary data:

$\Sigma s = 553$ $\Sigma t = 549$ $\Sigma st = 31\ 185$ $n = 12$ $\bar{s} = 46.0833$
$\bar{t} = 45.75$ $S_{ss} = 6193$

to work out the regression line of t on s.

b Find the value of t when s is 50

Solution:

a $S_{st} = 31185 - \frac{553 \times 549}{12} = 31185 - 25299.75 = 5885.25$

$b = \frac{5885.25}{6193} = 0.959(3067)$

$a = 45.75 - (0.950\ldots \times 46.083)$

$a = 45.75 - 43.792(988)$

$a = 1.957$

$t = 1.96 + 0.95s$

b $t = 1.96 + (0.95 \times 50)$

$t = 49.46$

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Regression
Exercise D, Question 3

Question:

A biologist recorded the breadth (x cm) and the length (y cm) of 12 beech leaves. The data collected can be summarised as follows.

$\Sigma x^2 = 97.73$ $\Sigma x = 33.1$ $\Sigma y = 66.8$ $\Sigma xy = 195.94$

a Calculate S_{xx} and S_{xy}.

b Find the equation of the regression line of y on x in the form $y = a + bx$.

c Predict the length of a beech leaf that has a breadth of 3.0 cm.

Solution:

a $S_{xx} = 97.73 - \frac{33.1 \times 33.1}{12} = 97.73 - 91.30 = \mathbf{6.43}$

$S_{xy} = 195.94 - \frac{33.1 \times 66.8}{12} = 195.94 - 184.26 = \mathbf{11.68}$

b $\bar{x} = \frac{33.1}{12} = 2.76 \bar{y} = \frac{66.8}{12} = 5.57$

$b = \frac{11.68}{6.43} = 1.82$

$a = 5.57 - (1.82 \times 2.76) = 5.57 - 5.02 = 0.55$

Equation is: $\boldsymbol{y = 0.55 + 1.82x}$

c length = 0.55 + (1.82 × 3) = **6.01** cm

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Regression
Exercise D, Question 4

Question:

Energy consumption is claimed to be a good predictor of Gross National Product. An economist recorded the energy consumption (x) and the Gross National Product (y) for eight countries. The data are shown in the table.

Energy Consumption x	3.4	7.7	12.0	75	58	67	113	131
Gross National Product y	55	240	390	1100	1390	1330	1400	1900

a Calculate S_{xy} and S_{xx}.

b Find the equation of the regression line of y on x in the form $y = a + bx$.

c Estimate the Gross National Product of a country that has an energy consumption of 100.

d Estimate the energy consumption of a country that has a Gross National Product of 3500.

e Comment on the reliability of your answer to **d**.

Solution:

a

$\Sigma x^2 = 43622.85$ $\Sigma x = 467.1$ $\Sigma y = 7805$ $\Sigma xy = 666045$

$S_{xx} = 43622.85 - \frac{467.1 \times 467.1}{8} = 43622.85 - 27272.80 \ldots. = \mathbf{16350.048\ldots}$

$S_{xy} = 666045 - \frac{467.1 \times 7805}{8} = 666045 - 455714.43 = \mathbf{210330.56\ldots}$

b $\bar{x} = 58.3875$ $\bar{y} = 975.625$

$b = \frac{210330.56\ldots}{16350.04\ldots} = 12.864\ldots\ldots$

$a = 975.625 - (12.86\ldots \times 58.3875) = 975.625 - 751.10939 = 224.515\ldots\ldots$

Equation is: $\mathbf{y = 224.52 + 12.86x}$

c Gross National product = 224.515.. + (12.86… × 100) = **1510.5**

d $3500 = 224.515\ldots + 12.864..x$

$12.864..x = 3500 - 224.515\ldots$

$x = \frac{3275.48..}{12.864..} = 254.6$

Energy consumption = **254.6**

e This answer is likely to be unreliable as it involves extrapolation. 3500 is well outside the limits of the data set used.

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Regression
Exercise D, Question 5

Question:

In an environmental survey on the survival of mammals the tail length t (cm) and body length m (cm) of a random sample of six small mammals of the same species were measured.

These data are coded such that $x = \frac{m}{2}$ and $y = t - 2$.

The data from the coded records are summarised below.

$\Sigma y = 13.5$ $\quad\Sigma x = 25.5$ $\quad\Sigma xy = 84.25$ $\quad S_{xx} = 59.88$

a Find the equation of the regression line of y on x in the form $y = ax + b$.

b Hence find the equation of the regression line of t on m.

c Predict the tail length of a mammal that has a body length of 10 cm.

Solution:

a $S_{xy} = 84.25 - \frac{25.5 \times 13.5}{6} = 84.25 - 57.375 = 26.875$

$\bar{x} = 4.25 \quad \bar{y} = 2.25$

$b = \frac{26.875}{59.88} = 0.4488 \ldots ..$

$a = 2.25 - (0.4488\ldots \times 4.25) = 2.25 - 1.9074 = 0.3425\ldots$

Equation is $\boldsymbol{y = 0.343 + 0.449x}$

b $t - 2 = 0.343.. + 0.448..\left(\frac{m}{2}\right)$

$\boldsymbol{t = 2.343 + 0.224m}$

c tail length = 2.343 + (0.224 × 10) = **4.58 cm**

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Regression
Exercise D, Question 6

Question:

A health clinic counted the number of breaths per minute (r) and the number of pulse beats (p) per minute for 10 people doing various activities. The data are shown in the table.

The data are coded such that $x = \frac{r-10}{2}$ and $y = \frac{p-50}{2}$.

x	3	5	5	7	8	9	9	10	12	13
y	4	9	10	11	17	15	17	19	22	27

(You may use $\Sigma x = 81$ $\Sigma x^2 = 747$ $\Sigma y = 151$ $\Sigma y^2 = 2695$ $\Sigma xy = 1413$.)

a Calculate S_{xy} and S_{xx}.

b Find the equation of the regression line of y on x in the form $y = a + bx$.

c Find the equation of the regression line for p on r.

d Estimate the number of pulse beats per minute for someone who is taking 22 breaths per minute.

e Comment on the reliability of your answer to **e**.

Solution:

a

$\Sigma x = 81$ $\Sigma y = 151$ $\Sigma x^2 = 747$ $\Sigma xy = 1413$ $\bar{x} = 8.1$ $\bar{y} = 15.1$

$S_{xx} = 747 - \frac{81 \times 81}{10} = 747 - 656.1 = \mathbf{90.9}$

$S_{xy} = 1413 - \frac{81 \times 151}{10} = 1413 - 1223.1 = \mathbf{189.9}$

b $b = \frac{189.9}{90.9} = 2.089\ldots..$

$\bar{x} = 8.1$ $\bar{y} = 15.1$

$\boldsymbol{a} = 15.1 - (2.089\ldots \times 8.1) = 15.1 - 16.9217\ldots. = -1.82\ldots$

Equation is: $\mathbf{y = -1.82 + 2.09x}$

c $\frac{p-50}{2} = -1.82 + 2.09\left(\frac{r-10}{2}\right)$ (multiply by 2)

$p - 50 = -3.64 + 2.09r - 20.9$

$\mathbf{p = 25.46 + 2.09r}$

d Pulse Beats = 25.46 + (2.09 × 22) = **71.44**

e The answer to d is reasonably reliable since it involves interpolation. 22 is within the range of the data set used.

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Regression
Exercise D, Question 7

Question:

A farm food supplier monitors the number of hens kept (x) against the weekly consumption of hen food (y kg) for a sample of 10 small holders. He records the data and works out the regression line for y on x to be $y = 0.16 + 0.79x$.

a Write down a practical interpretation of the figure 0.79.

b Estimate the amount of food that is likely to be needed by a small holder who has 30 hens.

c If food costs £12 for a 10 kg bag estimate the weekly cost of feeding 50 hens.

Solution:

a 0.79 kg is the average amount of food consumed in 1 week by 1 hen.

b $y = 0.16 + 0.79 \times 30 =$ **23.86 kg**

c Cost = (0.16 + 0.79 × 50)£12 = **£475.92**

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Regression
Exercise D, Question 8

Question:

Water voles are becoming very rare; they are often confused with water rats. A naturalist society decided to record details of the water voles in their area. The members measured the weight (y) to the nearest 10 grams, and the body length (x) to the nearest millimetre, of eight active healthy water voles. The data they collected are in the table.

Body Length (x) mm	140	150	170	180	180	200	220	220
Weight (y) grams	150	180	190	220	240	290	300	310

a Draw a scatter diagram of these data.

b Give a reason to support the calculation of a regression line for these data.

c Use the coding $l = \frac{x}{10}$ and $w = \frac{y}{10}$ to work out the regression line of w on l.

d Find the equation of the regression line for y on x.

e Draw the regression line on the scatter diagram.

f Use your regression line to calculate an estimate for the weight of a water vole that has a body length of 210 mm. Write down, with a reason, whether or not this is a reliable estimate.

The members of the society remove any water voles that seem unhealthy from the river and take them into care until they are fit to be returned.

They find three water voles on one stretch of river which have the following measurements.

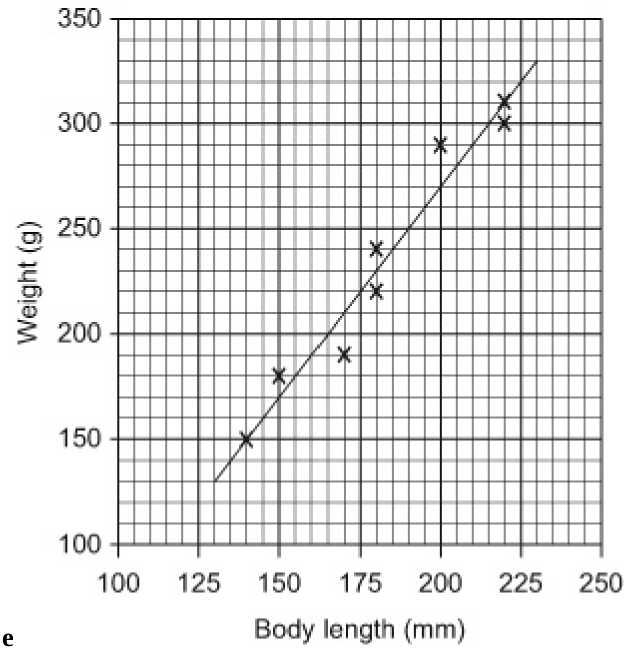
A: Weight 235 gm and body length 180 mm

B: Weight 180 gm and body length 200 mm

C: Weight 195 gm and body length 220 mm

g Write down, with a reason, which of these water voles were removed from the river.

Solution:



a & e

b There appears to be a linear relationship between body length and body weight.

c

l	14	15	17	18	18	20	22	22
w	15	18	19	22	24	29	30	31

$\Sigma l^2 = 2726$ $\Sigma l = 146$ $\Sigma w = 188$ $\Sigma lw = 3553$

$\bar{l} = 18.25$ $\bar{w} = 23.5$

$$S_{ll} = 2726 - \frac{146 \times 146}{8} = 2726 - 2664.5 = 61.5$$

$$S_{lw} = 3553 - \frac{146 \times 188}{8} = 3553 - 3431 = 122$$

$$b = \frac{122}{61.5} = 1.9837 \ldots \ldots$$

$a = 23.5 - (1.9837 \ldots \times 18.25) = 23.5 - 36.2032\ldots.. = -12.70\ldots$

Equation is: $\boldsymbol{w = -12.7 + 1.98l}$

d $\frac{y}{10} = -12.7 + \left(1.98 \times \frac{x}{10}\right)$ (multiply through by 10)

$\boldsymbol{y = -127 + 1.98x}$

e ***See diagram***

f Tail length = − 127 + 1.98 × 210 = **288.8mm**

This is a reliable estimate since it involves interpolation. 210 is within the range of the data.

g B and C are both underweight so were probably removed from the river. A is slightly overweight so was probably left in the river.

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Regression
Exercise D, Question 9

Question:

A mail order company pays for postage of its goods partly by destination and partly by total weight sent out on a particular day. The number of items sent out and the total weights were recorded over a seven day period. The data are shown in the table.

Number of items (*n*)	10	13	22	15	24	16	19
Weight in kg (*w*)	2800	3600	6000	3600	5200	4400	5200

a Use the coding $x = n - 10$ and $y = \frac{w}{400}$ to work out S_{xy} and S_{xx}.

b Work out the equation of the regression line for y on x.

c Work out the equation of the regression line for w on n.

d Use your regression equation to estimate the weight of 20 items.

e State why it would be unwise to use the regression equation to estimate the weight of 100 items.

Solution:

a

Coded number x	0	3	12	5	14	6	9
Coded weight y	7	9	15	9	13	11	13

$\Sigma x = 49$ $\Sigma x^2 = 491$ $\Sigma y = 77$ $\Sigma xy = 617$

$S_{xy} = 617 - \frac{49 \times 77}{7} = 617 - 539 = \mathbf{78}$

$S_{xx} = 491 - \frac{49^2}{7} = 491 - 343 = \mathbf{148}$

b $\bar{y} = 11$ $\bar{x} = 7$

$b = \frac{78}{148} = 0.5270 \ldots \ldots$

$a = 11 - (0.5270\ldots \times 7) = 11 - 3.6891\ldots\ldots = 7.3108\ldots$

Equation is: $\boldsymbol{y = 7.31 + 0.53\,x}$ (to two decimal places)

c $\frac{w}{400} = 7.31 + 0.527(n - 10)$ (multiply by 400)

$w = 2924 + 210.8(n - 10)$

$w = 2924 + 210.8\,n - 2108$

Equation is: $\boldsymbol{w = 816 + 211n}$

d $w = 816 + 211 \times 20 =$ **5036**

e This is far outside the range of values. This is extrapolation.

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Discrete random variables
Exercise A, Question 1

Question:

Write down whether or not each of the following is a discrete random variable.

Give a reason for your answer.

a The average lifetime of a battery.

b The number of days in a week.

c The number of moves it takes to complete a game of chess.

Solution:

i) This **is not** a discrete random variable. Time is a continuous variable.

ii) This **is not** a discrete random variable. It is always 7, so does not vary

iii) This **is** a discrete random variable. It is always a whole number but it does vary.

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Discrete random variables
Exercise A, Question 2

Question:

A fair die is thrown four times and the number of times it falls with a 6 on the top, Y, is noted. Write down all the possible values of y.

Solution:

y: 0 1 2 3 4

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Discrete random variables
Exercise A, Question 3

Question:

A bag contains two discs with the number 2 on them and two discs with the number 3 on them. A disc is drawn at random from the bag and the number noted. The disc is returned to the bag. A second disc is then drawn from the bag and the number noted.

a Write down the sample space.

The discrete random variable X is the sum of the two numbers.

b Write down the probability distribution for X.

c Write down the probability function for X.

Solution:

a S = (2,2), (2,3), (3,2), (3,3)

b

x	4	5	6
$P(X = x)$	$\frac{1}{4}$	$\frac{1}{2}$	$\frac{1}{4}$

c $P(X = x) = \begin{cases} \frac{1}{4} & x = 4,6, \\ \frac{1}{2} & x = 5. \end{cases}$

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Discrete random variables
Exercise A, Question 4

Question:

A discrete random variable X has the following probability distribution:

x	1	2	3	4
$\mathbf{P}(X = x)$	$\frac{1}{3}$	$\frac{1}{3}$	k	$\frac{1}{4}$

Find the value of k.

Solution:

$\frac{1}{3} + \frac{1}{3} + \frac{1}{4} + k = 1$

$k = 1 - \left(\frac{1}{3} + \frac{1}{3} + \frac{1}{4}\right)$

$= 1 - \frac{11}{12} = \mathbf{\frac{1}{12}}$

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Discrete random variables
Exercise A, Question 5

Question:

The random variable X has a probability function

$P(X = x) = kx \qquad x = 1, 2, 3, 4.$

Show that $k = \frac{1}{10}$.

Solution:

x	1	2	3	4
$P(X = x)$	k	$2k$	$3k$	$4k$

$k + 2k + 3k + 4k = 1$

$10k = 1$

$\boldsymbol{k = \frac{1}{10}}$

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Discrete random variables
Exercise A, Question 6

Question:

The random variable X has the probability function

$P(X = x) = \frac{x-1}{10}$ $\quad x = 1, 2, 3, 4, 5.$

Construct a table giving the probability distribution of X.

Solution:

x	1	2	3	4	5
$P(X = x)$	0	$\frac{1}{10}$	$\frac{1}{5}$	$\frac{3}{10}$	$\frac{2}{5}$

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Discrete random variables
Exercise A, Question 7

Question:

The random variable X has a probability function

$$P(X = x) = \begin{cases} kx & x = 1,3 \\ k(x-1) & x = 2,4 \end{cases}$$

where k is a constant.

a Find the value of k.

b Construct a table giving the probability distribution of X.

Solution:

a

x	1	2	3	4
$P(X = x)$	k	k	$3k$	$3k$

Using the fact that the probabilities add up to 1:

$$k + k + 3k + 3k = 1$$

$$8k = 1$$

$$\boldsymbol{k = \frac{1}{8}}$$

b

x	**1**	**2**	**3**	**4**
$P(X = x)$	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{3}{8}$	$\frac{3}{8}$

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Discrete random variables
Exercise A, Question 8

Question:

The discrete random variable X has a probability function

$$P(X = x) = \begin{cases} 0.1 & x = -2, -1 \\ \beta & x = 0,1 \\ 0.2 & x = 2 \end{cases}$$

a Find the value of β.

b Construct a table giving the probability distribution of X.

Solution:

x	-2	-1	0	1	2
$P(X = x)$	0.1	0.1	β	β	0.2

The probabilities add up to 1.

$$0.1 + 0.1 + \beta + \beta + 0.2 = 1$$

$$2\beta = 1 - 0.4$$

$$\beta = \mathbf{0.3}$$

b

x	-2	-1	0	1	2
$P(X = x)$	0.1	0.1	0.3	0.3	0.2

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Discrete random variables
Exercise A, Question 9

Question:

A discrete random variable has the probability distribution shown in the table below.

x	0	1	2
$\mathbf{P}(X = x)$	$\frac{1}{4} - a$	a	$\frac{1}{2} + a$

Find the value of a.

Solution:

$$\frac{1}{4} - a + a + \frac{1}{2} + a = 1$$

$$a = 1 - \frac{3}{4} = \frac{1}{4}$$

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Discrete random variables
Exercise B, Question 1

Question:

A discrete random variable X has probability distribution

x	0	1	2	3	4	5
$P(X = x)$	0.1	0.1	0.3	0.3	0.1	0.1

a Find the probability that $X < 3$.

b Find the probability that $X > 3$.

c Find the probability that $1 < X < 4$.

Solution:

a $P(X < 3) = P(0) + P(1) + P(2) = 0.1 + 0.1 + 0.3 = \mathbf{0.5}$

b $P(X > 3) = P(4) + P(5) = 0.1 + 0.1 = \mathbf{0.2}$

c $P(1 < X < 4) = P(2) + P(3) = 0.3 + 0.3 = \mathbf{0.6}$

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Discrete random variables
Exercise B, Question 2

Question:

A discrete random variable X has probability distribution

x	0	1	2	3
$\mathrm{P}(X = x)$	$\frac{1}{8}$	$\frac{1}{4}$	$\frac{1}{2}$	$\frac{1}{8}$

Find

a $\mathrm{P}(1 < X \le 3)$,　　　**b** $\mathrm{P}(X < 2)$.

Solution:

a $\mathrm{P}(1 < X \le 3) = \mathrm{P}(2) + \mathrm{P}(3) = \frac{1}{2} + \frac{1}{8} = \frac{5}{8}$

b $\mathrm{P}(X < 2) = \mathrm{P}(0) + \mathrm{P}(1) = \frac{1}{8} + \frac{1}{4} = \frac{3}{8}$

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Discrete random variables
Exercise B, Question 3

Question:

A discrete random variable X has a probability distribution

x	1	2	3	4	5	6
$\mathbf{P}(X = x)$	0.1	0.1	0.15	0.25	0.3	0.1

a Draw up a table to show the cumulative distribution function F(x).
b Write down F(5). **c** Write down F(2.2).

Solution:

a

x	**1**	**2**	**3**	**4**	**5**	**6**
$\mathbf{F}(x_0)$	**0.1**	**0.2**	**0.35**	**0.60**	**0.9**	**1**

b F(5) = **0.9**

c F(2.2) = **0.2**

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Discrete random variables
Exercise B, Question 4

Question:

A discrete random variable has a cumulative distribution function F(x) given in the table.

x	0	1	2	3	4	5	6
F(x)	0	0.1	0.2	0.45	0.5	0.9	1

a Draw up a table to show the probability distribution X.
b Write down $P(X < 5)$. **c** Find $P(2 \le X < 5)$.

Solution:

a

x	**1**	**2**	**3**	**4**	**5**	**6**
$P(X = x)$	**0.1**	**0.1**	**0.25**	**0.05**	**0.4**	**0.1**

b $P(X < 5) = 0.5$

c $P(2 \le X < 5) = 0.1 + 0.25 + 0.05 = 0.4$

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Discrete random variables
Exercise B, Question 5

Question:

5 The random variable X has a probability function

$$P(X = x) = \begin{cases} kx & x = 1,3,5 \\ k(x-1) & x = 2,4,6 \end{cases}$$

where k is a constant.

a Find the value of k.

b Draw a table giving the probability distribution of X.

c Find $P(2 \le X < 5)$.

d Find F(4).

e Find F(1.6).

Solution:

a

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

$k + k + 3k + 3k + 5k + 5k = 1$

$$k = \frac{1}{18}$$

b

x	**1**	**2**	**3**	**4**	**5**	**6**
P(*X* = *x*)	$\frac{1}{18}$	$\frac{1}{18}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{5}{18}$	$\frac{5}{18}$

c $P(2 \le X < 5) = P(2) + P(3) + P(4) = \frac{1}{18} + \frac{1}{6} + \frac{1}{6} = \frac{7}{18}$

d Remember F means the cumulative function

$$F(4) = 1 - (P(6) + P(5)) = 1 - \left(\frac{5}{18} + \frac{5}{18}\right) = \frac{8}{18} \text{ or } \frac{4}{9}$$

(This could also be done by adding P(1) P(2) P(3) and P(4).)

e 1.6 lies below 2 but above 1. Because this is a **discrete** random variable F(1.6) is the same as F(1) which is $\frac{1}{18}$

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Discrete random variables
Exercise B, Question 6

Question:

The discrete random variable X has the probability function

$$P(X = x) = \begin{cases} 0.1 & x = -2, -1 \\ \alpha & x = 0, 1 \\ 0.3 & x = 2 \end{cases}$$

a Find the value of α

b Draw a table giving the probability distribution of X.

c Write down the value of F(0.3).

Solution:

a

x	−2	−1	0	1	2
$P(X = x)$	0.1	0.1	α	α	0.3

$$\begin{aligned} 0.1 + 0.1 + \alpha + \alpha + 0.2 &= 1 \\ 2\alpha &= 0.6 \\ \alpha &= 0.3 \end{aligned}$$

b

x	−2	−1	0	1	2
$P(X = x)$	0.1	0.1	0.25	0.25	0.3

c F(0.3) = F (0) = 0.1 + 0.1 + 0.25 = 0.45

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Discrete random variables
Exercise B, Question 7

Question:

The discrete random variable X has a cumulative distribution function F(x) defined by

$$F(x)=\begin{cases}0 & x=0\\ \frac{1+x}{6} & x=1,2,3,4,5\\ 1 & x>5\end{cases}$$

a Find $P(X \le 4)$.

b Show that $P(X = 4)$ is $\frac{1}{6}$.

c Find the probability distribution for X.

Solution:

a

x	1	2	3	4	5
F(x)	$\frac{2}{6}$	$\frac{3}{6}$	$\frac{4}{6}$	$\frac{5}{6}$	1

$P(X \le 4) = \frac{5}{6}$

b $\mathbf{P(X = 4) = \frac{5}{6} - \frac{4}{6} = \frac{1}{6}}$

c

x	1	2	3	4	5
$P(X = x)$	$\frac{2}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Discrete random variables
Exercise B, Question 8

Question:

The discrete random variable X has a cumulative distribution function F(x) defined by $F(x) = \begin{cases} 0 & x = 0 \\ \frac{(x+k)^2}{16} & x = 1, 2 \text{ and } 3 \\ 1 & x > 3 \end{cases}$

a Find the value of k.

b Find the probability distribution for X.

Solution:

a $\frac{(x+k)^2}{16} = 1$ when $x = 3$

$\frac{(3+k)^2}{16} = 1$

$(3 + k)^2 = 16$

$3 + k = \pm 4$

$\boldsymbol{k} = \mathbf{1}$ (negative probabilities do not exist)

b

x	1	2	3
F(x)	$\frac{4}{16}$	$\frac{9}{16}$	1

So Probability distribution is

x	1	2	3
P(X = x)	$\frac{4}{16}$	$\frac{5}{16}$	$\frac{7}{16}$

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Discrete random variables
Exercise C, Question 1

Question:

Find E(X) and E(X^2) for the following distributions of x.

a

x	2	4	6	8
P(X = x)	0.3	0.3	0.2	0.2

b

x	1	2	3	4
P(X = x)	0.1	0.4	0.1	0.4

Solution:

a $\mathrm{E}(X) = (2 \times 0.3) + (4 \times 0.3) + (6 \times 0.2) + (8 \times 0.2)$

$= 0.6 + 1.2 + 1.2 + 1.6 = \mathbf{4.6}$

$\mathrm{E}(X^2) = (4 \times 0.3) + (16 \times 0.3) + (36 \times 0.2) + (64 \times 0.2)$

$= 1.2 + 4.8 + 7.2 + 12.8 = \mathbf{26}$

b $\mathrm{E}(X) = (1 \times 0.1) + (2 \times 0.4) + (3 \times 0.1) + (4 \times 0.4)$

$= 0.1 + 0.8 + 0.3 + 1.6 = \mathbf{2.8}$

$\mathrm{E}(X^2) = (1 \times 0.1) + (4 \times 0.4) + (9 \times 0.1) + (16 \times 0.4)$

$= 0.1 + 1.6 + 0.9 + 6.4 = \mathbf{9}$

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Discrete random variables
Exercise C, Question 2

Question:

A biased die has the probability distribution

x	1	2	3	4	5	6
$\mathbf{P}(X = x)$	0.1	0.1	0.1	0.2	0.4	0.1

Find $E(X)$ and $E(X^2)$.

Solution:

$$E(X) = (1 \times 0.1) + (2 \times 0.1) + (3 \times 0.1) + (4 \times 0.2) + (5 \times 0.4) + (6 \times 0.1)$$

$$= 0.1 + 0.2 + 0.3 + 0.8 + 2.0 + 0.6 = \mathbf{4}$$

$$E(X^2) = (1 \times 0.1) + (4 \times 0.1) + (9 \times 0.1) + (16 \times 0.2) + (25 \times 0.4) + (36 \times 0.1)$$

$$= 0.1 + 0.4 + 0.9 + 3.2 + 10 + 3.6 = \mathbf{18.2}$$

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Discrete random variables
Exercise C, Question 3

Question:

The random variable X has a probability function

$$P(X = x) = \begin{cases} \frac{1}{x} & x = 2,3,6 \\ 0 & \text{all other values} \end{cases}$$

a Construct a table giving the probability distribution of X.

b Work out E(X) and E(X^2).

c State with a reason whether or not $(E(X))^2 = E(X^2)$.

Solution:

a

x	2	3	6
$P(X = x)$	$\frac{1}{2}$	$\frac{1}{3}$	$\frac{1}{6}$

b

$$E(X) = \left(2 \times \frac{1}{2}\right) + \left(3 \times \frac{1}{3}\right) + \left(6 \times \frac{1}{6}\right)$$

$$= 1 + 1 + 1 = \mathbf{3}$$

$$E(X^2) = \left(4 \times \frac{1}{2}\right) + \left(9 \times \frac{1}{3}\right) + \left(36 \times \frac{1}{6}\right)$$

$$= 2 + 3 + 6 = \mathbf{11}$$

c

$(E(X))^2 = 3 \times 3 = 9$

$E(X^2) = 11$

Therefore $(E(X))^2$ does not equal $E(X^2)$

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Discrete random variables
Exercise C, Question 4

Question:

Two coins are tossed 50 times. The number of heads is noted each time.

a Construct a probability distribution table for the number of heads when the two coins are tossed once, assuming that the two coins are unbiased.

b Work out how many times you would expect to get 0, 1 or 2 heads.

The following table shows the actual results.

Number of heads (*h*)	0	1	2
Frequency (*f*)	7	22	21

c State whether or not the actual frequencies obtained support the statement that the coins are unbiased. Give a reason for your answer.

Solution:

a

Number of heads (h)	0	1	2
$P(H = h)$	0.25	0.5	0.25

b

$0.25 \times 50 = 12.5$

$0.5 \times 50 = 25$

We would expect to get **1 head 25 times** and **0 or 2 heads 12.5 times each**.

c The coins would appear to be biased. There were far too many times when 2 heads appeared and not enough when 0 heads appeared.

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Discrete random variables
Exercise C, Question 5

Question:

The random variable X has the following probability distribution.

x	1	2	3	4	5
$P(X = x)$	0.1	a	b	0.2	0.1

Given $E(X) = 2.9$ find the value of a and the value of b.

Solution:

x	1	2	3	4	5
$P(X = x)$	0.1	a	b	0.2	0.1

The probabilities add up to 1 so

$0.1 + a + b + 0.2 + 0.1 = 1$

$a + b = 0.6$ **(1)**

and

$2.9 = (1 \times 0.1) + (2 \times a) + (3 \times b) + (4 \times 0.2) + (5 \times 0.1)$

$2.9 = 0.1 + 2a + 3b + 0.8 + 0.5$

$2a + 3b = 1.5$ **(2)**

multiply **(1)** by **(2)**

$2a + 2b = 1.2$ **(3)**

(2) minus **(3)**

$b =$ **0.3**

so from 3 $2a + 0.6 = 1.2$

$a =$ **0.3**

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Discrete random variables
Exercise C, Question 6

Question:

A fair spinner with equal sections numbered 1 to 5 is thrown 500 times. Work out how many times it can be expected to land on 3.

Solution:

x	1	2	3	4	5
$P(X = x)$	0.2	0.2	0.2	0.2	0.2

$0.2 \times 500 = 100$

We can expect it to land on 3 100 times.

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Discrete random variables
Exercise D, Question 1

Question:

For the following probability distribution

x	-1	0	1	2	3
P(X = x)	$\frac{1}{5}$	$\frac{1}{5}$	$\frac{1}{5}$	$\frac{1}{5}$	$\frac{1}{5}$

a write down E(X).

b find Var(X).

Solution:

a By symmetry E(X) = 1

b Var X = E(X^2) − (E(X))^2

$$E(X^2) = \frac{1}{5} + 0 + \frac{1}{5} + \frac{4}{5} + \frac{9}{5} = 3$$

$(E(X))^2 = 1^2 = 1$

Var X = 3 − 1 = **2**

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Discrete random variables
Exercise D, Question 2

Question:

Find the expectation and variance of each of the following distributions of X.

a

x	1	2	3
P($X = x$)	$\frac{1}{3}$	$\frac{1}{2}$	$\frac{1}{6}$

b

x	−1	0	1
P($x = x$)	$\frac{1}{4}$	$\frac{1}{2}$	$\frac{1}{4}$

c

x	−2	−1	1	2
P($X = x$)	$\frac{1}{3}$	$\frac{1}{3}$	$\frac{1}{6}$	$\frac{1}{6}$

Solution:

a Mean $= \mathrm{E}(X) = \left(1 \times \frac{1}{3}\right) + \left(2 \times \frac{1}{2}\right) + \left(3 \times \frac{1}{6}\right) = \frac{1}{3} + 1 + \frac{1}{2} = \mathbf{1\frac{5}{6}}$

$\mathrm{E}(X^2) = \frac{1}{3} + 2 + \frac{9}{6} = 3\frac{5}{6}$

$\mathrm{Var}\ X = 3\frac{5}{6} - \left(1\frac{5}{6}\right)^2 = \frac{\mathbf{138}}{\mathbf{36}} - \frac{\mathbf{121}}{\mathbf{36}} = \frac{\mathbf{17}}{\mathbf{36}}$

b Mean $= \mathrm{E}(X) = \left(-1 \times \frac{1}{4}\right) + \left(0 \times \frac{1}{2}\right) + \left(1 \times \frac{1}{4}\right) = 0$ (or by symmetry = 0)

$\mathrm{E}(X^2) = \left(1 \times \frac{1}{4}\right) + \left(0 \times \frac{1}{2}\right) + \left(1 \times \frac{1}{4}\right) = \frac{1}{2}$

$\mathrm{Var}\ X = \frac{1}{2} - 0 = \frac{\mathbf{1}}{\mathbf{2}}$

c Mean $= \mathrm{E}(X) = \left(-2 \times \frac{1}{3}\right) + \left(-1 \times \frac{1}{3}\right) + \left(1 \times \frac{1}{6}\right) + \left(2 \times \frac{1}{6}\right) = -\frac{\mathbf{1}}{\mathbf{2}}$

$\mathrm{E}(X^2) = \left(4 \times \frac{1}{3}\right) + \left(1 \times \frac{1}{3}\right) + \left(1 \times \frac{1}{6}\right) + \left(4 \times \frac{1}{6}\right) = \mathbf{2\frac{1}{2}}$

$\mathrm{Var}\ X = 2\frac{1}{2} - \left(-\frac{1}{2}\right)^2 = \mathbf{2\frac{1}{4}}$

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Discrete random variables
Exercise D, Question 3

Question:

Given that Y is the score when a single unbiased die is rolled, find E(Y) and Var(Y).

Solution:

y	1	2	3	4	5	6
$P(Y = y)$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$

$$E(Y) = \frac{1}{6} + \frac{2}{6} + \frac{3}{6} + \frac{4}{6} + \frac{5}{6} + \frac{6}{6} = 3\frac{1}{2}$$

$$E(X^2) = \frac{1}{6} + \frac{4}{6} + \frac{9}{6} + \frac{16}{6} + \frac{25}{6} + \frac{36}{6} = 15\frac{1}{6}$$

$$\text{Var } X = 15\frac{1}{6} - \left(3\frac{1}{2}\right)^2 = \mathbf{2\frac{11}{12}}$$

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Discrete random variables
Exercise D, Question 4

Question:

Two fair cubical dice are rolled and S is the sum of their scores.

a Find the distribution of S.

b Find E(S).

c Find Var(S).

Solution:

a

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

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

b $E(S) = \frac{2+6+12+20+30+42+40+36+30+22+12}{36} = 7$ (or by symmetry = 7)

c $E(S^2) = \frac{4+18+48+100+180+294+320+324+300+242+144}{36}$

$= 54\frac{5}{6}$

$\text{Var}\,S = 54\frac{5}{6} - 7^2 = \mathbf{5\frac{5}{6}}$

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Discrete random variables
Exercise D, Question 5

Question:

Two fair tetrahedral (four-sided) dice are rolled and D is the difference between their scores.

a Find the distribution of D and show that $P(D = 3) = \frac{1}{8}$.

b Find E(D).

c Find Var(D).

Solution:

a

Diff.	1	2	3	4
1	0	1	2	3
2	1	0	1	2
3	2	1	0	1
4	3	2	1	0

d	0	1	2	3
$P(D = d)$	$\frac{1}{4}$	$\frac{3}{8}$	$\frac{1}{4}$	$\frac{1}{8}$

From Distribution Table it can be seen that $P(D = 3) = \mathbf{\frac{1}{8}}$

b $E(D) = 0 + \frac{3}{8} + \frac{1}{2} + \frac{3}{8} = \mathbf{1\frac{1}{4}}$

c $E(D^2) = 0 + \frac{3}{8} + 1 + 1\frac{1}{8} = \mathbf{2\frac{1}{2}}$

$\text{Var}\, D = 2\frac{1}{2} - \left(1\frac{1}{4}\right)^2 = \mathbf{\frac{15}{16}}$

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Discrete random variables
Exercise D, Question 6

Question:

A fair coin is tossed repeatedly until a head appears or three tosses have been made. The random variable T represents the number of tosses of the coin.

a Show that the distribution of T is

t	1	2	3
$P(T = t)$	$\frac{1}{2}$	$\frac{1}{4}$	$\frac{1}{4}$

b Find the expectation and variance of T.

Solution:

a $P(H) = \frac{1}{2}$

$P(TH) = \frac{1}{2} \times \frac{1}{2} = \frac{1}{4}$

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

$P(TTT) = \frac{1}{8}$

$P(T = 1) = \frac{1}{2}$

$P(T = 2) = \frac{1}{4}$

$P(T = 3) = \frac{1}{4}$

b $E(T) = 1 \times \frac{1}{2} + 2 \times \frac{1}{4} + 3 \times \frac{1}{4} = \mathbf{1\frac{3}{4}}$

$\text{Var}(T) = 1 \times \frac{1}{2} + 4 \times \frac{1}{4} + 9 \times \frac{1}{4} - \left(1\frac{3}{4}\right)^2 = \mathbf{\frac{11}{16}}$

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Discrete random variables
Exercise D, Question 7

Question:

The random variable X has the following distribution:

x	1	2	3
$\mathbf{P}(X = x)$	a	b	a

where a and b are constants.

a Write down E(X).

b Given that Var(X) = 0.75, find the values of a and b.

Solution:

a E(X) = 2 by symmetry

b $\sum \mathrm{p}(x) = 2a + b = 1 \qquad (1)$

$$\begin{aligned}\mathrm{Var}(X) &= \mathrm{E}(X^2) - (E(X))^2 \\ &= 10a + 4b - 2^2 \\ &= 10a + 4b - 4 = \frac{3}{4} \qquad (2)\end{aligned}$$

$$\begin{aligned}10a + 4b &= 4\frac{3}{4} \qquad \text{from (2)} \\ \underline{8a + 4b} &\underline{= 4} \qquad \text{from (1)} \times 4 \\ 2a &= \frac{3}{4}\end{aligned}$$

$$\begin{aligned}a &= \frac{3}{8} \\ b &= 1 - 2a \\ &= 1 - \frac{3}{4} \\ &= \mathbf{\frac{1}{4}}\end{aligned}$$

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Discrete random variables
Exercise E, Question 1

Question:

$E(X) = 4$, $Var(X) = 10$

Find

a $E(2X)$,

b Var $(2X)$.

Solution:

Remember mean is $E(X)$ and variance is Var X.

a $E(2X) = 2\ E(X) = 2 \times 4 =$ **8**

b Var $(2X) = 2^2$ Var $X = 4 \times 10 =$ **40**

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Discrete random variables
Exercise E, Question 2

Question:

$E(X) = 2$, $Var(X) = 6$

Find

a $E(3X)$,

b $E(3X + 1)$,

c $E(X - 1)$,

d $E(4 - 2X)$,

e $Var(3X)$,

f $Var(3X + 1)$,

g $Var(X - 1)$.

Solution:

a $E(3X) = 3\,E(X) = 3 \times 2 = \mathbf{6}$

b $E(3X + 1) = 3\,E(X) + 1 = (3 \times 2) + 1 = \mathbf{7}$

c $E(X - 1) = E(X) - 1 = 2 - 1 = \mathbf{1}$

d $E(4 - 2X) = 4 - 2E(X) = 4 - 2 \times 2 = \mathbf{0}$

e $Var(3X) = 3^2\,Var\,X = 9 \times 6 = \mathbf{54}$

f $Var(3X + 1) = 3^2\,Var\,X = \mathbf{54}$

g $Var(X - 1) = Var\,X = \mathbf{6}$

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Discrete random variables
Exercise E, Question 3

Question:

The random variable X has an expectation of 3 and a variance of 9.

Find

a $E(2X + 1)$,

b $E(2 + X)$,

c $\text{Var}(2X + 1)$,

d $\text{Var}(2 + X)$.

Solution:

a $E(2X + 1) = 2\,E(X) + 1 = (2 \times 3) + 1 = \mathbf{7}$

b $E(2 + X) = E(X + 2) = E(X) + 2 = 3 + 2 = \mathbf{5}$

c $\text{Var}(2X + 1) = 2^2\,\text{Var}(X) = 4 \times 9 = \mathbf{36}$

d $\text{Var}(2 + X) = \text{Var}(X + 2) = \text{Var}(X) = \mathbf{9}$

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Discrete random variables
Exercise E, Question 4

Question:

The random variable X has a mean μ and standard deviation σ.

Find, in terms of μ and σ

a $E(4X)$,

b $E(2X + 2)$,

c $E(2X - 2)$,

d $\text{Var}(2X + 2)$,

e $\text{Var}(2X - 2)$.

Solution:

a $E(4X) = 4\,E(X) = 4\mu$

b $E(2X + 2) = 2\,E(X) + 2 = \mathbf{2\,\mu + 2}$

c $E(2X - 2) = 2\,E(X) - 2 = \mathbf{2\,\mu - 2}$

d $\text{Var}(2X + 2) = 2^2\,\text{Var}(X) = \mathbf{4\sigma^2}$ (Remember Standard deviation is σ so variance is σ^2)

e $\text{Var}(2X - 2) = 2^2\,\text{Var}(X) = \mathbf{4\sigma^2}$

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Discrete random variables
Exercise E, Question 5

Question:

The random variable Y has mean 2 and variance 9.

Find:

a $E(3Y+1)$,

b $E(2-3Y)$,

c $\text{Var}(3Y+1)$,

d $\text{Var}(2-3Y)$,

e $E(Y^2)$,

f $E[(Y-1)(Y+1)]$.

Solution:

a $E(3Y+1) = 3E(Y) + 1 = 3 \times 2 + 1 = \mathbf{7}$

b $E(2-3Y) = 2 - 3E(Y) = 2 - 3 \times 2 = -\mathbf{4}$

c $\text{Var}(3Y+1) = 3^2\text{Var}(Y) = 9 \times 9 = \mathbf{81}$

d $\text{Var}(2-3Y) = (-3)^2\ \text{Var}(Y) = 9 \times 9 = \mathbf{81}$

e $E(Y^2) = \text{Var}(Y) + [E(Y)]^2 = 9 + 2^2 = \mathbf{13}$

f $E[(Y-1)(Y+1)] = E(Y^2-1) = E(Y^2) - 1 = 13 - 1 = \mathbf{12}$

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Discrete random variables
Exercise E, Question 6

Question:

The random variable T has a mean of 20 and a standard deviation of 5.

It is required to scale T by using the transformation $S = 3T + 4$.

Find E(S) and Var(S).

Solution:

$\mathrm{E}(S) = \mathrm{E}(3T + 4) = 3\ \mathrm{E}(S) + 4 = (3 \times 20) + 4 = \mathbf{64}$

$\mathrm{Var}(T) = 5 \times 5 = 25$

$\mathrm{Var}\ (S) = \mathrm{Var}(3T + 4) = 3^2 \mathrm{Var\ T} = 9 \times 25 = \mathbf{225}$

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Discrete random variables
Exercise E, Question 7

Question:

A fair spinner is made from the disc in the diagram and the random variable X represents the number it lands on after being spun.

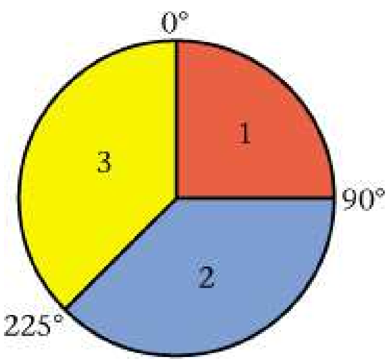
a Write down the distribution of X.

b Work out E(X).

c Find Var(X).

d Find E($2X + 1$).

e Find Var($3X - 1$).



Solution:

a

x	1	2	3
P($X = x$)	$\frac{1}{4}$	$\frac{3}{8}$	$\frac{3}{8}$

b $\mathrm{E}(X) = \left(1 \times \frac{1}{4}\right) + \left(2 \times \frac{3}{8}\right) + \left(3 \times \frac{3}{8}\right) = \mathbf{2\frac{1}{8}}$

c $\mathrm{E}(X^2) = \left(1 \times \frac{1}{4}\right) + \left(4 \times \frac{3}{8}\right) + \left(9 \times \frac{3}{8}\right) = \mathbf{5\frac{1}{8}}$

$\mathrm{Var}(X) = 5\frac{1}{8} - \left(2\frac{1}{8}\right)^2 = \mathbf{\frac{39}{64}}$

d $\mathrm{E}(2X + 1) = 2\mathrm{E}(X) + 1 = \left(2 \times 2\frac{1}{8}\right) + 1 = \mathbf{5\frac{1}{4}}$

e $\mathrm{Var}(3X - 1) = 3^2\ \mathrm{Var}\ X = 9 \times \mathbf{\frac{39}{64}} = \mathbf{5\frac{31}{64}}$

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Discrete random variables
Exercise E, Question 8

Question:

The discrete variable X has the probability distribution

x	−1	0	1	2
P(X = x)	0.2	0.5	0.2	0.1

a Find E(X),

b Find Var (X),

c Find $\mathrm{E}(\frac{1}{3}X+1)$,

d Find $\mathrm{Var}(\frac{1}{3}X+1)$.

Solution:

a $\mathrm{E}(X) = -0.2 + 0 + 0.2 + 0.2 = \mathbf{0.2}$

b $\mathrm{E}(X^2) = 0.2 + 0 + 0.2 + 0.4 = \mathbf{0.8}$

$\mathrm{Var}(X) = 0.8 - 0.2^2 = 0.8 - 0.04 = \mathbf{0.76}$

c $\mathrm{E}\left(\frac{1}{3}X+1\right) = \frac{1}{3}\mathrm{E}(X)+1 = \left(\frac{1}{3}\times 0.2\right)+1 = \left(\frac{1}{3}\times\frac{1}{5}\right)+1 = \mathbf{1\frac{1}{15}}(1.0\dot{6})$

d $\mathrm{Var}\left(\frac{1}{3}X+1\right) = \left(\frac{1}{3}\right)^2\mathrm{Var}X = \frac{1}{9}\times 0.76 = \frac{1}{9}\times\frac{19}{25} = \mathbf{\frac{19}{225}}(0.08\dot{4})$

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Discrete random variables
Exercise F, Question 1

Question:

X is a discrete uniform distribution over the numbers 1, 2, 3, 4 and 5. Work out the expectation and variance of X.

Solution:

Expectation $= \frac{n+1}{2} = \frac{5+1}{2} = \mathbf{3}$

Variance $= \frac{(n+1)(n-1)}{12} = \frac{(5+1)(5-1)}{12} = \mathbf{2}$

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Discrete random variables
Exercise F, Question 2

Question:

Seven similar balls are placed in a bag. The balls have the numbers 1 to 7 on them. A ball is drawn out of the bag. The variable X represents the number on the ball.

a Find E(X).

b Work out Var(X).

Solution:

a $n = 7$

$\mathrm{E}(X) = \frac{n+1}{2} = 4$

b $\mathrm{Var}(X) = \frac{(n+1)(n-1)}{12} = \mathbf{4}$

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Discrete random variables
Exercise F, Question 3

Question:

A fair die is thrown once and the random variable X represents the value on the upper face.

a Find the expectation and variance of X.

b Calculate the probability that X is within one standard deviation of the expectation.

Solution:

a Expectation $= \frac{n+1}{2} = \frac{6+1}{2} = \mathbf{3\frac{1}{2}}$

$\text{Var}(X) = \frac{(n+1)(n-1)}{12} = \frac{7 \times 5}{12} = \mathbf{2\frac{11}{12}} = \mathbf{2.91\dot{6}}$

b

x	1	2	3	4	5	6
$\text{P}(X = x)$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$

$\sigma = \sqrt{2.91\dot{6}} = 1.7078$

Therefore we want between 3.5 − 1.7078 = 1.7922 and 3.5 + 1.7078 = 5.2078

$\text{P}(1.7922 \le X \le 5.2078) = \text{p}(2) + \text{p}(3) + \text{p}(4) + \text{p}(5) = \mathbf{\frac{2}{3}}$

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Discrete random variables
Exercise F, Question 4

Question:

A card is selected at random from a pack of cards containing the even numbers 2, 4, 6, …, 20. The variable X represents the number on the card.

a Find $P(X > 15)$.

b Find the expectation and variance of X.

Solution:

a This is a uniform distribution.

x	2	4	6	8	10	12	14	16	18	20
$P(X = x)$	$\frac{1}{10}$	$\frac{1}{10}$	$\frac{1}{10}$	$\frac{1}{10}$	$\frac{1}{10}$	$\frac{1}{10}$	$\frac{1}{10}$	$\frac{1}{10}$	$\frac{1}{10}$	$\frac{1}{10}$

$P(X > 15) = P(16) + P(18) + P(20) = \mathbf{\frac{3}{10}}$

b Let R be a uniform distribution over the numbers 1,2, … 10.

Then $X = 2R$

$E(R) = \frac{n+1}{2} = 5.5$

$\text{Var}(R) = \frac{(n+1)(n-1)}{12} = \frac{99}{12} = 8\frac{1}{4} = 8.25$

$E(X) = 2E(R) = 2 \times 5.5 = \mathbf{11}$

$\text{Var}(X) = \text{Var}(2R) = 2^2\text{Var}(R) = 4 \times 8.25 = \mathbf{33}$

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Discrete random variables
Exercise F, Question 5

Question:

A card is selected at random from a pack of cards containing the odd numbers 1, 3, 5, …, 19. The variable X represents the number on the card.

a Find $P(X > 15)$.

b Find the expectation and variance of X.

Solution:

y	1	3	5	7	9	11	13	15	17	19
$P(Y = y)$	$\frac{1}{10}$	$\frac{1}{10}$	$\frac{1}{10}$	$\frac{1}{10}$	$\frac{1}{10}$	$\frac{1}{10}$	$\frac{1}{10}$	$\frac{1}{10}$	$\frac{1}{10}$	$\frac{1}{10}$

a $P(X > 15) = P(17) + P(19) = \mathbf{\frac{1}{5}}$

b $Y = X - 1$ (X as in previous question)

$E(Y) = E(X - 1) = E(X) - 1 = 11 - 1 = \mathbf{10}$

$\text{Var}(Y) = \text{Var}(X - 1) = \text{Var}(X) = \mathbf{33}$

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Discrete random variables
Exercise F, Question 6

Question:

A straight line is drawn on a piece of paper. The line is divided into four equal lengths and the segments are marked 1, 2, 3 and 4. In a party game a person is blindfolded and asked to mark a point on the line and the number of the segment is recorded. A discrete uniform distribution over the set (1, 2, 3, 4) is suggested as model for this distribution. Comment on this suggestion.

Solution:

A discrete uniform distribution is not likely to be a good model for this distribution. The game depends on the skill of the player. The points are quite likely to cluster around the middle.

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

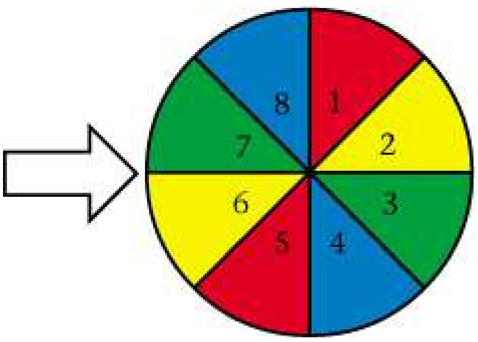
Discrete random variables
Exercise F, Question 7

Question:

In a fairground game the spinner shown is used.

It cost 5p to have a go on the spinner.

The spinner is spun and the number of pence shown is returned to the contestant.



If X is the number which comes up on the next spin,

a name a suitable model for X,

b find E(X),

c find Var(X),

d explain why you should not expect to make money at this game if you have a large number of goes.

Solution:

a A discrete uniform distribution

b $n = 8$

$\mathrm{E}(X) = \frac{n+1}{2} = \mathbf{4.5}$

c $\mathrm{Var}(X) = \frac{(n+1)(n-1)}{12} = \frac{63}{12} = \mathbf{5\frac{1}{4}}$ or $\mathbf{5.25}$

d The expected winnings are less than the 5p stake.

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Discrete random variables
Exercise G, Question 1

Question:

The random variable X has probability function

$P(X = x) = \frac{x}{21} x = 1, 2, 3, 4, 5, 6.$

a Construct a table giving the probability distribution of X.

Find

b $P(2 < X \le 5)$,

c $E(X)$,

d $Var(X)$,

e $Var(3 - 2X)$.

Solution:

a

x	1	2	3	4	5	6
$P(X = x)$	$\frac{1}{21}$	$\frac{2}{21}$	$\frac{3}{21}$	$\frac{4}{21}$	$\frac{5}{21}$	$\frac{6}{21}$

b $P(3) + P(4) + P(5) = \frac{3}{21} + \frac{4}{21} + \frac{5}{21} = \mathbf{\frac{12}{21}}$

c $E(X) = \left(1 \times \frac{1}{21}\right) + \left(2 \times \frac{2}{21}\right) + \left(3 \times \frac{3}{21}\right) + \left(4 \times \frac{4}{21}\right) + \left(5 \times \frac{5}{21}\right) + \left(6 \times \frac{6}{21}\right)$

$= \frac{91}{21} = \mathbf{4\frac{1}{3}} (4.\dot{3})$

d $E(X^2) = \left(1 \times \frac{1}{21}\right) + \left(4 \times \frac{2}{21}\right) + \left(9 \times \frac{3}{21}\right) + \left(16 \times \frac{4}{21}\right) + \left(25 \times \frac{5}{21}\right) + \left(36 \times \frac{6}{21}\right)$

$= \frac{441}{21} = \mathbf{21}$

$Var(X) = 21 - \left(4\frac{1}{3}\right)^2 = 21 - 18\frac{7}{9} = \mathbf{2\frac{2}{9}}$

e $Var(3 - 2X) = Var(-2X + 3) = (-2)^2 Var(X) = 4 \times 2\frac{2}{9} = \mathbf{8\frac{8}{9}}$

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Discrete random variables
Exercise G, Question 2

Question:

The discrete random variable X has the probability distribution shown.

x	-2	-1	0	1	2	3
$P(X = x)$	0.1	0.2	0.3	r	0.1	0.1

Find

a r,

b $P(-1 \leq X < 2)$,

c F(0.6),

d $E(2X + 3)$,

e $Var(2X + 3)$.

Solution:

a $0.1 + 0.2 + 0.3 + r + 0.1 + 0.1 = 1$

$r = 1 - 0.8 =$ **0.2**

b $P(-1) + P(0) + P(1) = 0.2 + 0.3 + 0.2 =$ **0.7**

c

x	-2	-1	0	1	2	3
$F(X)$	0.1	0.3	0.6	0.8	0.9	1

$F(0.6) = F(0) =$ **0.6**

d $E(X) = (-0.2) + (-0.2) + 0 + 0.2 + 0.2 + 0.3 = 0.3$

$E(2X + 3) = 2E(X) + 3 = (2 \times 0.3) + 3 =$ **3.6**

e $E(X^2) = 0.4 + 0.2 + 0 + 0.2 + 0.4 + 0.9 = 2.1$

$Var(X) = 2.1 - 0.3^2 = 2.01$

$Var(2X + 3) = 2^2 \, Var X = 4 \times 2.01 =$ **8.04**

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Discrete random variables
Exercise G, Question 3

Question:

A discrete random variable X has the probability distribution shown in the table below.

x	0	1	2
$\mathrm{P}(X = x)$	$\frac{1}{5}$	b	$\frac{1}{5} + b$

a Find the value of b.

b Show that $\mathrm{E}(X) = 1.3$.

c Find the exact value of $\mathrm{Var}(X)$.

d Find the exact value of $\mathrm{P}(X \le 1.5)$.

Solution:

a

x	0	1	2
$\mathrm{P}(X = x)$	$\frac{1}{5}$	b	$b + \frac{1}{5}$

$\frac{1}{5} + b + b + \frac{1}{5} = 1$

$2b = 1 - \frac{2}{5} = \frac{3}{5}$

$b = \frac{3}{10}$

b $\mathrm{E}(X) = \left(0 \times \frac{1}{5}\right) + \left(1 \times \frac{3}{10}\right) + \left(2 \times \frac{5}{10}\right) = 0 + \frac{3}{10} + 1 = 1.3$

c $\mathrm{E}(X^2) = \left(0 \times \frac{1}{5}\right) + \left(1 \times \frac{3}{10}\right) + \left(4 \times \frac{5}{10}\right) = 0 + \frac{3}{10} + 2 = 2.3$

$\mathrm{Var}(X) = 2.3 - 1.3^2 = 0.61$

d $\mathrm{P}(0) + \mathrm{P}(1) = \frac{1}{5} + \frac{3}{10} = 0.5$

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Discrete random variables
Exercise G, Question 4

Question:

The discrete random variable X has a probability function

$$P(X = x) = \begin{cases} k(1-x) & x = 0,1 \\ k(x-1) & x = 2,3 \\ 0 & \text{otherwise} \end{cases}$$

where k is a constant.

a Show that $k = \frac{1}{4}$.

b Find $E(X)$ and show that $E(X^2) = 5.5$.

c Find $\text{Var}(2X - 2)$.

Solution:

a $k(1 - 0) + k(1 - 1) + k(2 - 1) + k(3 - 1) = 1$

$k + k + 2k = 1$

$4k = 1$

$k = \mathbf{\frac{1}{4}}$

b

x	0	1	2	3
$P(X = x)$	$\frac{1}{4}$	0	$\frac{1}{4}$	$\frac{1}{2}$

$E(X) = 0 + 0 + \frac{1}{2} + \frac{3}{2} = 2$

$E(X^2) = \left(0 \times \frac{1}{4}\right) + (1 \times 0) + \left(4 \times \frac{1}{4}\right) + \left(9 \times \frac{1}{2}\right) = 1 + 4.5 = \mathbf{5.5}$

c $\text{Var}(X) = 5.5 - 4 = 1.5$

$\text{Var}(2X - 2) = 4 \times 1.5 = \mathbf{6}$

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Discrete random variables
Exercise G, Question 5

Question:

A discrete random variable X has the probability distribution,

x	0	1	2	3
P(*X* = *x*)	$\frac{1}{4}$	$\frac{1}{2}$	$\frac{1}{8}$	$\frac{1}{8}$

Find

a $P(1 < X \le 2)$,

b F(1.5),

c E(X),

d E($3X - 1$),

e Var(X).

Solution:

a $P(1 < X \le 2) = P(2) = \mathbf{\frac{1}{8}}$

b $F(1.5) = F(1) = \frac{1}{4} + \frac{1}{2} = \mathbf{\frac{3}{4}}$

c $E(X) = 0 + \frac{1}{2} + \frac{1}{4} + \frac{3}{8} = \frac{9}{8} = \mathbf{1\frac{1}{8}}$

d $E(3X - 1) = 3\,E(X) - 1 = 3\frac{3}{8} - 1 = \mathbf{2\frac{3}{8}}$

e $E(X^2) = 0 + \frac{1}{2} + \frac{1}{2} + \frac{9}{8} = \mathbf{2\frac{1}{8}}$

$\text{Var}(X) = 2\frac{1}{8} - \left(1\frac{1}{8}\right)^2 = \mathbf{\frac{55}{64}}$

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Discrete random variables
Exercise G, Question 6

Question:

A discrete random variable is such that each of its values is assumed to be equally likely.

a Write the name of the distribution.

b Give an example of such a distribution.

A discrete random variable *X* as defined above can take values 0, 1, 2, 3, and 4.

Find

c E(*X*),

d Var (*X*).

Solution:

a A discrete uniform distribution

b Any distribution where all the probabilities are the same. An example is throwing a fair die.

c

x	0	1	2	3	4
$P(X = x)$	$\frac{1}{5}$	$\frac{1}{5}$	$\frac{1}{5}$	$\frac{1}{5}$	$\frac{1}{5}$

$E(X) = 0 + \frac{1}{5} + \frac{2}{5} + \frac{3}{5} + \frac{4}{5} = \mathbf{2}$ (or $\frac{5+1}{2} - 1 = 2$) (OR use symmetry)

d $E(X^2) = 0 + \frac{1}{5} + \frac{4}{5} + \frac{9}{5} + \frac{16}{5} = \mathbf{6}$ $\left(\text{or } \mathrm{Var}(X) = \frac{(5+1)(5-1)}{12} = 2\right)$

Var $(X) = 6 - 4 = \mathbf{2}$

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Discrete random variables
Exercise G, Question 7

Question:

The random variable X has a probability distribution

x	1	2	3	4	5
P(X = x)	0.1	p	q	0.3	0.1

a Given that E(X) = 3.1, write down two equations involving p and q.

Find

b the value of p and the value of q,

c Var(X),

d Var($2X - 3$).

Solution:

a $0.1 + p + q + 0.3 + 0.1 = 1$

$p + q = 0.5$ **1**

$0.1 + 2p + 3q + 1.2 + 0.5 = 3.1$

$\mathbf{2p + 3q = 1.3}$ **2**

b $2p + 3q = 1.3$ (equation 2)

$2p + 2q = 1$ (equation 1 times 2) **3**

$\mathbf{q = 0.3}$ (equation 2–3)

$p + \mathbf{0.3} = 0.5$

$\mathbf{p = 0.2}$

c

x	1	2	3	4	5
P(X = x)	0.1	0.2	0.3	0.3	0.1

$E(X) = 0.1 + 0.4 + 0.9 + 1.2 + 0.5 = 3.1$

$E(X^2) = 0.1 + 0.8 + 2.7 + 4.8 + 2.5 = 10.9$

$\text{Var}(X) = 10.9 - 3.1^2 = \mathbf{1.29}$

d Var $(2X + 3) = 4$ Var $(X) = 4 \times 1.29 =$ **5.16**

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Discrete random variables
Exercise G, Question 8

Question:

The random variable X has probability function

$$P(X = x) = \begin{cases} kx & x = 1,2 \\ k(x-2) & x = 3,4,5 \end{cases}$$

where k is a constant.

a Find the value of k.

b Find the exact value of E(X).

c Show that, to three significant figures, Var(X) = 2.02.

d Find, to one decimal place, Var(3 − 2X).

Solution:

a

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

$k + 2k + k + 2k + 3k = 1$

$9k = 1$

k$= \frac{1}{9}$

b $E(X) = \frac{1}{9} + \frac{4}{9} + \frac{3}{9} + \frac{8}{9} + \frac{15}{9} = \frac{31}{9} = \mathbf{3}\frac{4}{9}$

c $E\left(X^2\right) = \frac{1}{9} + \frac{8}{9} + 1 + \frac{32}{9} + \frac{75}{9} = \frac{125}{9} = 13\frac{8}{9}$

$\text{Var}(X) = 13\frac{8}{9} - \left(3\frac{4}{9}\right)^2 = 13.888 - 11.864 = \mathbf{2.02}$ to 3 sig figs

d Var (3 − 2X) = $(-2)^2$ Var (X) = 4 × 2.02 = **8.1** to 1dp

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Discrete random variables
Exercise G, Question 9

Question:

The random variable X has the discrete uniform distribution

$\mathrm{P}(X = x) = \frac{1}{6}\ \ x = 1,2,3,4,5,6.$

a Write down E(X) and show that $\mathrm{Var}(X) = \frac{35}{12}$.

b Find E(2X − 1).

c Find Var(3 − 2X).

Solution:

a $\mathrm{E}(X) = 3.5 = 3\frac{1}{2}$

$\mathrm{E}\left(X^2\right) = \frac{1}{6} + \frac{4}{6} + \frac{9}{6} + \frac{16}{6} + \frac{25}{6} + \frac{36}{6} = \frac{91}{6} = 15\frac{1}{6}$

$\mathrm{Var}\ (X) = 15\frac{1}{6} - \left(3\frac{1}{2}\right)^2 = \frac{91}{6} - \frac{49}{4} = \mathbf{\frac{35}{12}}$

b E(2X − 1) = 2 E(X) − 1 = 7 − 1 = **6**

c $\mathrm{Var}\ (3 - 2X) = 4\ \mathrm{Var}\ (X) = 4 \times \frac{35}{12} = \frac{35}{3} = \mathbf{11\frac{2}{3}}$ or **11.67** to 2 dp

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Discrete random variables
Exercise G, Question 10

Question:

The random variable X has probability function

$p(x) = \frac{(3x-1)}{26}$ $x = 1,2,3,4.$

a Construct a table giving the probability distribution of X.

Find

b $P(2 < X \leq 4)$,

c the exact value of $E(X)$.

d Show that $Var(X) = 0.92$ to two significant figures.

e Find $Var(1 - 3X)$.

Solution:

a

x	1	2	3	4
$P(X = x)$	$\frac{2}{26}$	$\frac{5}{26}$	$\frac{8}{26}$	$\frac{11}{26}$

b $P(2 < X \leq 4) = P(3) + P(4) = \mathbf{\frac{19}{26}}$

c $E(X) = \frac{2}{26} + \frac{10}{26} + \frac{24}{26} + \frac{44}{26} = \frac{80}{26} = \mathbf{3\frac{1}{13}}$

d $E\left(X^2\right) = \frac{2}{26} + \frac{20}{26} + \frac{72}{26} + \frac{176}{26} = \frac{270}{26} = 10\frac{10}{26} = \mathbf{10\frac{5}{13}}$

$Var\ (X) = 10\frac{5}{13} - \left(3\frac{1}{13}\right)^2 = 10.385\ldots - 9.467\ldots = \mathbf{0.92}$

e $Var(1 - 3X) = (-3)^2\ Var\ (X) = 9 \times 0.92 = \mathbf{8.28}$

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Normal distribution
Exercise A, Question 1

Question:

Use tables of the normal distribution to find the following.

a P(Z < 2.12)

b P(Z < 1.36)

c P(Z > 0.84)

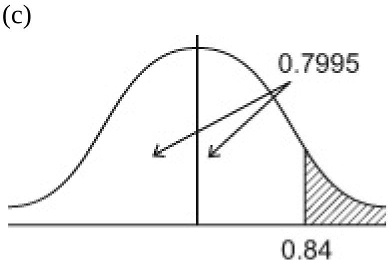
d P(Z < −0.38)

Solution:

(a) P(z < 2.12) = <u>0.9830</u>

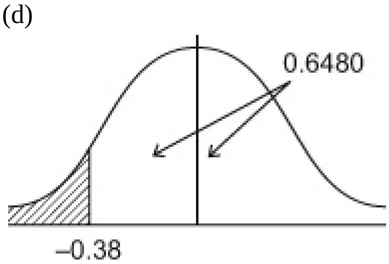
(b) P(z < 1.36) = <u>0.9131</u>

(c)



$$P(Z > 0.84) = 1 - 0.7995 = \underline{0.2005}$$

(d)



$$P(Z < -0.38) = 1 - 0.6480 = \underline{0.352}$$

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Normal distribution
Exercise A, Question 2

Question:

Use tables of the normal distribution to find the following.

a $P(Z > 1.25)$

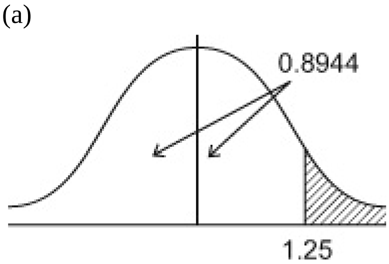
b $P(Z > -1.68)$

c $P(Z < -1.52)$

d $P(Z < 3.15)$

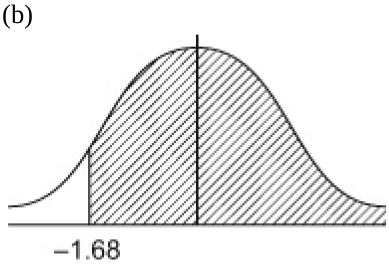
Solution:

(a)



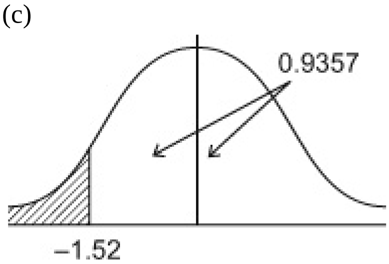
$$\begin{aligned} P(Z > 1.25) &= 1 - 0.8944 \\ &= \underline{\underline{0.1056}} \end{aligned}$$

(b)



$$P(Z > -1.68) = \underline{\underline{0.9535}}$$

(c)



$$\begin{aligned} P(Z < -1.52) &= 1 - 0.9357 \\ &= \underline{\underline{0.0643}} \end{aligned}$$

(d) $P(Z < 3.15) = \underline{\underline{0.9992}}$

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Normal distribution
Exercise A, Question 3

Question:

Use tables of the normal distribution to find the following.

a $P(Z > -2.24)$

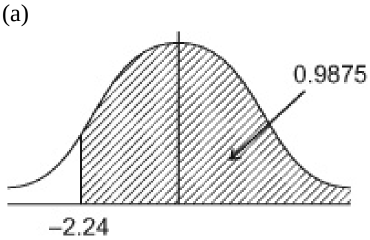
b $P(0 < Z < 1.42)$

c $P(-2.30 < Z < 0)$

d $P(Z < -1.63)$

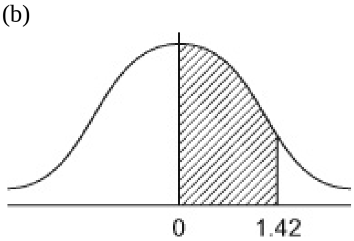
Solution:

(a)



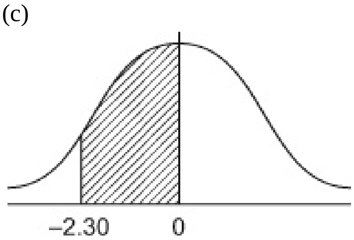
$P(Z > -2.24) = \underline{\underline{0.9875}}$

(b)



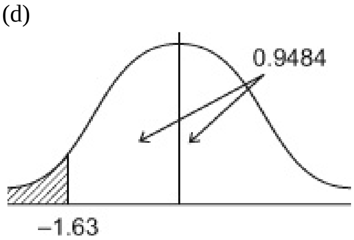
$$\begin{aligned} P(0 < Z < 1.42) &= 0.9222 - 0.5 \\ &= \underline{\underline{0.4222}} \end{aligned}$$

(c)



$$\begin{aligned} P(-2.30 < Z < 0) &= 0.9893 - 0.5 \\ &= \underline{\underline{0.4893}} \end{aligned}$$

(d)



$$\begin{aligned} P(Z < -1.63) &= 1 - 0.9484 \\ &= \underline{\underline{0.0516}} \end{aligned}$$

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Normal distribution
Exercise A, Question 4

Question:

Use tables of the normal distribution to find the following.

a $P(1.25 < Z < 2.16)$

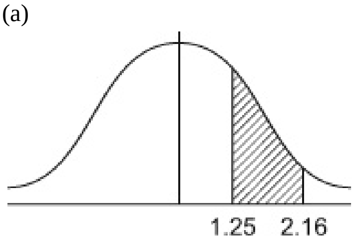
b $P(-1.67 < Z < 2.38)$

c $P(-2.16 < Z < -0.85)$

d $P(-1.57 < Z < 1.57)$

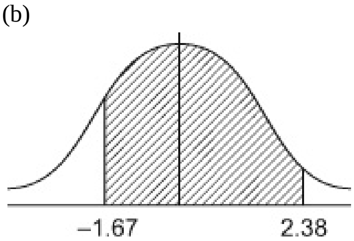
Solution:

(a)



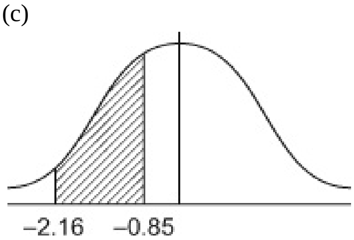
$P(1.25 < Z < 2.16)$
$= P(Z < 2.16) - P(Z < 1.25)$
$= 0.9846 - 0.8944$
$= \underline{\underline{0.0902}}$

(b)



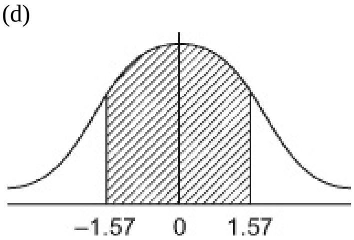
$P(-1.67 < Z < 2.38)$
$= P(Z < 2.38) - [1 - P(Z < 1.67)]$
$= 0.9913 - [1 - 0.9525]$
$= 0.9913 - 0.0475$
$= \underline{\underline{0.9438}}$

(c)



$P(-2.16 < Z < -0.85)$
$= P(Z < 2.16) - P(Z < 0.85)$
$= 0.9846 - 0.8023$
$= \underline{\underline{0.1823}}$

(d)



$P(-1.57 < Z < 1.57)$
$= 2 \times P(0 < Z < 1.57)$
$= 2 \times [0.9418 - 0.5]$
$= 2 \times 0.4418$
$= \underline{\underline{0.8836}}$

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Normal distribution
Exercise B, Question 1

Question:

Find the value of *a* in the following.

a $P(Z < a) = 0.9082$

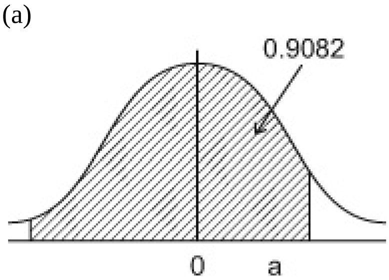
b $P(Z > a) = 0.0314$

c $P(Z < a) = 0.3372$

d $P(Z > a) = 0.6879$

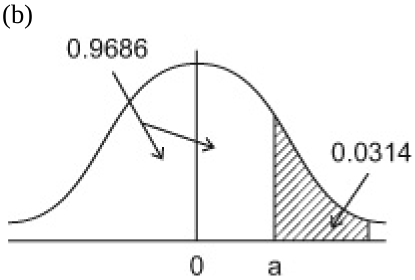
Solution:

(a)



$\underline{a = 1.33}$

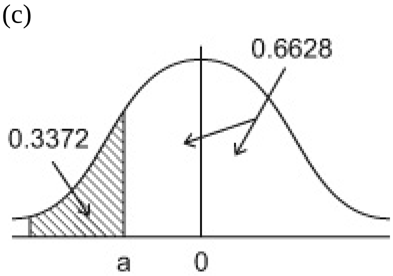
(b)



$1 - 0.0314 = 0.9686$

$\therefore \underline{\underline{a = 1.86}}$

(c)



$1 - 0.3372 = 0.6628$

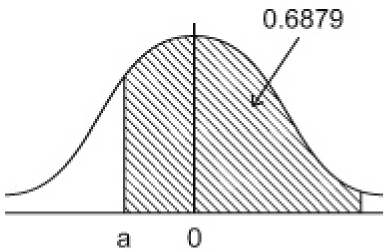
N.B.$0.3372 < 0.5 \therefore a < 0$

$\therefore \underline{\underline{a = -0.42}}$

(d)

[N.B.$a < 0$again]

$\underline{\underline{a = -0.49}}$


0.6879
a
0

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Normal distribution
Exercise B, Question 2

Question:

Find the value of a in the following.

a $P(Z < a) = 0.9938$

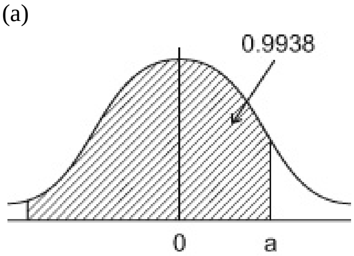
b $P(Z > a) = 0.4129$

c $P(Z > a) = 0.7611$

d $P(Z > a) = 0.2000$

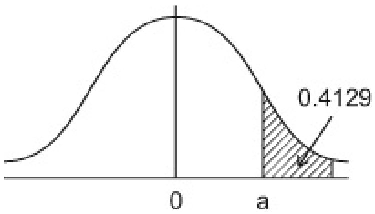
Solution:

(a)



$\underline{\underline{a = 2.50}}$

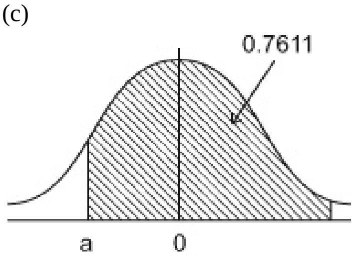
(b)



$1 - 0.4129 = 0.5871$

$\therefore\ \underline{\underline{a = 0.22}}$

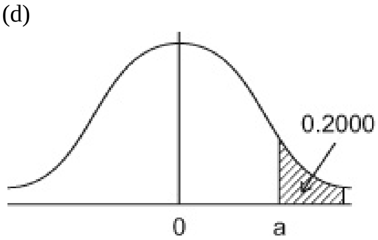
(c)



[N.B. $a < 0 \because P(Z > a) > 0.5$]

$\underline{\underline{a = -0.71}}$

(d)



Using table of percentage points

$\underline{\underline{a = 0.8416}}$

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Normal distribution
Exercise B, Question 3

Question:

Find the value of a in the following.

a $P(Z > a) = 0.1500$

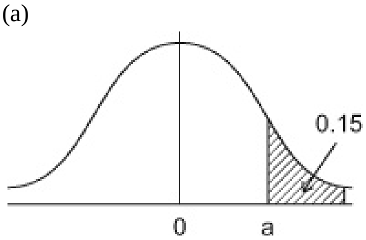
b $P(Z > a) = 0.9500$

c $P(Z > a) = 0.1112$

d $P(Z < a) = 0.9990$

Solution:

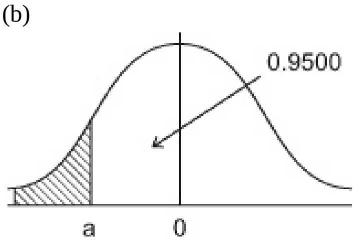
(a)



Use table of percentage points with $p = 0.15$

$\underline{\underline{a = 1.0364}}$

(b)

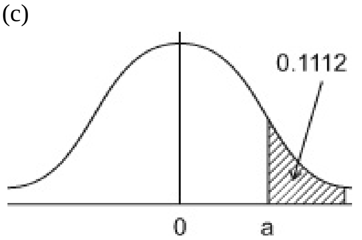


$1 - 0.9500 = 0.05 = p$

Using table of percentage points

$\underline{\underline{a = -1.6449}}$

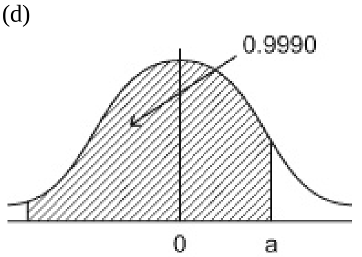
(c)



$1 - 0.1112 = 0.8888$

$\underline{\underline{a = 1.22}}$

(d)



$1 - 0.9990 = 0.0010$

Use table of percentage points with

$p = 0.0010$

$\underline{\underline{a = 3.0902}}$

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Normal distribution
Exercise B, Question 4

Question:

Find the value of a in the following.

a $P(0 < Z < a) = 0.3554$

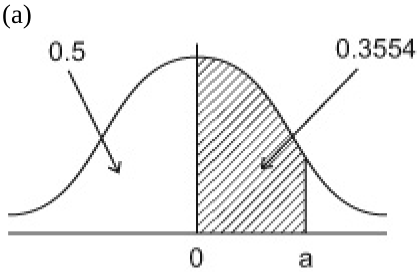
b $P(0 < Z < a) = 0.4946$

c $P(-a < Z < a) = 0.5820$

d $P(-a < Z < a) = 0.8230$

Solution:

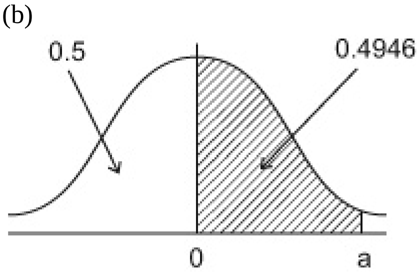
(a)



$P(Z < a) = 0.8554$

$\therefore \underline{\underline{a = 1.06}}$

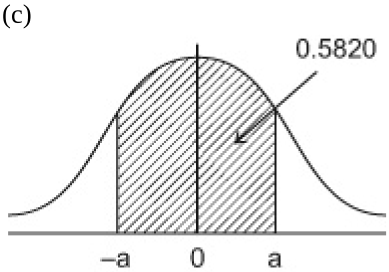
(b)



$P(Z < a) = 0.9946$

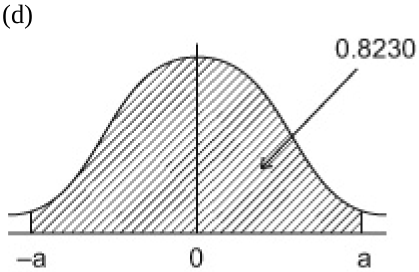
$\therefore \underline{\underline{a = 2.55}}$

(c)



$$\begin{aligned} P(0 < Z < a) &= \frac{1}{2} \times 0.5820 \\ &= 0.2910 \\ \therefore P(Z < a) &= 0.7910 \\ \therefore \underline{\underline{a = 0.81}} \end{aligned}$$

(d)



$$\begin{aligned} P(0 < Z < a) &= \frac{1}{2} \times 0.8230 \\ &= 0.4115 \end{aligned}$$

$\therefore$ P($Z < a$) = 0.9115

$\therefore$ $\underline{\underline{a = 1.35}}$

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Normal distribution
Exercise B, Question 5

Question:

Find the value of a in the following.

a $p(0 < Z < a) = 0.10$

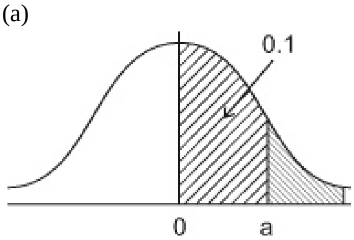
b $p(0 < Z < a) = 0.35$

c $p(-a < Z < a) = 0.80$

d $p(-a < Z < a) = 0.40$

Solution:

(a)

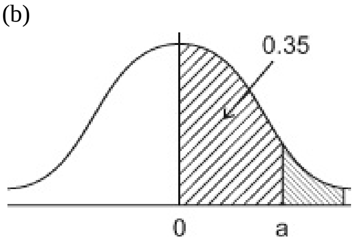


$p = 0.5 - 0.1 = 0.4000$

Use table of percentage points

$\underline{\underline{a = 0.2533}}$

(b)

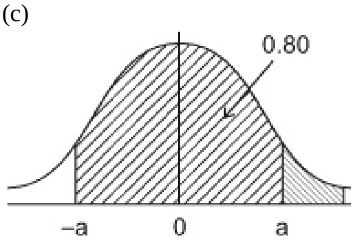


$p = 0.5 - 0.35 = 0.1500$

Use table of percentage points

$\underline{\underline{a = 1.0364}}$

(c)

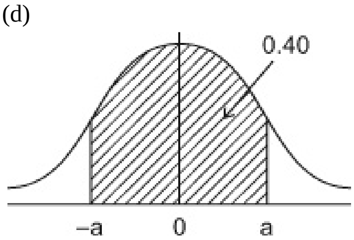


$p = 0.5 - \frac{1}{2} \times 0.8 = 0.1000$

Use table of percentage points

$\underline{\underline{a = 1.2816}}$

(d)



$P(0 < Z < a) = \frac{1}{2} \times 0.4 = 0.20$

$p = 0.5 - 0.2 = 0.3000$

Use table of percentage points

$\underline{\underline{a = 0.5244}}$

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Normal distribution
Exercise C, Question 1

Question:

The random variable $X \sim N(30, 2^2)$.

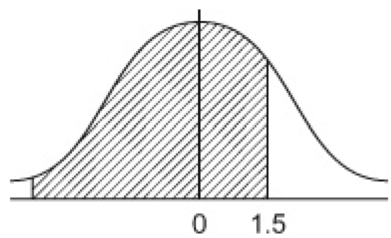
Find **a** $P(X < 33)$,

b $P(X > 26)$.

Solution:

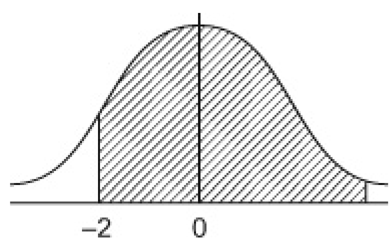
(a)

$$\begin{aligned} P(X < 33) &= P\left(Z < \frac{33-30}{2}\right) \\ &= P(Z < 1.5) \\ &= \underline{\underline{0.9332}} \end{aligned}$$



(b)

$$\begin{aligned} P(X > 26) &= P\left(Z > \frac{26-30}{2}\right) \\ &= P(Z > -2) \\ &= \underline{\underline{0.9772}} \end{aligned}$$



Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Normal distribution
Exercise C, Question 2

Question:

The random variable $X \sim N(40, 9)$.

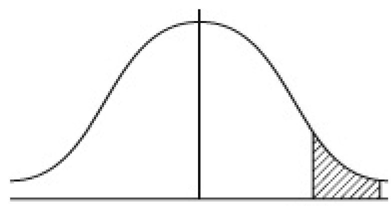
Find **a** $P(X > 45)$,

b $P(X < 38)$.

Solution:

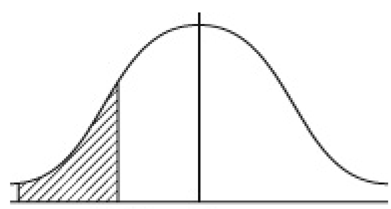
(a)

$$\begin{aligned} P(X > 45) &= P\left(Z > \frac{45-40}{\sqrt{9}}\right) \\ &= P(Z > 1.67) \\ &= 1 - 0.9525 \\ &= \underline{\underline{0.0475}} \text{ (allow AWRT } \underline{\underline{0.048}}) \end{aligned}$$



(b)

$$\begin{aligned} P(X < 38) &= P\left(Z < \frac{38-40}{3}\right) \\ &= P(Z < -0.67) \\ &= 1 - 0.7486 \\ &= \underline{\underline{0.2514}} \text{ (allow AWRT 0.251 or 0.252)} \end{aligned}$$



Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Normal distribution
Exercise C, Question 3

Question:

The random variable $Y \sim N(25, 25)$.

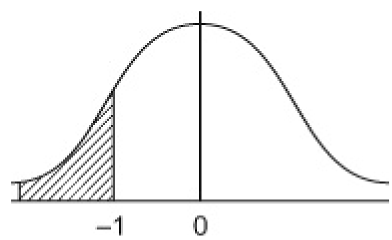
Find **a** $P(Y < 20)$,

b $P(18 < Y < 26)$.

Solution:

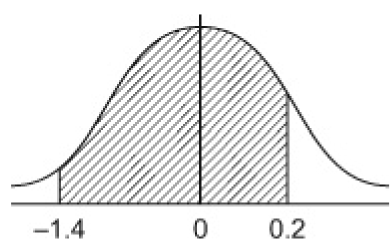
(a)

$$\begin{aligned} P(Y < 20) &= P\left(Z < \frac{20-25}{\sqrt{25}}\right) \\ &= P(Z < -1) \\ &= 1 - 0.8413 \\ &= \underline{\underline{0.1587}} \end{aligned}$$



(b)

$$\begin{aligned} P(18 < Y < 26) &= P\left(\frac{18-25}{5} < Z < \frac{26-25}{5}\right) \\ &= P(-1.4 < Z < 0.2) \\ &= (0.5793 - 0.5) + (0.9192 - 0.5) \\ &= \underline{\underline{0.4985}} \end{aligned}$$



Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Normal distribution
Exercise C, Question 4

Question:

The random variable $X \sim \mathrm{N}(18, 10)$.

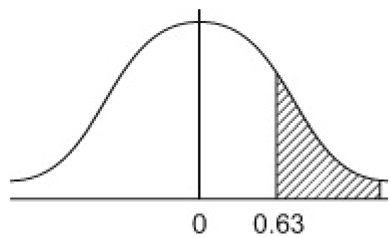
Find **a** $\mathrm{P}(X > 20)$,

b $\mathrm{P}(X < 15)$.

Solution:

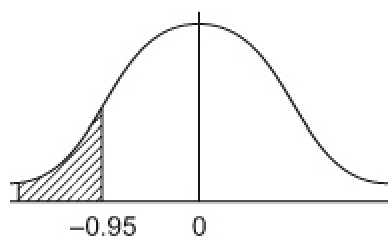
(a)

$$\begin{aligned}\mathrm{P}(X > 20) &= \mathrm{P}\left(Z > \frac{20-18}{\sqrt{10}}\right)\\ &= \mathrm{P}(Z > 0.6324\ldots) \text{ Use } 0.63\\ &= 1 - 0.7357\\ &= \underline{\underline{0.2643}} \text{ (Calculator } 0.26354\ldots\text{.)}\\ &\text{allow AWRT } \underline{\underline{0.264}} \text{ or } \underline{\underline{0.263}}\end{aligned}$$



(b)

$$\begin{aligned}\mathrm{P}(X < 15) &= \mathrm{P}\left(Z < \frac{15-18}{\sqrt{10}}\right)\\ &= \mathrm{P}(Z < -0.9486\ldots) \text{ [Use} -0.95]\\ &= 1 - 0.8289\\ &= \underline{\underline{0.1711}} \text{ (Calculator: } 0.17139\ldots)\\ &\text{allow AWRT } \underline{\underline{0.171}}\end{aligned}$$



Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Normal distribution
Exercise C, Question 5

Question:

The random variable $X \sim N(20, 8)$.

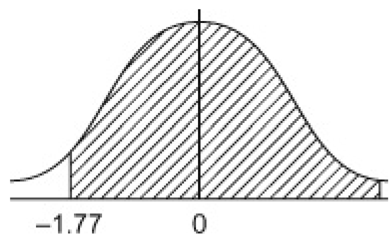
Find **a** $P(X > 15)$,

b the value of a such that $P(X < a) = 0.8051$.

Solution:

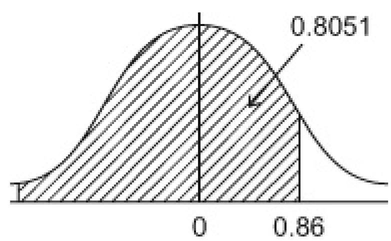
(a)

$$\begin{aligned} P(X > 15) &= P\left(Z > \frac{15-20}{\sqrt{8}}\right) \\ &= P(Z > -1.767\ldots) \text{ Use } -1.77 \\ &= 0.9616 \ \ (\text{Calculator: } 0.96145\ldots) \\ &\text{allow AWRT } \underline{0.961 \text{ or } 0.962} \end{aligned}$$



(b)

$$\begin{aligned} P(X < a) &= 0.8051 \\ P\left(Z < \frac{a-20}{\sqrt{8}}\right) &= 0.8051 \\ \therefore \frac{a-20}{\sqrt{8}} &= 0.86 \\ \therefore a &= \underline{\underline{22.43}} \ \ (\text{allow AWRT } 22.4) \end{aligned}$$



Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Normal distribution
Exercise C, Question 6

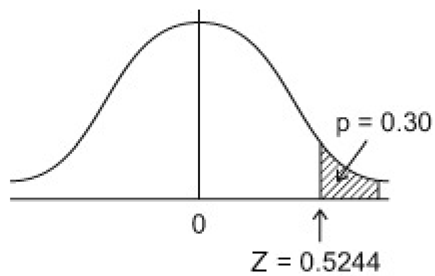
Question:

The random variable $Y \sim N(30, 5^2)$.

Find the value of a such that $P(Y > a) = 0.30$.

Solution:

$$\begin{aligned} P(Y > a) &= 0.30 \\ \frac{a-30}{5} &= 0.5244 \\ \therefore a &= 5 \times 0.5244 + 30 \\ a &= 32.622 \text{ or } \underline{\underline{32.6}} \text{ (3sf)} \end{aligned}$$



Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Normal distribution
Exercise C, Question 7

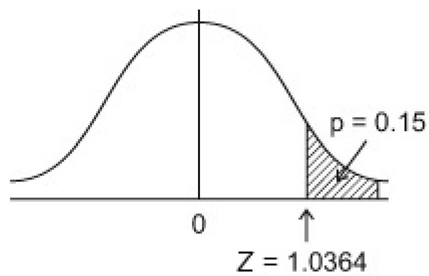
Question:

The random variable $X \sim N(15, 3^2)$.

Find the value of a such that $P(X > a) = 0.15$.

Solution:

$$\begin{aligned} p(X > a) &= 0.15 \\ \frac{a-15}{3} &= 1.0364 \\ \therefore a &= 3 \times 1.0364 + 15 \\ a &= 18.1092 \\ a &= \underline{\underline{18.1}} \text{ (3sf)} \end{aligned}$$



Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Normal distribution
Exercise C, Question 8

Question:

The random variable $X \sim N(20, 12)$.

Find the value of a and the value of b such that

a $P(X < a) = 0.40$,

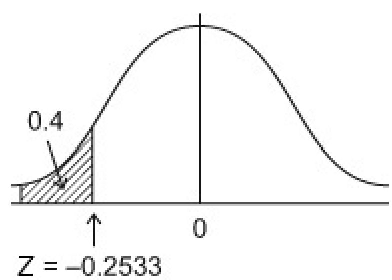
b $P(X > b) = 0.6915$.

c Write down $P(b < X < a)$.

Solution:

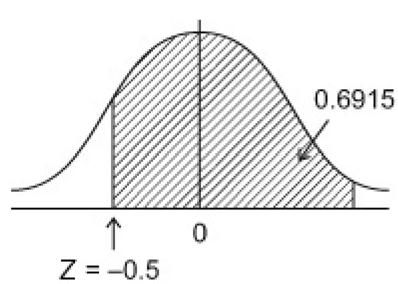
(a)

$$p(X < a) = 0.40 \quad \text{Use} \quad P = 0.4000$$
$$\frac{a-20}{\sqrt{12}} = -0.2533$$
$$a = 19.122\ldots \quad \therefore \quad a = \underline{\underline{19.1}} \ (3sf)$$



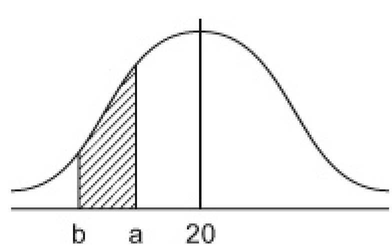
(b)

$$P(X > b) = 0.6915$$
$$\frac{b-20}{\sqrt{12}} = -0.5$$
$$\therefore b = 18.267\ldots \quad \therefore b = \underline{\underline{18.3}} \ (3sf)$$



(c)

$$p(b < X < a)$$
$$= 0.40 - [1 - 0.6915]$$
$$= \underline{\underline{0.0915}}$$



Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Normal distribution
Exercise C, Question 9

Question:

The random variable $Y \sim N(100, 15^2)$.

Find the value of a and the value of b such that

a $P(Y > a) = 0.975$,

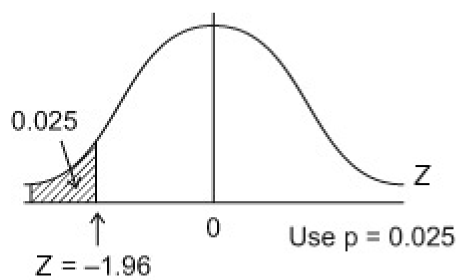
b $P(Y < b) = 0.10$.

c Write down $P(a < Y < b)$.

Solution:

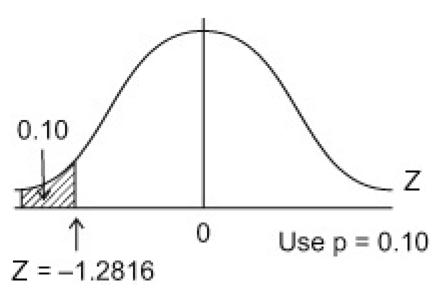
(a)

$$P(Y > a) = 0.975$$
$$\therefore \frac{a-100}{15} = -1.96$$
$$\therefore a = \underline{\underline{70.6}}$$



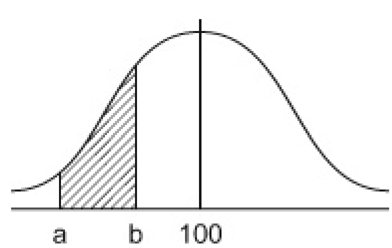
(b)

$$P(Y < b) = 0.10$$
$$\therefore \frac{b-100}{15} = -1.2816$$
$$\therefore b = 80.776 \text{ or } \underline{\underline{80.8}} \text{ (3sf)}$$



(c)

$$P(a < Y < b)$$
$$= 0.10 - 0.025$$
$$= \underline{\underline{0.075}}$$



Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Normal distribution
Exercise C, Question 10

Question:

The random variable X ~ N(80, 16).

Find the value of a and the value of b such that

a $P(X > a) = 0.40$,

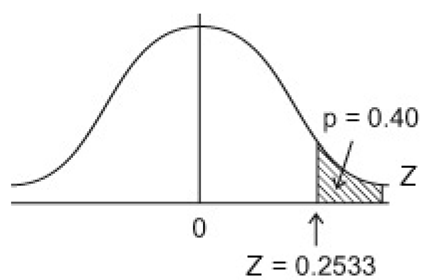
b $P(X < b) = 0.5636$.

c Write down $P(b < X < a)$.

Solution:

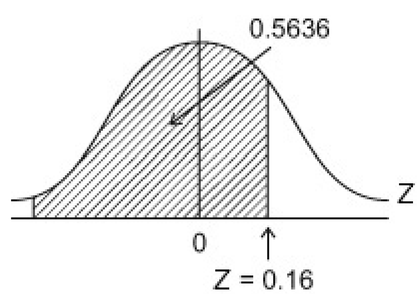
(a)

$$P(X > a) = 0.40$$
$$\therefore \frac{a-80}{\sqrt{16}} = 0.2533$$
$$\therefore a = \underline{\underline{81.0}} \text{ (3sf)}$$



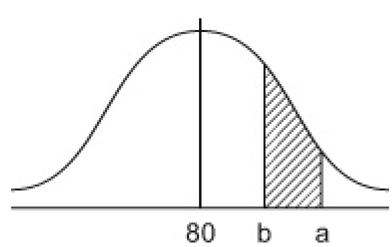
(b)

$$P(X < b) = 0.5636$$
$$\therefore \frac{b-80}{4} = 0.16$$
$$\therefore b = \underline{\underline{80.64}}$$



(c)

$$P(b < X < a)$$
$$= [1 - 0.4] - 0.5636$$
$$= 0.6 - 0.5636$$
$$= \underline{\underline{0.0364}}$$



Solutionbank S1
Edexcel AS and A Level Modular Mathematics

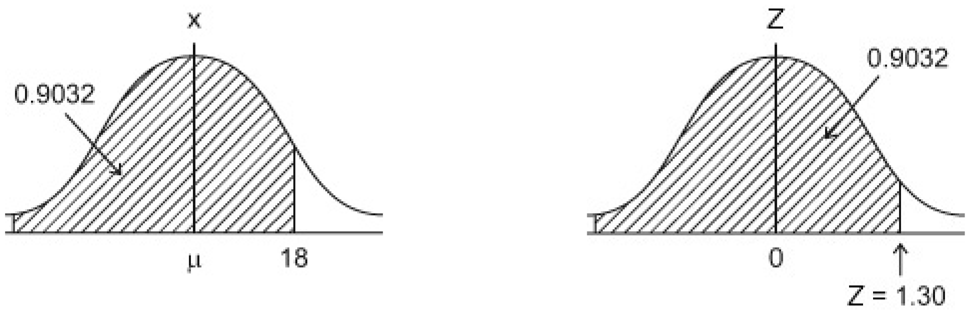
Normal distribution
Exercise D, Question 1

Question:

The random variable $X \sim N(\mu, 5^2)$ and $P(X < 18) = 0.9032$.

Find the value of μ.

Solution:



$$Z = \frac{X - \mu}{\sigma} \quad \Rightarrow \quad 1.30 = \frac{18 - \mu}{5}$$

$$\therefore \quad \mu = 18 - 5 \times 1.30$$

$$\mu = \underline{\underline{11.5}}$$

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

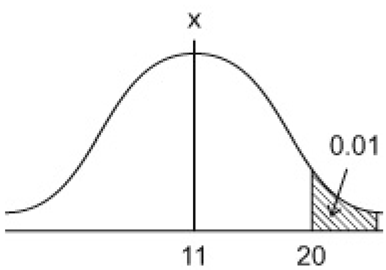
Normal distribution
Exercise D, Question 2

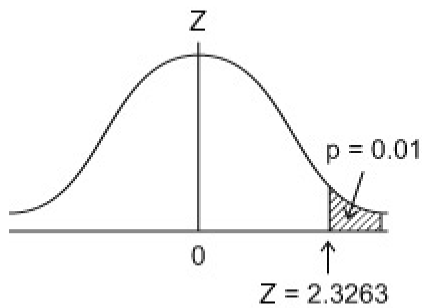
Question:

The random variable $X \sim N(11, \sigma^2)$ and $P(X > 20) = 0.01$.

Find the value of σ.

Solution:





$$Z = \frac{X - \mu}{\sigma} \Rightarrow \quad 2.3263 = \frac{20 - 11}{\sigma}$$

$$\therefore \quad \sigma = \frac{9}{2.3263}$$

$$= 3.8688\ldots$$

$$\therefore \quad \sigma = \underline{\underline{3.87}} \quad \text{(3sf)}$$

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

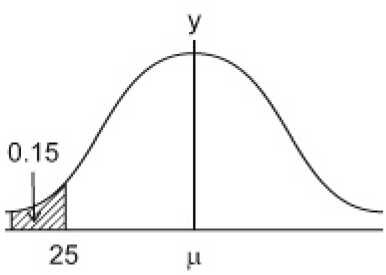
Normal distribution
Exercise D, Question 3

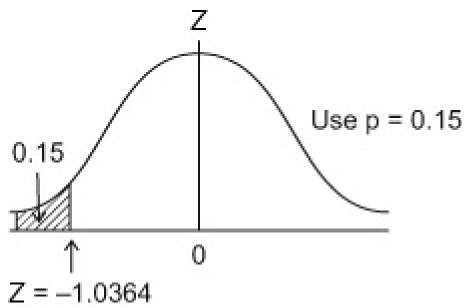
Question:

The random variable $Y \sim N(\mu, 40)$ and $P(Y < 25) = 0.15$.

Find the value of μ.

Solution:





$$Z = \frac{X - \mu}{\sigma} \quad \Rightarrow \quad -1.0364 = \frac{25 - \mu}{\sqrt{40}}$$

$$\mu = 31.554\ldots.$$

$$= \underline{\underline{31.6}} \quad \text{(3sf)}$$

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

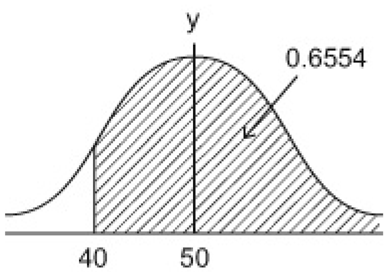
Normal distribution
Exercise D, Question 4

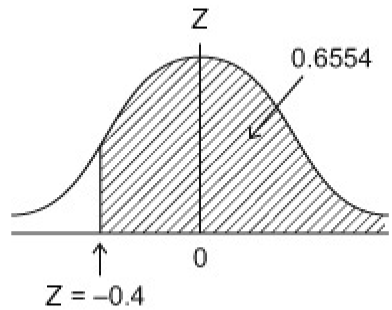
Question:

The random variable $Y \sim N(50, \sigma^2)$ and $P(Y > 40) = 0.6554$.

Find the value of σ.

Solution:





$$Z = \frac{X-\mu}{\sigma} \quad \Rightarrow \quad -0.4 = \frac{40-50}{\sigma}$$

$$\therefore \quad \sigma = \frac{10}{0.4}$$

$$\sigma = \underline{\underline{25}}$$

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

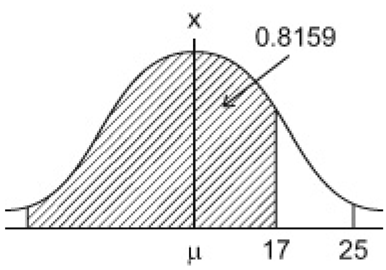
Normal distribution
Exercise D, Question 5

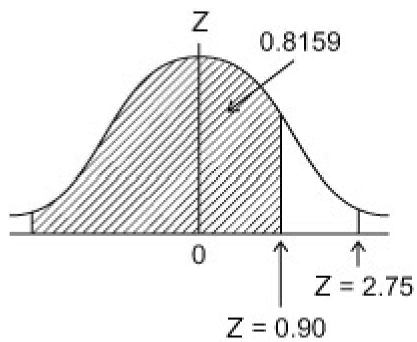
Question:

The random variable $X \sim N(\mu, \sigma^2)$.

Given that $P(X < 17) = 0.8159$ and $P(X < 25) = 0.9970$, find the value of μ and the value of σ.

Solution:





$Z = \frac{X-\mu}{\sigma} \Rightarrow \quad 0.90\ \sigma = 17 - \mu$

$\underline{2.75\sigma = 25 - \mu}$

Subtract $\quad 1.85\ \sigma = 8$

$\therefore \quad \sigma = \frac{8}{1.85} = 4.3243$

$\mu = 17 - 0.90 \times \sigma.$

$= 13.1081$

$\therefore \quad \underline{\underline{\mu = 13.1}} \quad \sigma = \underline{\underline{4.32}} \quad (3sf)$

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

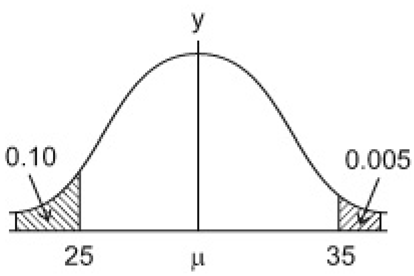
Normal distribution
Exercise D, Question 6

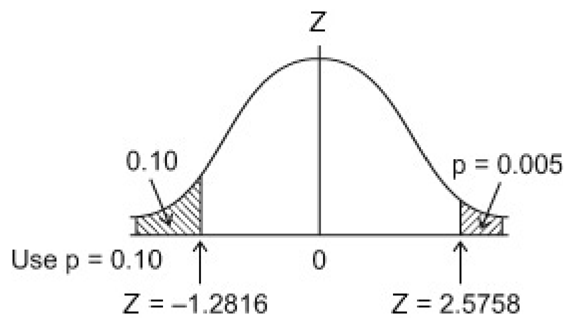
Question:

The random variable $Y \sim N(\mu, \sigma^2)$.

Given that $P(Y < 25) = 0.10$ and $P(Y > 35) = 0.005$, find the value of μ and the value of σ.

Solution:





$$Z = \frac{X-\mu}{\sigma} \Rightarrow 2.5758\sigma = 35 - \mu$$
$$\underline{-1.2816\sigma = 25 - \mu.}$$

Subtract $3.8574\sigma = 10$

$\therefore \sigma = 2.59241 \ldots..$

$\therefore \mu = 35 - 2.5758 \times 2.59241 \ldots.$

$= 28.322 \ldots$

$\therefore \mu = \underline{\underline{28.3}}$, $\sigma = \underline{\underline{2.59}}$ (3sf)

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

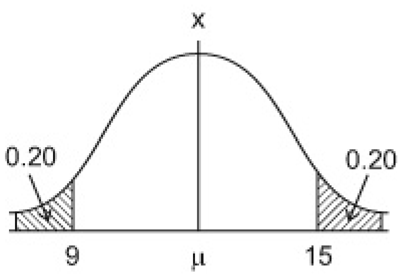
Normal distribution
Exercise D, Question 7

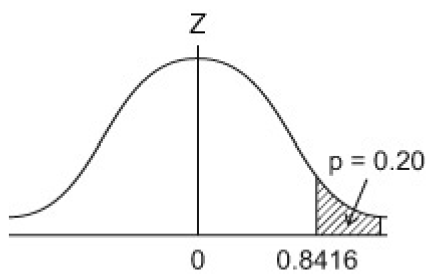
Question:

The random variable $X \sim N(\mu, \sigma^2)$.

Given that $P(X > 15) = 0.20$ and $P(X < 9) = 0.20$, find the value of μ and the value of σ.

Solution:





By symmetry $\mu = \frac{1}{2}(9 + 15) = \underline{\underline{12}}$

$$Z = \frac{X - \mu}{\sigma} \Rightarrow 0.8416 = \frac{15 - \mu}{\sigma} = \frac{3}{\sigma}$$

$$\therefore \quad \sigma = \frac{3}{0.8416}$$

$$\sigma = \underline{\underline{3.56}}$$

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Normal distribution
Exercise D, Question 8

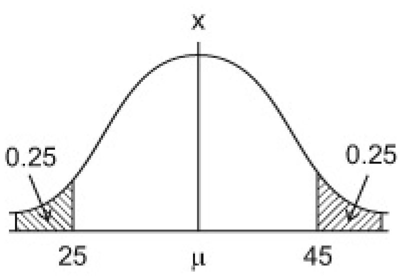
Question:

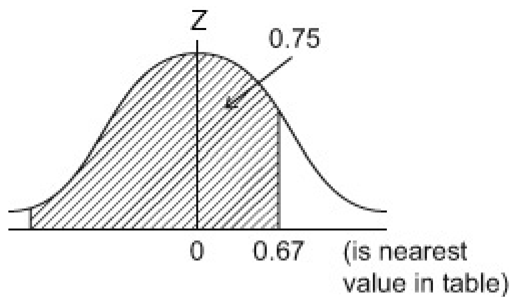
The random variable $X \sim N(\mu, \sigma^2)$.

The lower quartile of X is 25 and the upper quartile of X is 45.

Find the value of μ and the value of σ.

Solution:





By Symmetry $\mu = \frac{1}{2}(25+45) = \underline{\underline{35}}$

$$Z = \frac{X-\mu}{\sigma} \quad \Rightarrow \quad 0.67 = \frac{45-35}{\sigma}$$

$$\therefore \quad \sigma = \frac{10}{0.67}$$

$$\sigma = 14.92 \quad (\text{or } 14.826\ldots \text{ on Calculator})$$

$$\therefore \quad \text{accept} \quad \sigma = \underline{\underline{14.9 \text{ or } 14.8}}$$

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

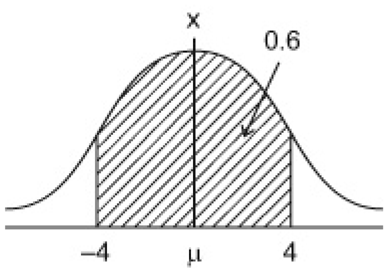
Normal distribution
Exercise D, Question 9

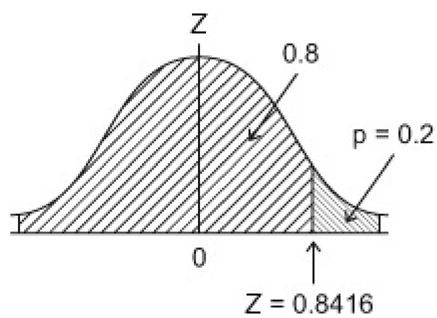
Question:

The random variable $X \sim N(0, \sigma^2)$.

Given that $P(-4 < X < 4) = 0.6$, find the value of σ.

Solution:





By Symmetry $\mu = \underline{\underline{0}}$

$$Z = \frac{X-\mu}{\sigma} \quad \Rightarrow \quad 0.8416 = \frac{4}{\sigma}$$

$$\therefore \quad \sigma = \frac{4}{0.8416}$$

$$\sigma = \underline{\underline{4.75}} \quad \text{3sf}$$

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

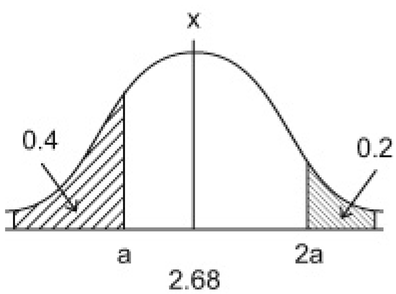
Normal distribution
Exercise D, Question 10

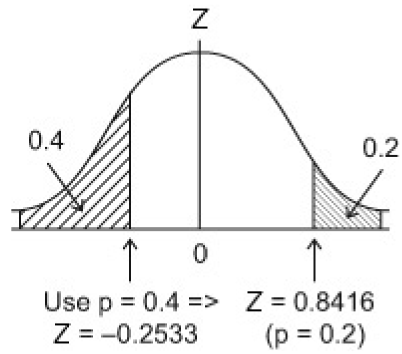
Question:

The random variable $X \sim N(2.68, \sigma^2)$.

Given that $P(X > 2a) = 0.2$ and $P(X < a) = 0.4$, find the value of σ and the value of a.

Solution:





$$Z = \frac{X-\mu}{\sigma} \Rightarrow \quad 0.8416\sigma = 2a - 2.68$$

$$\underline{-0.2533\sigma = a - 2.68} \times 2$$
$$-0.5066\sigma = 2a - 5.36$$
$$\underline{0.8416\sigma = 2a - 2.68}$$

Subtract : $1.3482\sigma = 2.68$

$\sigma = 1.9878\ldots$

$\sigma = \underline{\underline{1.99}}$ 3sf

$a = 2.68 - 0.2533\sigma$

$\therefore$ $a = 2.176\ldots$

$a = \underline{\underline{2.18}}$ 3sf

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Normal distribution
Exercise E, Question 1

Question:

The heights of a large group of men are normally distributed with a mean of 178 cm and a standard deviation of 4 cm.

A man is selected at random from this group.

a Find the probability that he is taller than 185 cm.

A manufacturer of door frames wants to ensure that fewer than 0.005 men have to stoop to pass through the frame.

b On the basis of this group, find the minimum height of a door frame.

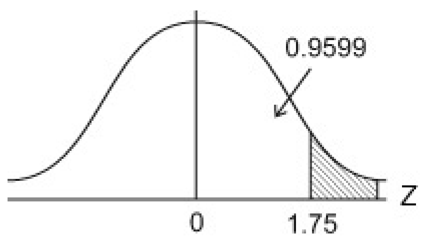
Solution:

M~ N(178, 4^2)

(a)

$$P(M > 185) = P\left(Z > \frac{185 - 178}{4}\right)$$

$$= P(Z > 1.75)$$
$$= 1 - 0.9599$$
$$= \underline{\underline{0.0401}}$$

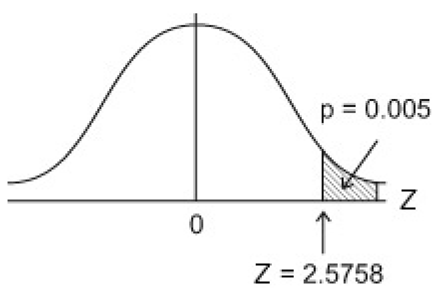


(b)

P(M>h) = 0.005

$$Z = \frac{X - \mu}{\sigma} \Rightarrow \quad 2.5758 = \frac{h - 178}{4}$$

$$\therefore \quad h = 188.3032$$
$$\therefore \quad h = \underline{\underline{188cm}}$$



Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Normal distribution
Exercise E, Question 2

Question:

The weights of steel sheets produced by a factory are known to be normally distributed with mean 32.5 kg and standard deviation 2.2 kg.

a Find the percentage of sheets that weigh less than 30 kg.

Bob requires sheets that weigh between 31.6 kg and 34.8 kg.

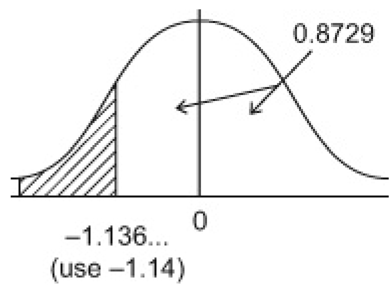
b Find the percentage of sheets produced that satisfy Bob's requirements.

Solution:

$W \sim N(32.5, 2.2^2)$

(a)

$$\begin{aligned} P(W < 30) &= P\left(Z < \frac{30 - 32.5}{2.2}\right) \\ &= P(Z < -1.14) \\ &= 1 - 0.8729 \\ &= 0.1271 \quad \underline{\underline{12.7\,\% \text{ or } 12.8\,\%}} \end{aligned}$$

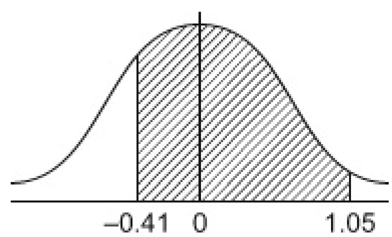


(Calculator gives 0.1279.. so allow AWRT (0.127 - 0.128))

(b)

$P(31.6 < W < 34.8)$

$$\begin{aligned} &= P\left(\frac{31.6 - 32.5}{2.2} < Z < \frac{34.8 - 32.5}{2.2}\right) \\ &= P(-0.41 < Z < 1.05) \\ &= 0.8531 - (1 - 0.6591) \\ &= 0.5122 \end{aligned}$$



(Calculator gives 0.510856 … So allow AWRT 0.511 or 0.512)

So 51.1% or 51.2% of sheets satisfy Bob's requirements

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Normal distribution
Exercise E, Question 3

Question:

The time a mobile phone battery lasts before needing to be recharged is assumed to be normally distributed with a mean of 48 hours and a standard deviation of 8 hours.

a Find the probability that a battery will last for more than 60 hours.

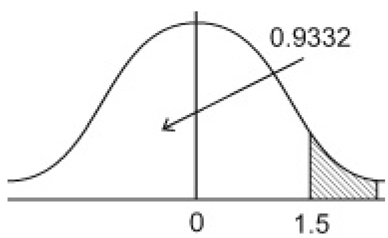
b Find the probability that the battery lasts less than 35 hours.

Solution:

T ~ N(48, 8^2)

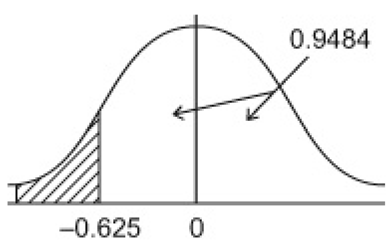
(a)

$$\begin{aligned} P(T>60) &= P\left(Z > \frac{60-48}{8}\right) \\ &= P(Z > 1.5) \\ &= 1 - 0.9332 \\ &= \underline{\underline{0.0668}} \end{aligned}$$



(b)

$$\begin{aligned} P(T<35) &= P\left(Z < \frac{35-48}{8}\right) \\ &= P(Z < -1.63) \\ &= 1 - 0.9484 \\ &= 0.0516 \end{aligned}$$



(Calculator gives 0.05208… so allow AWRT <u>0.052</u>)

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Normal distribution
Exercise E, Question 4

Question:

The random variable $X \sim N(24, \sigma^2)$.

Given that $P(X > 30) = 0.05$, find

a the value of σ,

b $P(X < 20)$,

c the value of d so that $P(X > d) = 0.01$.

Solution:

$X \sim N(24, \sigma^2)$

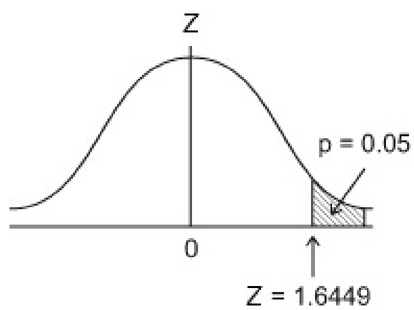
(a)

$P(X > 30) = 0.05$

$$Z = \frac{X-\mu}{\sigma} \Rightarrow \quad 1.6449 = \frac{30-24}{\sigma}$$

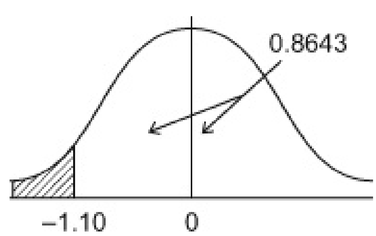
$$\therefore \quad \sigma = \frac{6}{1.6449} = 3.6476\ldots$$

$$\therefore \quad \sigma = \underline{3.65(3 \text{ sf})}$$



(b)

$$P(X < 20) = P\left(Z < \frac{20-24}{\sigma}\right)$$
$$= P(Z < -1.10)$$
$$= 1 - 0.8643$$
$$= \underline{\underline{0.1357}}$$



(Calculator gives 0.1364… so allow AWRT $\underline{0.136}$)

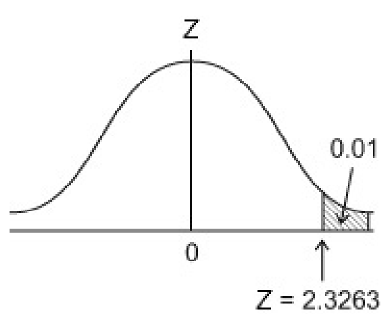
(c)

$P(X > d) = 0.01$

$$Z = \frac{X-\mu}{\sigma} \Rightarrow 2.3263 = \frac{d-24}{\sigma}$$

$$\therefore \quad d = 32.485\ldots.$$

$$d = \underline{\underline{32.5}} \quad (3\text{sf})$$



Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Normal distribution
Exercise E, Question 5

Question:

A machine dispenses liquid into plastic cups in such a way that the volume of liquid dispensed is normally distributed with a mean of 120 ml. The cups have a capacity of 140 ml and the probability that the machine dispenses too much liquid so that the cup overflows is 0.01.

a Find the standard deviation of the volume of liquid dispensed.

b Find the probability that the machine dispenses less than 110 ml.

Ten percent of customers complain that the machine has not dispensed enough liquid.

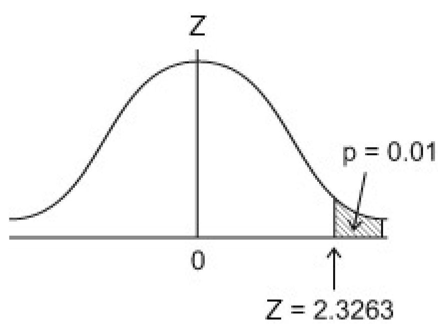
c Find the largest volume of liquid that will lead to a complaint.

Solution:

$L \sim N(120, \sigma^2)$

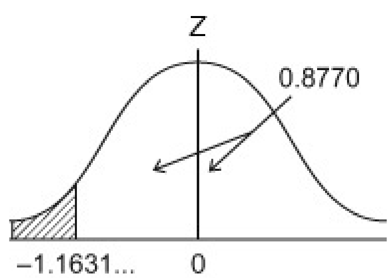
(a)

$$\begin{aligned} P(L > 140) &= 0.01 \\ Z = \frac{X-\mu}{\sigma} \Rightarrow 2.3263 &= \frac{140-120}{\sigma} \\ \therefore \quad \sigma &= 8.5973\ldots = \underline{\underline{8.60\ \text{ml}}} \end{aligned}$$



(b)

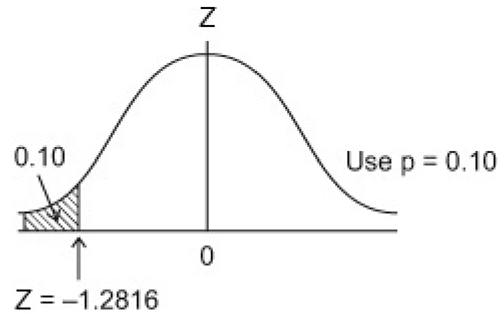
$$\begin{aligned} P(L < 110) &= P\left(Z < \frac{110-120}{\sigma}\right) \\ &= P(Z < -1.16) \\ &= 1 - 0.8770 \\ &= 0.123 \end{aligned}$$



(Calculator gives 0.12238… so allow AWRT 0.122 or 0.123)

(c)

$$\begin{aligned} P(L < c) &= 0.10 \\ Z = \frac{X-\mu}{\sigma} \Rightarrow -1.2816 &= \frac{c-120}{\sigma} \\ \therefore \quad c &= 108.98\ldots \\ &= \underline{\underline{109\ \text{ml}}} \quad \text{(3sf)} \end{aligned}$$


Z
0.10
Use p = 0.10
0
Z = –1.2816

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

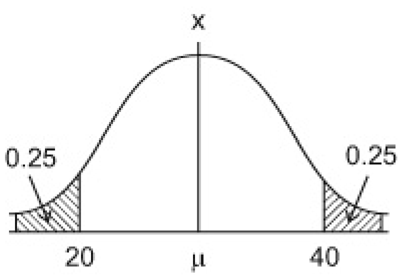
Normal distribution
Exercise E, Question 6

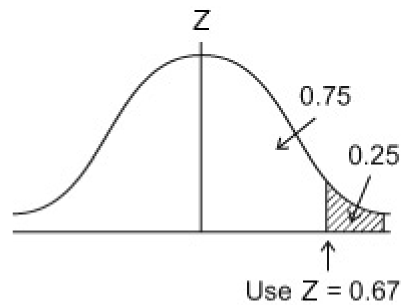
Question:

The random variable $X \sim N(\mu, \sigma^2)$. The lower quartile of X is 20 and the upper quartile is 40. Find μ and σ.

Solution:

$X \sim N(\mu, \sigma^2)$





By symmetry $\mu = \frac{1}{2}(20 + 40) = \underline{\underline{30}}$

$$Z = \frac{X - \mu}{\sigma} \Rightarrow \frac{40 - 30}{\sigma} = 0.67$$

$$\therefore \quad \sigma = 14.925\ldots..$$

(Calculator gives 14.82… so allow AWRT <u>14.8 or 14.9</u>)

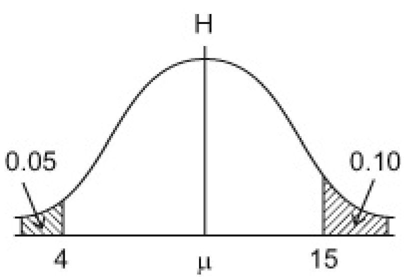
Solutionbank S1
Edexcel AS and A Level Modular Mathematics

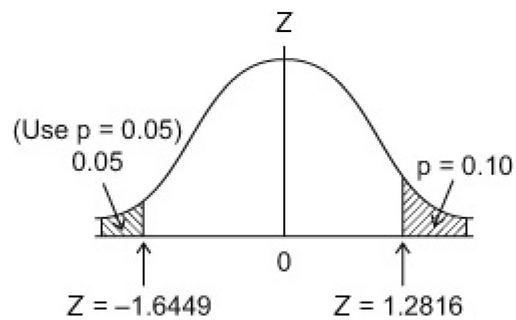
Normal distribution
Exercise E, Question 7

Question:

The heights of seedlings are normally distributed. Given that 10% of the seedlings are taller than 15 cm and 5% are shorter than 4 cm, find the mean and standard deviation of the heights.

Solution:





$Z\sigma = X - \mu \Rightarrow \quad -1.6449\ \sigma = 4 - \mu$

$\underline{1.2816\ \sigma = 15 - \mu}$

Subtract $\quad 2.9265\ \sigma = 11$

$\therefore \quad \sigma = 3.7587\ldots = \underline{\underline{3.76}}$ cm (3sf)

$\mu = 15 - 1.2816\ \sigma$

$\therefore \quad \mu = 10.2$ cm (3sf)

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Normal distribution
Exercise E, Question 8

Question:

A psychologist gives a student two different tests. The first test has a mean of 80 and a standard deviation of 10 and the student scored 85.

a Find the probability of scoring 85 or more on the first test.

The second test has a mean of 100 and a standard deviation of 15. The student scored 105 on the second test.

b Find the probability of a score of 105 or more on the second test.

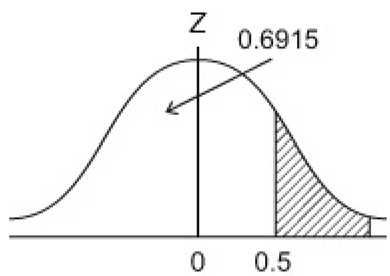
c State, giving a reason, which of the student's two test scores was better.

Solution:

$T \sim N(80, 10^2)$

(a)

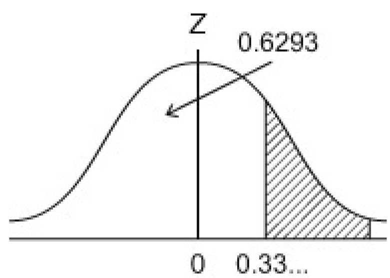
$$\begin{aligned} P(T > 85) &= P\left(Z > \frac{85-80}{10}\right) \\ &= P(Z > 0.5) \\ &= 1 - 0.6915 \\ &= \underline{\underline{0.3085}} \end{aligned}$$



(b)

$S \sim N(100, 15^2)$

$$\begin{aligned} P(S > 105) &= P\left(Z > \frac{105-100}{15}\right) \\ &= P(Z > 0.33) \\ &= 1 - 0.6293 \\ &= 0.3707 \end{aligned}$$



(Calculator gives 0.36944… so allow 0.369, 0.370 or 0.371)

(c) 1st score is best since a lower proportion of scores will beat it. (or Z value of 1st test in higher so this is the better result)

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Normal distribution
Exercise E, Question 9

Question:

Jam is sold in jars and the mean weight of the contents is 108 grams. Only 3% of jars have contents weighing less than 100 grams. Assuming that the weight of jam in a jar is normally distributed find

a the standard deviation of the weight of jam in a jar,

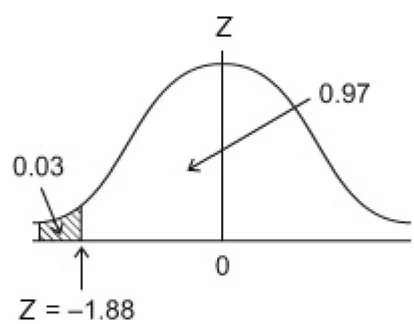
b the proportion of jars where the contents weigh more than 115 grams.

Solution:

J~ N(108, σ^2)

(a)

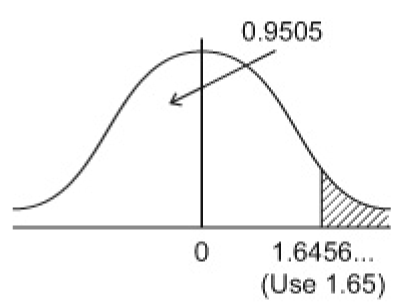
$$\text{P}(J < 100) = 0.03$$
$$Z = \frac{X-\mu}{\sigma} \Rightarrow -1.88 = \frac{100-108}{\sigma}$$
$$\therefore \quad \sigma = 4.255 \ldots..$$



(Calculator gives 4.2535… so allow AWRT <u>4.25–4.26</u>)

(b)

$$\text{P}(J > 115) = \text{P}\left(Z > \frac{115-108}{\sigma}\right)$$
$$= \text{P}(Z > 1.65)$$
$$= 1 - 0.9505$$
$$= 0.0495$$



(Calculator gives: 0.0499… so allow AWRT 0.050)

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Normal distribution
Exercise E, Question 10

Question:

The waiting time at a doctor's surgery is assumed to be normally distributed with standard deviation of 3.8 minutes. Given that the probability of waiting more than 15 minutes is 0.0446, find

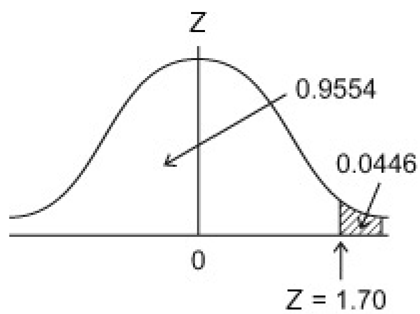
a the mean waiting time,

b the probability of waiting fewer than 5 minutes.

Solution:

$T \sim N(\mu, 3.8^2)$

$P(T > 15) = 0.0446$



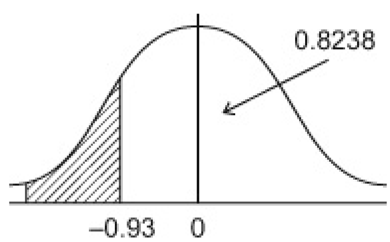
(a) $Z = \frac{X-\mu}{\sigma} \Rightarrow 1.70 = \frac{15-\mu}{3.8}$

$\therefore \mu = 15 - 3.8 \times 1.70$

$\mu = \underline{8.54}$ (3sf) minutes

(b)

$$P(T < 5) = P\left(Z < \frac{5-8.54}{3.8}\right)$$
$$= P(Z < -0.93\ldots)$$
$$= 1 - 0.8238$$
$$= \underline{\underline{0.1762}}$$



(Calculater gives 0.17577… so allow AWRT $\underline{0.176}$)

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Normal distribution
Exercise E, Question 11

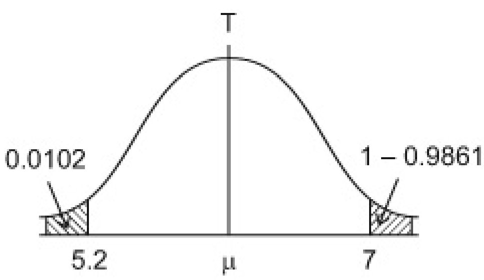
Question:

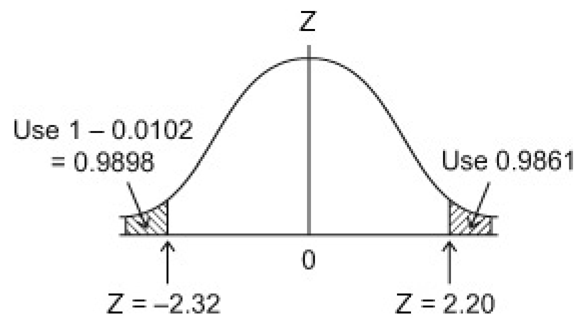
The thickness of some plastic shelving produced by a factory is normally distributed. As part of the production process the shelving is tested with two gauges. The first gauge is 7 mm thick and 98.61% of the shelving passes through this gauge. The second gauge is 5.2 mm thick and only 1.02% of the shelves pass through this gauge.

Find the mean and standard deviation of the thickness of the shelving.

Solution:

$T \sim N(\mu, \sigma^2)$





$$Z\sigma = X - \mu \quad \Rightarrow 2.20\ \sigma = 7 - \mu$$
$$\underline{-2.32\ \sigma = 5.2 - \mu}$$

Subtract $4.52\ \sigma = 1.8$
$\sigma = 0.3982\ldots.$

$\mu = 7 - 2.20\ \sigma \ \Rightarrow\ \mu = 6.1238\ldots.$
$\therefore\ \underline{\underline{\mu = 6.12\ \text{mm}, \sigma = 0.398\ \text{mm}}}$

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Normal distribution
Exercise E, Question 12

Question:

The random variable $X \sim \mathrm{N}(14, 9)$. Find

a $\mathrm{P}(X \geq 11)$,

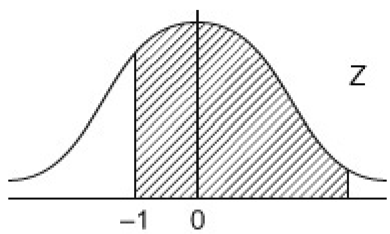
b $\mathrm{P}(9 < X < 11)$.

Solution:

$X \sim \mathrm{N}(14, 3^2)$

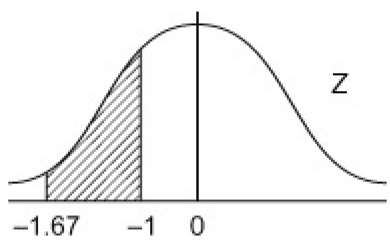
(a)

$$\begin{aligned}\mathrm{P}(X \geq 11) &= \mathrm{P}\left(Z \geq \frac{11-14}{3}\right)\\ &= \mathrm{P}(Z \geq -1)\\ &= \underline{\underline{0.8413}}\end{aligned}$$



(b)

$$\begin{aligned}\mathrm{P}(9 < X < 11) &= \mathrm{P}\left(\frac{9-14}{3} < Z < \frac{11-14}{3}\right)\\ &= \mathrm{P}(-1.67 < Z < -1)\\ &= 0.9525 - 0.8413\\ &= 0.1112\end{aligned}$$



(Calculater gives: 0.11086 … so allow AWRT $\underline{0.111}$)

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Normal distribution
Exercise E, Question 13

Question:

The random variable $X \sim \mathrm{N}(20, 5^2)$. Find

a $\mathrm{P}(X \leq 16)$,

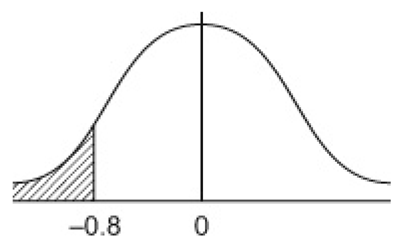
b the value of d such that $\mathrm{P}(X < d) = 0.95$.

Solution:

$X \sim N(20, 5^2)$

(a)

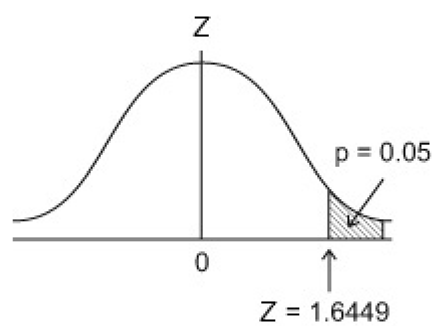
$$\begin{aligned} \mathrm{P}(X \leq 16) &= \mathrm{P}\left(Z < \frac{16-20}{5}\right) \\ &= \mathrm{P}(Z < -0.8) \\ &= 1 - 0.7881 \\ &= \underline{\underline{0.2119}} \end{aligned}$$



(b)

$\mathrm{P}(X < d) = 0.95$

$$Z = \frac{X-\mu}{\sigma} \Rightarrow 1.6449 = \frac{d-20}{5}$$
$$\therefore d = 20 + 5 \times 1.6449$$
$$d = 28.2$$



Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Examination style paper
Exercise A, Question 1

Question:

1 A fair die has six faces numbered 1, 1, 1, 2, 2 and 3. The die is rolled twice and the number showing on the uppermost face is recorded.

Find the probability that the sum of the two numbers is at least three. (5)

Solution:

3	_4_	_4_	_4_	_5_	_5_	_6_
2	_3_	_3_	_3_	_4_	_4_	_5_
2	_3_	_3_	_3_	_4_	_4_	_5_
1	2	2	2	_3_	_3_	_4_
1	2	2	2	_3_	_3_	_4_
1	2	2	2	_3_	_3_	_4_
Second First	**1**	**1**	**1**	**2**	**2**	**3**

The easiest solution involves drawing a diagram to represent the sample space. Each square is the sum of the scores on the die. The first method mark is for attempting the diagram and the second is an accuracy mark for all the values correct.

Each of the values that are 'at least 3' are underlined; 3, 4, 5 & 6.

M1A1A1

$$\text{P(Sum at least } 3) = \frac{27}{36} = \frac{3}{4}$$

There are 27 values underlined and 36 values in the sample space. Then cancel the fraction.
M1A1

ALTERNATIVE SOLUTION

Let D_1 = the number on the first die
and D_2 = the number on the second die

$$\begin{aligned}\mathrm{P}(D_1 + D_2 \geq 3) &= 1 - \mathrm{P}(D_1 + D_2 = 2) \\ &= 1 - \mathrm{P}(D_1 = 1 \text{ and } D_2 = 1) \\ &= 1 - \mathrm{P}(D_1 = 1) \times \mathrm{P}(D_2 = 1) \\ &= 1 - \frac{1}{2} \times \frac{1}{2} \\ &= \frac{3}{4}\end{aligned}$$

This is a slightly quicker solution.

$\mathrm{P}(D = 1) = 0.5$ and D_1 and D_2 are independent so the probabilities are multiplied together.

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Examination style paper
Exercise A, Question 2

Question:

Jars are filled with jam by a machine. Each jar states it contains 450 g of jam.

The actual weight of jam in each jar is normally distributed with mean 460 g and standard deviation of 10 g.

a Find the probability that a jar contains less then the stated weight of 450 g. (3)

Jars are sold in boxes of 30.

b Find the expected number of jars containing less than the stated weight. (2)

The standard deviation is still 10 g. The mean weight is changed so that 1% of the jars contain less than the stated weight of 450 g of jam.

c Find the new mean weight of jam. (4)

Solution:

(a)

$$\mathrm{P}(X < 450) = \mathrm{P}\left(Z < \frac{450 - 460}{10}\right) = \mathrm{P}(Z < -1.0)$$

$$= 1 - 0.8413 = 0.1587$$

Standardise by subtracting the mean and dividing by the standard deviation gets the first method mark and the z value of -1.0 gets the accuracy mark.

(b)

Expected number of jars = 30×0.1587
$= 4.761$ or 4.76 or 4.8

(c)

$$\mathrm{P}(X < 450) = 0.01$$

$$\frac{450 - \mu}{10} = -2.3263$$

$$\mu = 473.263 = 473 \text{ to } 3 \text{ sf}$$

Forming the correct equation with the new mean as an unknown gets the method and accuracy mark, the B mark is awarded for getting −2.3263 from the tables

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Examination style paper
Exercise A, Question 3

Question:

A discrete random variable X has a probability distribution as shown in the table below

x	0	1	2	3
P(X = x)	0.2	0.3	b	$2a$

where a and b are constants.

If E(X) = 1.6,

a show that $b = 0.2$ and find the value of a. (5)

Find

b E(5 − 2X), (2)

c Var(X), (3)

d Var(5 − 2X). (2)

Solution:

(a)

$0.5 + b + 2a = 1$

Remember that adding all the probabilities together equals 1.

$0.3 + 2b + 6a = 1.6$
Solving
$a = 0.15$, $b = 0.2$

The second equation is formulated from the value of the expectation. Multiply the values of X by the associated probabilities and equate to 1.6.

(b)

$$\begin{aligned} E(5-2X) &= 5 - 2E(X) \\ &= 5 - 2 \times 1.6 = 1.8 \end{aligned}$$

(c)

$$\text{Var}(X) = 1^2 \times 0.3 + 2^2 \times 0.2 + 3^2 \times 0.3 - 1.6^2$$
$$= 1.24$$

For the variance you square each value of x and multiply by the probability. Remember to subtract the square of the expectation.

(d)

Var (5 − 2X) = 4 × Var(X) = 4.96

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Examination style paper
Exercise A, Question 4

Question:

The attendance at college of a group of 20 students was recorded for a four week period. The numbers of students attending each of the 16 classes are shown below.

20	20	19	19
18	19	18	20
20	16	19	20
17	19	20	18

a Calculate the mean and the standard deviation of the attendance. (4)

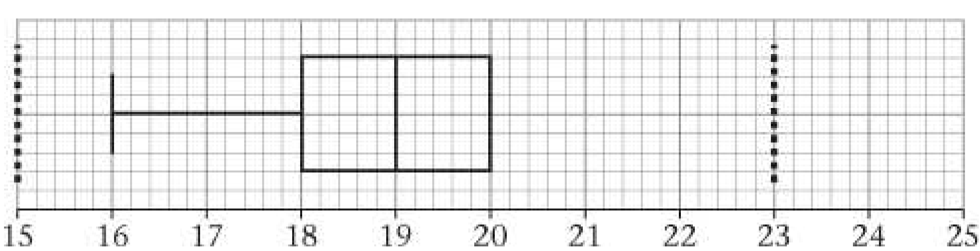
b Express the mean as a percentage of the 20 students in the group. (1)

In the same four week period, the attendance of a different group of 22 students was recorded.

22	18	20	21
17	16	16	17
20	17	18	19
18	20	17	16

c Find the mode, median and inter-quartile range for this group of students. (3)

A box plot for the first group of students is drawn below.



d Using the same scale draw a box plot for the second group of students. (3)

The mean percentage attendance and standard deviation for the second group of students are 82.95 and 1.82 respectively.

e Compare and contrast the attendance of each group of students. (3)

Solution:

(a)

$\overline{x} = \frac{302}{16} = 18.875$

standard deviation is $\sqrt{\frac{5722}{16} - 18.875^2} = \sqrt{1.359375}$

Set out your working clearly so you will still be given the method mark if you make a calculator error.

$= 1.16592\ldots$

(b) mean % attendance is $\frac{18.875}{20} \times 100 = 94.375$

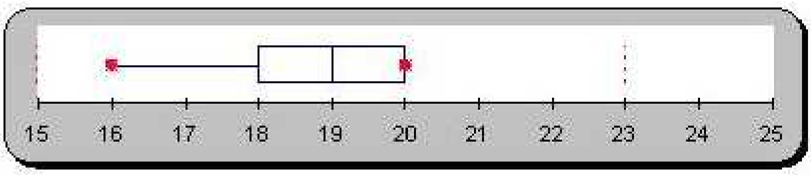
(c) Mode is 17

Median is 18

IQR is $20 - 17 = 3$

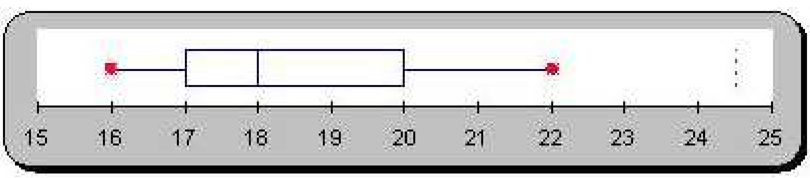
(d)

First Group:



Second Group:

Put the box plots side by side so you can compare easily.



(e)

First mean % > Second mean %

First IQR < Second IQR

First sd < Second sd

First range < Second range

First negative skew, given by whiskers, symmetric by box

Second positive skew.

There are 3 marks for this part, so 3 different correct comments are required. Try to comment about location, spread and shape.

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Examination style paper
Exercise A, Question 5

Question:

The random variable X has the distribution

$P(X = x) = \frac{1}{n}$ for $x = 1, 2, \ldots, n$.

a Write down the name of the distribution. (1)

Given that $E(X) = 10$,

b show that $n = 19$, (2)
c Find $\text{Var}(X)$. (2)

Solution:

(a) Discrete uniform distribution

(b)

$$\frac{(n+1)}{2} = 10$$
$$n = 19$$

Learning the details of the uniform distribution and formulae for mean and variance make this question easier.

(c) $\frac{(n+1)(n-1)}{12} = 30$

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Examination style paper
Exercise A, Question 6

Question:

A researcher thinks that there is a link between a person's confidence and their height. She devises a test to measure the confidence, c, of nine people and their height, h cm. The data are shown in the table below.

h	179	169	187	166	162	193	161	177	168
c	569	561	579	561	540	598	542	565	573

[$\Sigma h = 1562$, $\Sigma c = 5088$, $\Sigma hc = 884\ 484$, $\Sigma h^2 = 272094$, $\Sigma c^2 = 2878966$]

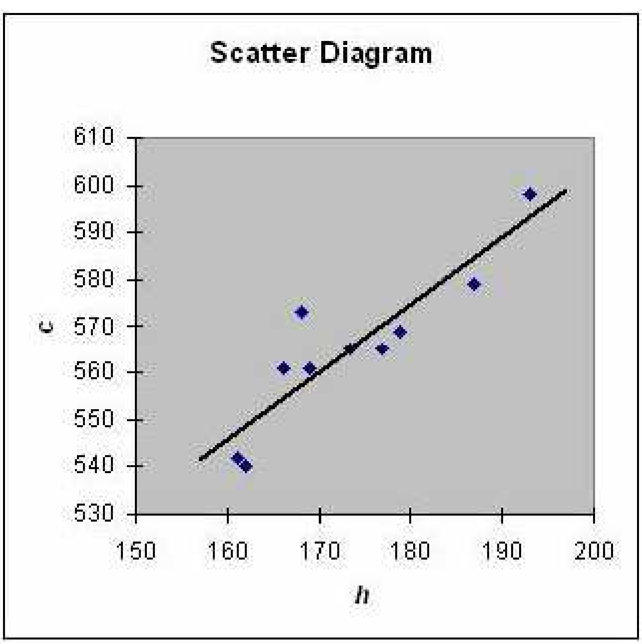
a Draw a scatter diagram to represent these data. (2)
b Find the values of S_{hh}, S_{cc} and S_{hc}. (3)
c Calculate the value of the product moment correlation coefficient. (3)
d Calculate the equation of the regression line of c on h. (4)
e Draw this line on your scatter diagram. (2)
f Interpret the gradient of the regression line. (1)

The researcher decides to use this regression model to predict a person's confidence.

g Find the proposed confidence for the person who has a height of 172 cm. (2)

Solution:

(a) & (e)



Be careful when plotting the points. Make sure the regression line passes through this point. $(\bar{h}, \bar{c})$

B1B1 (2) for points, B1B1 (2) for line.

(b)

$$S_{hh} = 272094 - \frac{1562^2}{9} = 1000.\dot{2}$$
$$S_{cc} = 2878966 - \frac{5088^2}{9} = 2550$$
$$S_{hc} = 884484 - \frac{1562 \times 5088}{9} = 1433.\dot{3}$$

Make sure you work accurately as all these marks are for the answers.

(c)

$$r = \frac{S_{hc}}{\sqrt{S_{hh}S_{cc}}} = \frac{1433.\dot{3}}{\sqrt{1000.\dot{2} \times 2550}} = 0.897488$$

Don't forget the square root.

(d)

$$b = \frac{1433.\dot{3}}{1000.\dot{2}} = 1.433015$$
$$a = \frac{5088}{9} - b \times \frac{1562}{9} = 316.6256$$
$$c = 1.43h + 317$$

Set out your working carefully and remember to subtract the b in the second equation.

(e) See Graph

(f)

For every 1 cm increase in height, the confidence measure increases by 1.43.

This must be in context i.e. it relates to 'height' and 'confidence measure'.

(g)

$$h = 172$$
$$c = 1.43 \times 172 + 317 = 563 \text{ to 3 sf}$$

Substituting h = 172 into your equation gets the method mark.

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Examination style paper
Exercise A, Question 7

Question:

A fairground game involves trying to hit a moving target with an air rifle pellet.

Each player has up to three pellets in a round. Five points are scored if a pellet hits the target, but the round is over if a pellet misses the target.

Jean has a constant probability of 0.4 of hitting the target.

The random variable X is the number of points Jean scores in a round.

Find

a the probability that Jean scores 15 points in a round, (2)
b the probability distribution of X. (5)

A game consists of two rounds.

c Find the probability that Jean scores more points in her second round than her first. (6)

Solution:

(a) P(Scores 15 points)

$= \text{P(hit,hit,hit)} = 0.4 \times 0.4 \times 0.4 = 0.064$ There is only one way of scoring 15 points.

(b)

x	0	5	10	15
$\text{P}(X = x)$	0.6	0.4×0.6	$0.4^2 \times 0.6$	
	0.6	0.24	0.096	0.064

Set out the distribution in a table

(c)

P(Jean scores more in round two than round one)
$= \text{P}(X = 0 \text{ then } X = 5, 10 \text{ or } 15)$
$+\text{P}(X = 5 \text{ then } X = 10 \text{ or } 15)$
$+\text{P}(X = 10 \text{ then } X = 15)$

There is only 1 way of scoring each value as the round ends if Jean misses.

$= 0.6 \times (0.24 + 0.096 + 0.064)$
$+0.24 \times (0.096 + 0.064)$
$+0.096 \times 0.064$

Consider the possible score for the first round in turn and the corresponding scores on the second round.

$= 0.284544$
$= 0.285$ (3 sf)

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Review Exercise
Exercise A, Question 1

Question:

In a factory, machines A, B and C are all producing metal rods of the same length. Machine A produces 35% of the rods, machine B produces 25% and the rest are produced by machine C. Of their production of rods, machines A, B and C produce 3%, 6% and 5% defective rods respectively.

a Draw a tree diagram to represent this information.

b Find the probability that a randomly selected rod is

i produced by machine A and is defective,

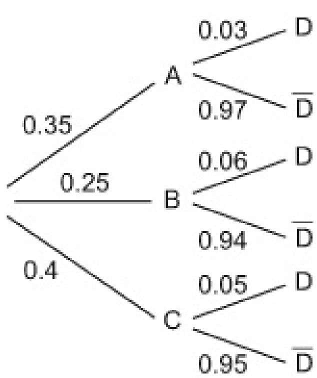
ii is defective.

c Given that a randomly selected rod is defective, find the probability that it was produced by machine C.

Solution:

Give your answers to at least 3 significant figures unless otherwise stated in the question.

a



Remember to label all the branches and put all the probabilities on.

bi

$$P(A \cap D) = 0.35 \times 0.03, = \underline{\mathbf{0.0105}} \quad \text{or} \quad \frac{21}{2000}$$

ii

$$P(D) = P(A \cap D) \text{ or } P(B \cap D) \text{ or } P(C \cap D)$$

$$= 0.0105 + (0.25 \times 0.06) + (0.4 \times 0.05)$$

If often helps to write down which combinations you want.

$$= \underline{\mathbf{0.0455}} \quad \text{or} \quad \frac{91}{2000}$$

c $$P(C \mid D) = \frac{P(C \cap D)}{P(D)} = \frac{0.4 \times 0.05}{0.0455}$$

$$= \underline{\mathbf{0.440}} \quad \text{or} \quad \frac{40}{91}$$

The words “given that” in the question tell you to use conditional probability.

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Review Exercise
Exercise A, Question 2

Question:

Summarised opposite are the distances, to the nearest mile, travelled to work by a random sample of 120 commuters.

Distance (to the nearest mile)	Number of commuters
0–9	10
10–19	19
20–29	43
30–39	25
40–49	8
50–59	6
60–69	5
70–79	3
80–89	1

For this distribution,

a describe its shape,

b use linear interpolation to estimate its median.

The mid-point of each class was represented by x and its corresponding frequency by f giving $\Sigma fx = 3550$ and $\Sigma fx^2 =$ 138 020

c Estimate the mean and standard deviation of this distribution.

One coefficient of skewness is given by

$$\frac{3(\text{mean} - \text{median})}{\text{standarddeviation}}$$

d Evaluate this coefficient for this distribution.

e State whether or not the value of your coefficient is consistent with your description in part **a**. Justify your answer.

f State, with a reason, whether you should use the mean or the median to represent the data in this distribution.

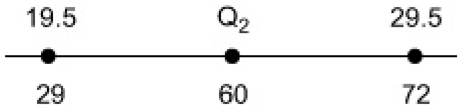
g State the circumstance under which it would not matter whether you used the mean or the median to represent a set of data.

Solution:

a Positive skew

b We do not need to do any rounding as the data is

Median : $\frac{120}{2} = 60^{\text{th}}$ term so in the $20 - 29$ class grouped.



Draw a diagram if it helps to get the fractions. Remember to use the class boundaries. Here they are 19.5 and 29.5 as there are gaps in the data. The end values on the bottom are the cumulative frequency for the previous class and this class.

$$\frac{Q_2 - 19.5}{29.5 - 19.5} = \frac{60 - 29}{72 - 29}$$

$$Q_2 = 19.5 + \frac{31}{43} \times 10$$

$$= \underline{\mathbf{26.7}}$$

c Mean $= \frac{3550}{120} = \underline{\mathbf{29.6}}$

Using $\frac{\sum fx}{\sum f}$

$$\sigma = \sqrt{\frac{138020}{120} - \left(\frac{3550}{120}\right)^2}$$

$$= 16.5829$$

$$= \underline{\mathbf{16.6}}$$

Use $\sqrt{\frac{\sum fx^2}{\sum f} - \left(\frac{\sum fx}{\sum f}\right)^2}$ and the accurate value for the mean

d $\frac{3(29.6 - 26.7)}{16.6} = 0.5199 \ldots$

$$= \underline{\mathbf{0.520}}$$

Use the accurate figures to do this calculation. Although you can write down the numbers to 3 s.f. If you use the rounded off figures you get 0.524 which may lose a mark.

e Yes it is as 0.520 > 0

f Use **<u>Median</u>**

Since the data is skewed

g If the data are **<u>symmetrical</u>** or **<u>skewness is zero</u>**

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Review Exercise
Exercise A, Question 3

Question:

A teacher recorded, to the nearest hour, the time spent watching television during a particular week by each child in a random sample. The times were summarised in a grouped frequency table and represented by a histogram.

One of the classes in the grouped frequency distribution was 20–29 and its associated frequency was 9. On the histogram the height of the rectangle representing that class was 3.6 cm and the width was 2 cm.

a Give a reason to support the use of a histogram to represent these data.

b Write down the underlying feature associated with each of the bars in a histogram.

c Show that on this histogram each child was represented by 0.8 cm^2.

The total area under the histogram was 24 cm^2.

d Find the total number of children in the group.

Solution:

a Time is a continuous variable

b Area is proportional to frequency

c Area of bar = 3.6 × 2 = 7.2

Remember it is the area of the bar which represents the frequency in a histogram.

9 children represented by 7.2 so: $\frac{7.2}{9} = 0.8$

or $\frac{7.2}{0.8} = 9$ or $0.8 \times 9 = 7.2$

1 child represented by 0.8

d Total $= \frac{24}{0.8} =$ **30**

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Review Exercise
Exercise A, Question 4

Question:

a Give two reasons to justify the use of statistical models.

It has been suggested that there are seven stages involved in creating a statistical model. They are summarised below, with stages 3, 4 and 7 missing.

Stage 1. The recognition of a real-world problem.

Stage 2. A statistical model is devised.

Stage 3.

Stage 4.

Stage 5. Comparisons are made against the devised model.

Stage 6. Statistical concepts are used to test how well the model describes the real-world problem.

Stage 7.

b Write down the missing stages.

Solution:

a Any 2 lines from

Used to simplify or represent a real world problem.

Cheaper or quicker (than producing the real situation) or more easily modified

To improve understanding of the real world problem

Used to predict outcomes from a real world problem (idea of predictions)

b (3) Model used to make predictions.

(4) Experimental data collected

(7) Model is refined. (Steps 2(or 3) to 5(or 6) are repeated)

You could put 3 and 4 the other way round.

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Review Exercise
Exercise A, Question 5

Question:

The following table summarises the distances, to the nearest km, that 134 examiners travelled to attend a meeting in London.

Distance (km)	Number of examiners
41–45	4
46–50	19
51–60	53
61–70	37
71–90	15
91–150	6

a Give a reason to justify the use of a histogram to represent these data.

b Calculate the frequency densities needed to draw a histogram for these data.

(*Do not draw the histogram.*)

c Use interpolation to estimate the median Q_2, the lower quartile Q_1, and the upper quartile Q_3 of these data.

The mid-point of each class is represented by x and the corresponding frequency by f. Calculations then give the following values

$\Sigma fx = 8379.5$ and $\Sigma fx^2 = 557\,489.75$

d Calculate an estimate of the mean and an estimate of the standard deviation for these data.

One coefficient of skewness is given by

$$\frac{Q_3 - 2Q_2 + Q_1}{Q_3 - Q_1}.$$

e Evaluate this coefficient and comment on the skewness of these data.

f Give another justification of your comment in **e**.

Solution:

a Distance is a continuous variable.

b $\text{F.D.} = \dfrac{\text{frequency}}{\text{class width}} \Rightarrow 0.8, 3.8, 5.3, 3.7, 0.75, 0.1$

c Q_2 is the $\dfrac{134}{2} = 67^{\text{th}}$ term so is in the class $51 - 60$

$$\frac{Q_2 - 50.5}{60.5 - 50.5} = \frac{67 - 23}{76 - 23}$$

$$\mathbf{Q_2} = \underline{\mathbf{58.8}}$$

Q_1 is the $\frac{134}{4} = 33.5^{\text{th}}$ term

$$\frac{Q_1 - 50.5}{60.5 - 50.5} = \frac{33.5 - 23}{76 - 23}$$

$$Q_1 = \underline{\mathbf{52.5}}$$

Q_3 is the $3 \times \frac{134}{4} = 100.5^{\text{th}}$ term

$$\frac{Q_3 - 60.5}{70.5 - 60.5} = \frac{100.5 - 76}{113 - 76}$$

$$Q_3 = \underline{\mathbf{67.1}}$$

d Mean $= \frac{8379.5}{134}$

$= \underline{\mathbf{62.5}}$

Using $\frac{\sum fx}{\sum f}$

Standard deviation $= \sqrt{\frac{557489.75}{134} - \left(\frac{8379.5}{134}\right)^2}$

$= \underline{\mathbf{15.8}}$

Using $\sqrt{\frac{\sum fx^2}{\sum f} - \left(\frac{\sum fx}{\sum f}\right)^2}$ and the accurate value for the mean. Using the rounded off version will give the wrong anser.

e $\frac{Q_3 - 2Q_2 + Q_1}{Q_3 - Q_1} = \frac{67.1 - 2 \times 58.8 + 52.5}{67.1 - 52.5}$

= $\underline{\mathbf{0.137}} \Rightarrow$ **POSITIVE SKEW**

f For positive skew Mean > median and 62.5 > 58.8

or

$Q_3 - Q_2$ (8.3) > $Q_2 - Q_1$ (6.3)

or

$\frac{3(62.5 - 58.8)}{15.8} = 0.703 > 0$

Therefore positive skew.

You should put the figures as well as the reason to show skew. Simply stating for example mean > median is not enough.

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Review Exercise
Exercise A, Question 6

Question:

Aeroplanes fly from City *A* to City *B*. Over a long period of time the number of minutes delay in take-off from City *A* was recorded. The minimum delay was five minutes and the maximum delay was 63 minutes. A quarter of all delays were at most 12 minutes, half were at most 17 minutes and 75% were at most 28 minutes. Only one of the delays was longer than 45 minutes.

An outlier is an observation that falls either 1.5 × (interquartile range) above the upper quartile or 1.5 × (interquartile range) below the lower quartile.

a On graph paper, draw a box plot to represent these data.

b Comment on the distribution of delays. Justify your answer.

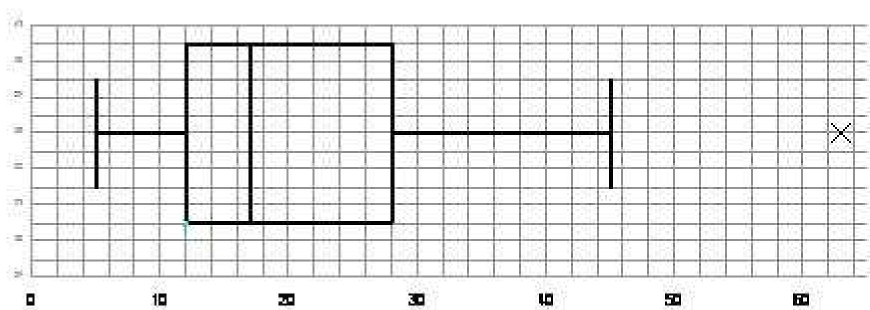
c Suggest how the distribution might be interpreted by a passenger who frequently flies from City *A* to City *B*.

Solution:

a $1.5\,(Q_3 - Q_1) = 1.5\,(28 - 12) = 24$

$Q_3 + 24 = 52 \Rightarrow 63$ is outlier

$Q_1 - 24 = -12 < 0$ no outliers



You must always check for outliers before drawing a box plot. Show your working and state what the outliers are if there are any. Or say no outliers if there are none.

The Outlier is marked by the cross.

The end of the whisker is put 45 since the question states that only one of the values is greater than 45. If we had not been told this the whisker would have gone at 52.

Always label the axis

b The distribution is positive skew since
$Q_2 - Q_1\ (5) < Q_3 - Q_2\ (11)$

Use figures to back up your reason

c Many delays are small so passengers should find these acceptable.

Always interpret in the context of the question

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Review Exercise
Exercise A, Question 7

Question:

In a school there are 148 students in Years 12 and 13 studying Science, Humanities or Arts subjects. Of these students, 89 wear glasses and the others do not. There are 30 Science students of whom 18 wear glasses. The corresponding figures for the Humanities students are 68 and 44 respectively.

A student is chosen at random.

Find the probability that this student

a is studying Arts subjects,

b does not wear glasses, given that the student is studying Arts subjects.

Amongst the Science students, 80% are right-handed. Corresponding percentages for Humanities and Arts students are 75% and 70% respectively.

A student is again chosen at random.

c Find the probability that this student is right-handed.

d Given that this student is right-handed, find the probability that the student is studying Science subjects.

Solution:

	Glasses	No glasses	Totals
Science	18	12	30
Arts	27	23	50
Humanities	44	24	68
Totals	89	59	148

Drawing a two way table enables you to see the data clearly. A Venn diagram would be too complicated as 5 circles would be needed

a $\text{P(Arts)} = \frac{50}{148} = \frac{25}{74} = \underline{\mathbf{0.338}}$

b $\text{P(no glasses / Arts)} = \frac{\text{P}(G' \cap A)}{\text{P}(A)} = \frac{23/148}{50/148} = \frac{23}{50} = \underline{\mathbf{0.46}}$

The words "given that" in the question tell you to use conditional probability.

c $\text{P(Right handed)} = \text{P(S}\cap\text{RH)} + \text{P}(A \cap \text{RH}) + \text{P(H}\cap\text{RH)}$

It often helps to write down which combinations you want.

$$= \frac{30}{148} \times 0.8 + \frac{50}{148} \times 0.7 + \frac{68}{148} \times 0.75$$

$$= \frac{55}{74} = \underline{\mathbf{0.743}}$$

d $\text{P(Science / right handed)} = \frac{\text{P(S}\cap\text{RH)}}{\text{P(RH)}} = \frac{\frac{30}{148} \times 0.8}{\frac{55}{74}}$

The words "given that" in the question tell you to use conditional probability.

$$= \frac{12}{55} = \underline{\mathbf{0.218}}$$

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Review Exercise
Exercise A, Question 8

Question:

Over a period of time, the number of people x leaving a hotel each morning was recorded. These data are summarised in the stem and leaf diagram below.

Number leaving	3|2 means 32	Totals
2	7 9 9	(3)
3	2 2 3 5 6	(5)
4	0 1 4 8 9	(5)
5	2 3 3 6 6 6 8	(7)
6	0 1 4 5	(4)
7	2 3	(2)
8	1	(1)

For these data,

a write down the mode,

b find the values of the three quartiles.

Given that $\Sigma x = 1335$ and $\Sigma x^2 = 71\,801$, find

c the mean and the standard deviation of these data.

One measure of skewness is found using

$\frac{3(\text{mean} - \text{mode})}{\text{standard deviation}}$.

d Evaluate this measure to show that these data are negatively skewed.

e Give two other reasons why these data are negatively skewed.

Solution:

a Mode is **56**

b Q_1 is the $\frac{27}{4} = 6.75$ which rounds up to 7^{th} term $Q_1 =$ **35**

Rounding needs to be used as the data is discrete

Q_2 is the $\frac{27}{2} = 13.5$ which rounds up to 14^{th} term $Q_2 =$ **52**

Q_3 is the $3 \times \frac{27}{4} = 20.25$ which rounds up to 21^{st} term $Q_3 =$ **60**

c Mean $= \frac{1335}{27} = 49.\dot{4}$ or $\underline{\mathbf{49\frac{4}{9}}}$

Using $\frac{\sum fx}{\sum f}$

Standard deviation $= \sqrt{\frac{71801}{27} - \left(\frac{1335}{27}\right)^2} =$ **14.6**

Using $\sqrt{\frac{\sum fx^2}{\sum f} - \left(\frac{\sum fx}{\sum f}\right)^2}$

d $\frac{3(49.4-56)}{14.4} = -\underline{\mathbf{1.343}}$

e For negative skew;

Mean<median<mode (49.4 < 52 < 56)

$Q_2 - Q_1$ (17) > $Q_3 - Q_2$ (8)

$\frac{3(49.4-52)}{14.6} = -0.534$

Any 2 of these would be acceptable but you should put the figures as well as the reason to show skew. Simply stating for example mean < median < mode is not enough.

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Review Exercise
Exercise A, Question 9

Question:

A bag contains nine blue balls and three red balls. A ball is selected at random from the bag and its colour is recorded. The ball is not replaced. A second ball is selected at random and its colour is recorded.

a Draw a tree diagram to represent the information.

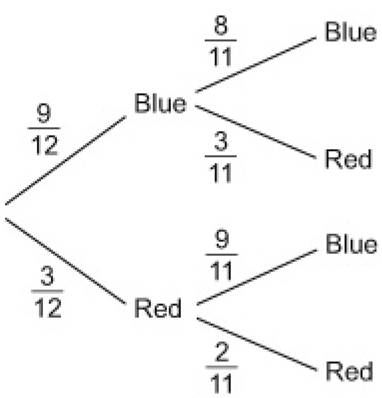
Find the probability that

b the second ball selected is red,

c both balls selected are red, given that the second ball selected is red.

Solution:

a



Remember to label all the branches and put all the probabilities on.

b P(Second ball is red) $= \frac{9}{12} \times \frac{3}{11} + \frac{3}{12} \times \frac{2}{11}$

$= \underline{\mathbf{\frac{1}{4}}}$

It often helps to write down which combinations you want.

c P(Both are red | Second ball is red) $= \frac{\frac{3}{12} \times \frac{2}{11}}{\frac{1}{4}} = \underline{\mathbf{\frac{2}{11}}}$

The words "given that" in the question tell you to use conditional probability.

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Review Exercise
Exercise A, Question 10

Question:

For the events A and B,

$P(A \cap B') = 0.32$, $P(A' \cap B) = 0.11$ and $P(A \cup B) = 0.65$.

a Draw a Venn diagram to illustrate the complete sample space for the events A and B.

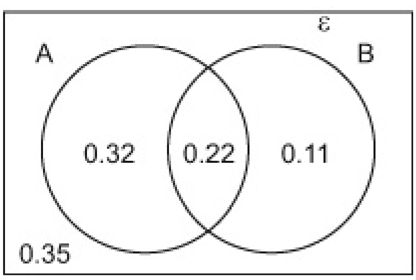
b Write down the value of $P(A)$ and the value of $P(B)$.

c Find $P(A|B')$.

d Determine whether or not A and B are independent.

Solution:

a $P(A \cup B) = P(A) + P(B) - P(A \cap B)$
$0.65 = 0.32 + 0.11 + P(A \cap B)$
$P(A \cap B) = 0.22$
$1 - 0.32 - 0.22 - 0.11 = 0.35$



When drawing a Venn diagram remember to draw a rectangle around the circles and add the probability 0.35.

b $P(A) = 0.32 + 0.22 = \underline{\mathbf{0.54}}$
$P(\boldsymbol{B}) = \underline{\mathbf{0.33}}$

c $P(A \mid B') = \frac{P(A \cap B')}{P(B')} = \frac{0.32}{0.67} = \underline{\frac{\mathbf{32}}{\mathbf{67}}}$

The words "given that" in the question tell you to use conditional probability.

d For independence $P(A \cap B) = P(A)P(B)$
For these data $0.22 \neq 0.54 \times 0.33 = 0.1782$
$\therefore$ NOT independent

Solutionbank S1

Edexcel AS and A Level Modular Mathematics

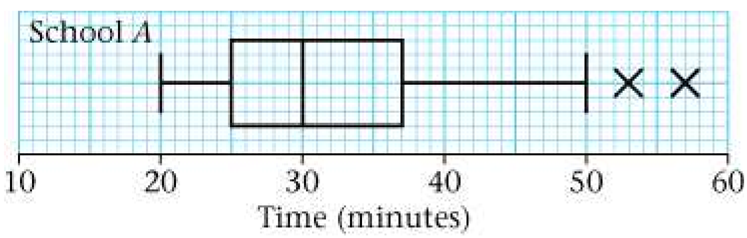
Review Exercise

Exercise A, Question 11

Question:

a Describe the main features and uses of a box plot.

Children from schools *A* and *B* took part in a fun run for charity. The times, to the nearest minute, taken by the children from school *A* are summarised in Figure 1.



bi Write down the time by which 75% of the children in school *A* had completed the run.

ii State the name given to this value.

c Explain what you understand by the two crosses (×) on Figure 1.

For school *B* the least time taken by any of the children was 25 minutes and the longest time was 55 minutes.

The three quartiles were 30, 37 and 50 respectively.

d On graph paper, draw a box plot to represent the data from school *B*.

e Compare and contrast these two box plots.

Solution:

a Indicates max/median/min/upper quartile/lower quartile

Indicates outliers

Indicates skewness

Allows comparisons

Indicates range / IQR/ spread

bi 37 minutes

ii Upper quartile /third quartile /75th percentile

c Outliers

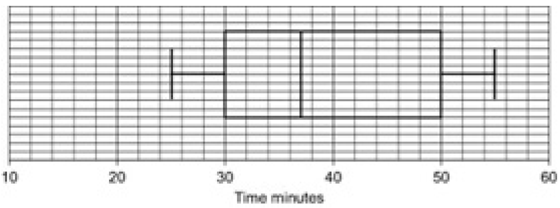
Observations that are different from the other observations and need to be treated with caution.

These two children probably walked/ took a lot longer.

d IQR = 20

$Q_1 - 1.5 \times 20 = 0$ therefore no outliers

$Q_3 + 1.5 \times 20 = 80$ therefore no outliers



You must always check for outliers before drawing a box plot. Show your working and state what the outliers are if there are any. Or say no outliers if there are none.

Use the same scale as the one in the question so they are easy to compare. As there are no outliers you must use the highest and lowest values for the end of the whiskers.

e Children from school A generally took less time.

50% of B ≤ 37 mins, 75% of A < 37 mins (similarly for 30)

Median of A < median of B

A has outliers, (B does not)

Both positive skew

IQR of A < IQR of B, range of A > range of B

You should always try to compare a measure of location, a measure of spread and skewness. Doing say the IQR and range will be treated as one comparison not two

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Review Exercise
Exercise A, Question 12

Question:

Sunita and Shelley talk to each other once a week on the telephone. Over many weeks they recorded, to the nearest minute, the number of minutes spent in conversation on each occasion. The following table summarises their results.

Time (to the nearest minute)	Number of conversations
5–9	2
10–14	9
15–19	20
20–24	13
25–29	8
30–34	3

Two of the conversations were chosen at random.

a Find the probability that both of them were longer than 24.5 minutes.

The mid-point of each class was represented by x and its corresponding frequency by f, giving $\Sigma fx = 1060$.

b Calculate an estimate of the mean time spent on their conversations.

During the following 25 weeks they monitored their weekly conversations and found that at the end of the 80 weeks their overall mean length of conversation was 21 minutes.

c Find the mean time spent in conversation during these 25 weeks.

d Comment on these two mean values.

Solution:

a $\text{P(both longer than } 24.5) = \frac{11}{55} \times \frac{10}{54}$

Note the numbers reduce in the fractions as they are different conversations

$= \mathbf{\frac{1}{27}}$ or $0.\dot{0}3\dot{7}$ or 0.037 to 3 sf

b Estimate of mean time spent

on their conversations is $= \frac{1060}{55}$

$= \mathbf{19\frac{3}{11}}$ or $19.\dot{2}\dot{7}$

c $\frac{1060 + \sum fy}{80} = 21$

Using $\frac{\text{total time}}{\text{total number}}$

$\sum fy = 620$

$\bar{y} = \frac{620}{25} = \underline{\mathbf{24.8}}$

d Increase in mean value.

Length of conversations increased considerably during the 25 weeks relative to the 55 weeks.	You need to put your comment in the context of the question

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Review Exercise
Exercise A, Question 13

Question:

A group of 100 people produced the following information relating to three attributes. The attributes were wearing glasses, being left-handed and having dark hair.

Glasses were worn by 36 people, 28 were left-handed and 36 had dark hair. There were 17 who wore glasses and were left-handed, 19 who wore glasses and had dark hair and 15 who were left-handed and had dark hair. Only 10 people wore glasses, were left-handed and had dark hair.

a Represent these data on a Venn diagram.

A person was selected at random from this group.

Find the probability that this person

b wore glasses but was not left-handed and did not have dark hair,

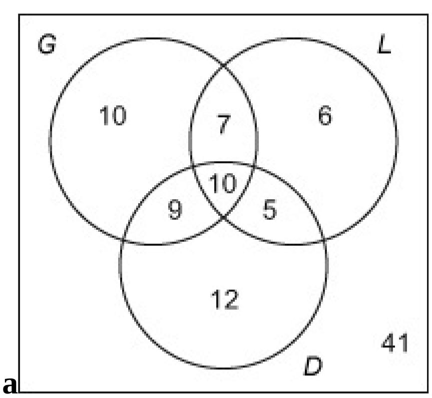
c did not wear glasses, was not left-handed and did not have dark hair,

d had only two of the attributes,

e wore glasses given they were left-handed and had dark hair.

Solution:

a



Start in the middle of the Venn diagram and work outwards. Remember the rectangle and those not in any of the circles. Your numbers should total 100.

b $P(G, \overline{L}, \overline{D}) = \frac{10}{100} = \underline{\mathbf{\frac{1}{10}}}$

c $P(\overline{G}, \overline{L}, \overline{D}) = \underline{\mathbf{\frac{41}{100}}}$

d $P(\text{Only two attributes}) = \frac{9+7+5}{100} = \underline{\mathbf{\frac{21}{100}}}$

e $P(G \mid L \,\&\, D) = \frac{P(G \mid L \,\&\, D)}{P(L \mid D)} = \frac{\frac{10}{100}}{\frac{15}{100}} = \frac{10}{15} = \underline{\mathbf{\frac{2}{3}}}$

The words “given that” in the question tell you to use conditional probability.

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Review Exercise
Exercise A, Question 14

Question:

A survey of the reading habits of some students revealed that, on a regular basis, 25% read quality newspapers, 45% read tabloid newspapers and 40% do not read newspapers at all.

a Find the proportion of students who read both quality and tabloid newspapers.

b Draw a Venn diagram to represent this information.

A student is selected at random. Given that this student reads newspapers on a regular basis,

c find the probability that this student only reads quality newspapers.

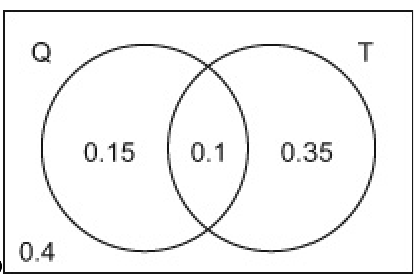
Solution:

a $\mathrm{P}(Q \cup T) = \mathrm{P}(Q) + \mathrm{P}(T) - \mathrm{P}(Q \cap T)$

$0.6 = 0.25 + 0.45 - \mathrm{P}(Q \cap T)$

$\mathrm{P}(Q \cap T) = \underline{\mathbf{0.1}}$

b



When drawing a Venn diagram remember to draw a rectangle around the circles and add the probability 0.4.

Remember it is the total in circle A = 0.25 and the total in circle B = 0.45.

c $\mathrm{P}(Q \cap T' \mid Q \cup T) = \frac{0.15}{0.6} = \frac{1}{4} = \underline{\mathbf{0.25}}$

The words “given that” in the question tell you to use conditional probability.

Solutionbank S1

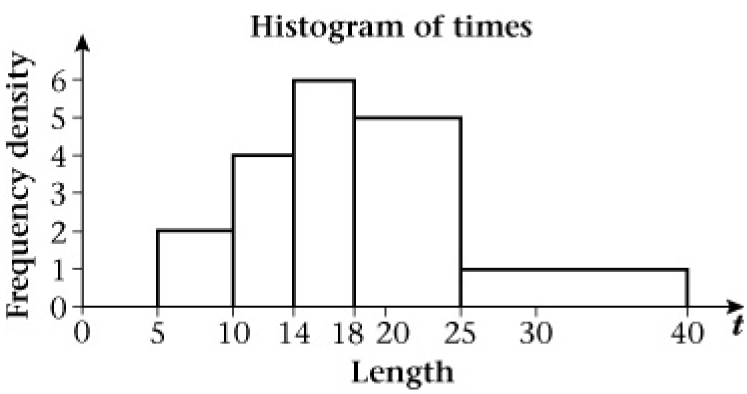
Edexcel AS and A Level Modular Mathematics

Review Exercise

Exercise A, Question 15

Question:

Figure 2 shows a histogram for the variable *t* which represents the time taken, in minutes, by a group of people to swim 500 m.



a Copy and complete the frequency table for *t*.

t	5–10	10–14	14–18	18–25	25–40
Frequency	10	16	24		

b Estimate the number of people who took longer than 20 minutes to swim 500 m.

c Find an estimate of the mean time taken.

d Find an estimate for the standard deviation of *t*.

e Find the median and quartiles for *t*.

One measure of skewness is found using

$$\frac{3(\text{mean} - \text{median})}{\text{standard deviation}}$$

f Evaluate this measure and describe the skewness of these data.

Solution:

a $18 - 25$ group, area $= 7 \times 5 = \underline{\mathbf{35}}$

$25 - 40$group, area $= 15 \times 1 = \underline{\mathbf{15}}$

Check the bars given to see whether the area equals the frequency or is proportional to the frequency. Here it equals the frequency.

b $(25 - 20) \times 5 + (40 - 25) \times 1 = \underline{\mathbf{40}}$

Calculatimg the area on the histogram for time > 20

c Mid points are 7.5, 12, 16, 21.5, 32.5

$\sum f = 100$

The word estimate is used in the question when the midpoints are used to calculate the mean and standard deviation.

$\sum ft = (10 \times 7.5) + (16 \times 12) + (24 \times 16) + (35 \times 21.5) + (15 \times 32.5)$
$= 1891$

$\frac{\sum ft}{\sum f} = \frac{1891}{100} = \underline{\mathbf{18.91}}$

d $\sum ft^2 = (10 \times 7.5^2) + (16 \times 12^2) + (24 \times 16^2) + (35 \times 21.5^2) + (15 \times 32.5^2) = 41033$

$\sigma = \sqrt{\frac{41033}{100} - \left(\frac{1891}{100}\right)^2} = \underline{\mathbf{7.26}}$

Using $\sqrt{\frac{\sum fx^2}{\sum f} - \left(\frac{\sum fx}{\sum f}\right)^2}$

e Q_2 is the $\frac{100}{2} = 50^{\text{th}}$ term

No Rounding is needed when calculating the quartiles etc as the data is grouped.

$Q_2 = \underline{\mathbf{18}}$

The 50th term is at the end of the 14 − 18 class so it is 18 as there are no gaps in the classes

Q_1 is the $\frac{100}{4} = 25^{\text{th}}$ term

$\frac{Q_1 - 10}{14 - 10} = \frac{25 - 10}{26 - 10}$

$Q1 = \underline{\mathbf{13.75}}$

Q_3 is the $3 \times \frac{100}{4} = 75^{\text{th}}$ term

$\frac{Q_3 - 18}{25 - 18} = \frac{75 - 50}{85 - 50}$

$Q_3 = \underline{\mathbf{23}}$

f $\frac{3(18.91 - 18)}{7.26} = \underline{\mathbf{0.376}}$ therefore it is positive skew

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Review Exercise
Exercise A, Question 16

Question:

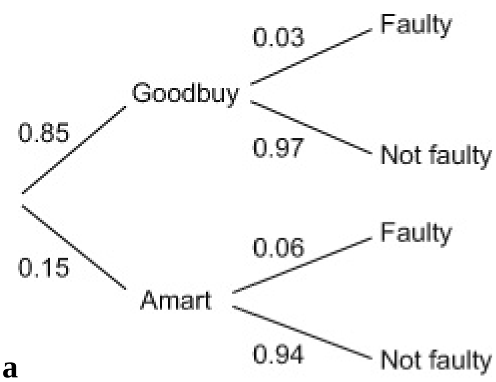
A company assembles drills using components from two sources. Goodbuy supplies 85% of the components and Amart supplies the rest. It is known that 3% of the components supplied by Goodbuy are faulty and 6% of those supplied by Amart are faulty.

a Represent this information on a tree diagram.

An assembled drill is selected at random.

b Find the probability that it is not faulty.

Solution:



a

Remember to label all the branches and put all the probabilities on.

b $P(\text{Not faulty}) = P(G \cap NF) + P(A \cap NF)$

$= (0.85 \times 0.97) + (0.15 \times 0.94)$
$= \underline{\mathbf{0.9655}}$

It often helps to write down which combinations you want.

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Review Exercise
Exercise A, Question 17

Question:

Data is coded using $y = \frac{x - 120}{5}$. The mean of the coded data is 24 and the standard deviation is 2.8. Find the mean and the standard deviation of the original data.

Solution:

$\frac{x - 120}{5} = 24$

$x = 240$ therefore mean = $\underline{\mathbf{240}}$

standard deviation $= 2.8 \times 5$

$= \underline{\mathbf{14}}$

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Review Exercise
Exercise A, Question 1

Question:

As part of a statistics project, Gill collected data relating to the length of time, to the nearest minute, spent by shoppers in a supermarket and the amount of money they spent. Her data for a random sample of 10 shoppers are summarised in the table below, where t represents time and £m the amount spent over £20.

t (minutes)	£m
15	−3
23	17
5	−19
16	4
30	12
6	−9
32	27
23	6
35	20
27	6

a Write down the actual amount spent by the shopper who was in the supermarket for 15 minutes.

b Calculate S_{tt}, S_{mm} and S_{tm}.

(You may use $\Sigma t^2 = 5478$, $\Sigma m^2 = 2101$, and $\Sigma tm = 2485$)

c Calculate the value of the product moment correlation coefficient between t and m.

d Write down the value of the product moment correlation coefficient between t and the actual amount spent. Give a reason to justify your value.

On another day Gill collected similar data. For these data the product moment correlation coefficient was 0.178

e Give an interpretation to both of these coefficients.

f Suggest a practical reason why these two values are so different.

Solution:

a $20 - 3 = \underline{\mathbf{£\,17}}$

b $\sum t = 212$ and $\sum m = 61$

$$S_{tm} = 2485 - \frac{61 \times 212}{10} = \underline{\mathbf{1191.8}}$$

$$S_{tt} = 5478 - \frac{212^2}{10} = \underline{\mathbf{983.6}}$$

$$S_{mm} = 2101 - \frac{61^2}{10} = \underline{\mathbf{1728.9}}$$

c

$$r = \frac{1191.8}{\sqrt{983.6 \times 1728.9}}$$
$$\underline{= \mathbf{0.914}}$$

d **<u>0.914</u>**

e.g. linear transformation, coding does not affect coefficient

e

0.914 suggests that the longer spent shopping the more money spent. (Idea more time, more spent)	Interpretation must be done in the context of the question
0.178 suggests that different amounts spent for same time.	

f e.g. might spend short time buying 1 expensive item <u>OR</u> might spend a long time checking for bargains, talking, buying lots of cheap items.

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Review Exercise
Exercise A, Question 2

Question:

The random variable X has probability function

$P(X = x) = \frac{(2x - 1)}{36}$ $x = 1, 2, 3, 4, 5, 6.$

a Construct a table giving the probability distribution of X.

Find

b $P(2 < X \le 5)$,

c the exact value of $E(X)$.

d Show that $Var(X) = 1.97$ to three significant figures.

e Find $Var(2 - 3X)$.

Solution:

a

x	1	2	3	4	5	6
$P(X = x)$	$\frac{1}{36}$	$\frac{3}{36}$	$\frac{5}{36}$	$\frac{7}{36}$	$\frac{9}{36}$	$\frac{11}{36}$
	0.0278, 0.0833, 0.139, 0.194, 0.25, 0.306					

b $P(3) + P(4) + P(5) = \underline{\mathbf{\frac{21}{36}} \text{ or } \mathbf{\frac{7}{12}} \text{ or } \mathbf{0.58\dot{3}}}$

c $E(X) = \frac{1}{36}[1 + 2 \times 3 + 3 \times 5 + 4 \times 7 + 5 \times 9 + 6 \times 11], = \underline{\mathbf{\frac{161}{36}} \text{ or } \mathbf{4.47\dot{2}} \text{ or } \mathbf{4\frac{17}{36}}}$

d $E(X^2) = \frac{1}{36}[1 + 2^2 \times 3 + 3^2 \times 5 + 4^2 \times 7 + 5^2 \times 9 + 6^2 \times 11],$ Using $\Sigma x^2 p$

$= \underline{\mathbf{\frac{791}{36}} \text{ or } \mathbf{21.97\dot{2}} \text{ or } \mathbf{21\frac{35}{36}} \text{ or awrt } \mathbf{21.97}}$

You must show all the steps when you are asked to show that $Var(X) = 1.97$

$Var(X) = \frac{791}{36} - \left(\frac{161}{36}\right)^2 = \underline{\mathbf{1.9714} \ldots}$

e $Var(2 - 3X) = 9 \times 1.97$ or $(-3)^2 \times 1.97 = \underline{\mathbf{17.73}}$ Using $Var(aX + b) = a^2 Var(X)$

more accurate: $\underline{17.74}$

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Review Exercise
Exercise A, Question 3

Question:

The measure of intelligence, IQ, of a group of students is assumed to be Normally distributed with mean 100 and standard deviation 15.

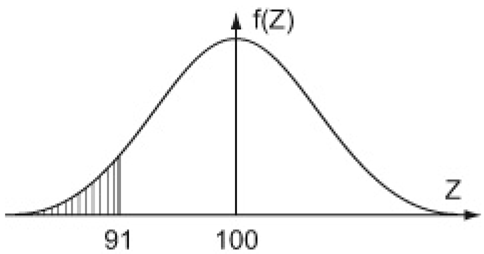
a Find the probability that a student selected at random has an IQ less than 91.

The probability that a randomly selected student as an IQ of at least 100 + k is 0.2090.

b Find, to the nearest integer, the value of k.

Solution:

a



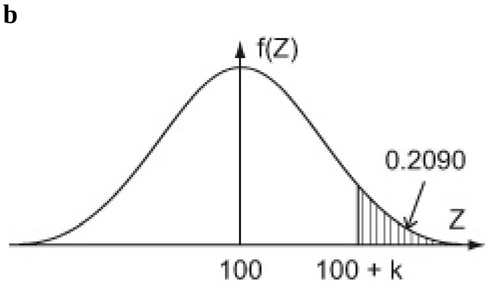
Drawing a diagram will help you to work out the correct area

$$\begin{aligned} P(X < 91) &= P\left(Z < \frac{91 - 100}{15}\right) \\ &= P(Z < -0.6) \\ &= 1 - 0.7257 \\ &= \underline{\mathbf{0.2743}} \end{aligned}$$

Using $z = \frac{x-\mu}{\sigma}$. As 91 is to the left of 100 your z value should be negative.

The tables give P(Z < 0.6) = P(Z > − 0.6) so you want 1 – this probability.

b



As 0.2090 is not in the table of percentage points you must work out the largest area

$1 - 0.2090 = 0.7910$

$P(X > 100 + k) = 0.2090$ <u>or</u> $P(X < 100 + k) = 0.791$

$$\frac{100 + k - 100}{15} = 0.81$$

$\underline{\boldsymbol{k} \mathbf{= 12}}$

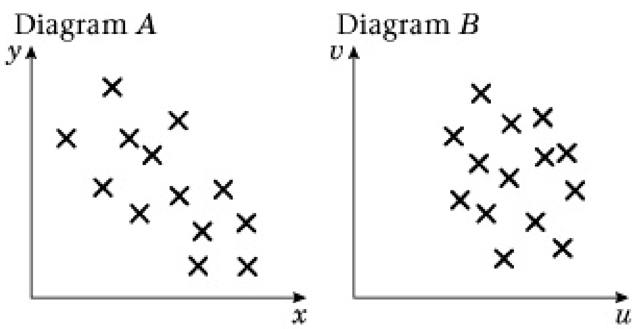
Use the first table or calculator to find the z value. It is positive as 100 + k is to the right of 100

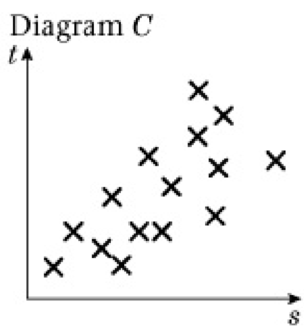
Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Review Exercise
Exercise A, Question 4

Question:

The scatter diagrams below were drawn by a student.





The student calculated the value of the product moment correlation coefficient for each of the sets of data.

The values were

0.68 −0.79 0.08

Write down, with a reason, which value corresponds to which scatter diagram.

Solution:

Diagram A : y & x : $\mathbf{r} = -\,\mathbf{0.79}$; as x increases, y decreases or most points lie in the 2nd and 4th quadrant

You must identify clearly which diagram each value goes with. And use the letters labelled on the axes.

Diagram B: v & u: $\mathbf{r} = \mathbf{0.08}$; no real pattern. Several values of v for one value of u or points lie in all four quadrants, randomly selected

Diagram C : t and s: $\mathbf{r} = \mathbf{0.68}$; As s increases, t increases or most points lie in the 1st and 3rd quadrants.

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Review Exercise
Exercise A, Question 5

Question:

A long distance lorry driver recorded the distance travelled, m miles, and the amount of fuel used, f litres, each day. Summarised below are data from the driver's records for a random sample of eight days.

The data are coded such that $x = m - 250$ and $y = f - 100$.

$\Sigma x = 130$	$\Sigma y = 48$
$\Sigma xy = 8880$	$S_{xx} = 20\,487.5$

a Find the equation of the regression line of y on x in the form $y = a + bx$.

b Hence find the equation of the regression line of f on m.

c Predict the amount of fuel used on a journey of 235 miles.

Solution:

a $S_{xx} = 20487.5$

$$S_{xy} = 8880 - \frac{130 \times 48}{8} = 8100$$

$$b = \frac{S_{xy}}{S_{xx}} = \frac{8100}{20487.5} = 0.395$$

$$a = \frac{48}{8} - (0.395363\ldots)\frac{130}{8} = -0.425$$

$\underline{\boldsymbol{y = -\ 0.425 + 0.395x}}$

b $f - 100 = -\ 0.4246\ \ldots.. + 0.395\ldots(m - 250)$ Just substitute in for x and y.

$\underline{\boldsymbol{f = 0.735 + 0.395\ m}}$ You must use the accurate values for a and b otherwise you get an incorrect answer of 0.825 instead of 0.735

c $m = 235 \Rightarrow \underline{\boldsymbol{f = 93.6}}$

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Review Exercise
Exercise A, Question 6

Question:

The random variable X has probability function

$$P(X = x) = \begin{cases} kx & x = 1, 2, 3, \\ k(x+1) & x = 4, 5 \end{cases}$$

where k is a constant.

a Find the value of k.

b Find the exact value of E(X).

c Show that, to three significant figures, Var(X) = 1.47.

d Find, to one decimal place, Var (4 − 3X).

Solution:

a

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

Draw a probability distribution table. Substitute 1, 2 and 3 into kx, and then 4 and 5 into $k(x+1)$ to work out the probabilities.

$$k + 2k + 3k + 5k + 6k = 1$$
$$17k = 1$$
$$k = \underline{\mathbf{\frac{1}{17}}}$$

The sum of the probabilities = 1

b $$E(X) = 1 \times \frac{1}{17} + 2 \times \frac{2}{17} + 3 \times \frac{3}{17} + 4 \times \frac{5}{17} + 5 \times \frac{6}{17}$$
$$= \frac{64}{17}$$
$$= \underline{\mathbf{3\frac{13}{17}}}$$

Question requires an exact answer therefore it is best to work in fractions

c $$E(X^2) = 1 \times \frac{1}{17} + 4 \times \frac{2}{17} + 9 \times \frac{3}{17} + 16 \times \frac{5}{17} + 25 \times \frac{6}{17}$$
$$= \mathbf{\frac{266}{17}}$$

Using $\Sigma x^2 p$

You must show all the steps when you are asked to show that Var(X) = 1.47.

$$\text{Var}(X) = \frac{266}{17} - \left(3\frac{13}{17}\right)^2$$
$$= 1.474 = \underline{\mathbf{1.47}}$$

d $$\text{Var}(4 - 3X) = 9\text{Var}(X)$$

Using Var (aX + b) = a^2Var(X)

$= 9 \times 1.47$
$= \underline{\mathbf{13.3}}$

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Review Exercise
Exercise A, Question 7

Question:

A scientist found that the time taken, M minutes, to carry out an experiment can be modelled by a normal random variable with mean 155 minutes and standard deviation 3.5 minutes.

Find

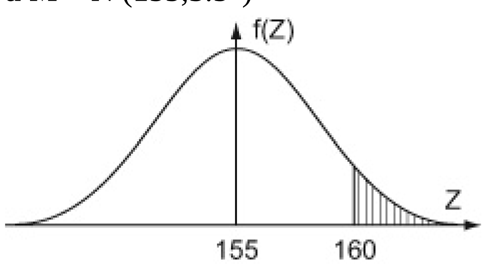
a $P(M > 160)$.

b $P(150 \leq M \leq 157)$.

c the value of m, to one decimal place, such that $P(M \leq m) = 0.30$.

Solution:

a $M \sim N(155, 3.5^2)$



Drawing a diagram will help you to work out the correct area

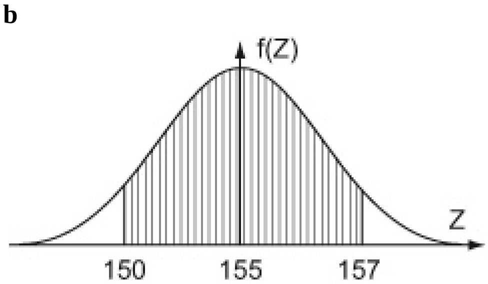
$$P(M > 160) = P\left(z > \frac{160-155}{3.5}\right)$$

Using $z = \frac{x-\mu}{\sigma}$. As 160 is to the right of 155 your z value should be positive

$$= P(z > 1.43)$$
$$= 1 - 0.9236$$
$$= \underline{0.7640} \qquad \underline{(0.0766 \text{ if calc used})}$$

The tables give $P(Z < 1.43)$ so you want 1 − this probability.

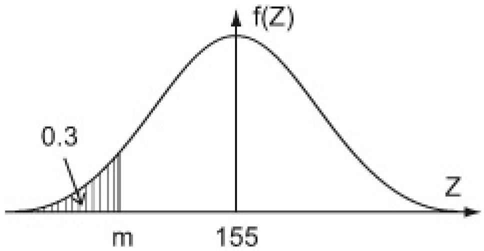
b



$$P(150 < M < 157) = P\left(\frac{150-155}{3.5} < z < \frac{157-155}{3.5}\right)$$
$$= P(-1.43 \leq z \leq 0.57)$$
$$= P(z \leq 0.57) - P(z \leq -1.43)$$
$$= 0.7157 - (1 - 0.9236)$$
$$= \underline{0.6393} \qquad \underline{(0.6396 \text{ if calc used})}$$

The tables give $P(Z > -1.43)$ so you want 1 − this probability.

c



$\text{P}(M \leq m) = 0.3 \Rightarrow \frac{m-155}{3.5} = -0.5244$

$\underline{\boldsymbol{m = 153.2}}$

Use the table of percentage points or calculator to find z. You must use at least the 4 decimal places given in the table. It is a negative value since *m* is to the left of 155

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Review Exercise
Exercise A, Question 8

Question:

The random variable X has probability distribution

x	1	2	3	4	5
$\mathbf{P}(X = x)$	0.1	p	0.20	q	0.30

a Given that E(X) = 3.5, write down two equations involving p and q.

Find

b the value of p and the value of q,

c Var (X),

d Var (3 − 2X).

Solution:

a $0.1 + p + 0.2 + q + 0.3 = 1$ $\qquad \sum p = 1$

$\mathbf{p + q = 0.4} \qquad \mathbf{(1)}$

$1 \times 0.1 + 2 \times p + 3 \times 0.2 + 4 \times q + 5 \times 0.3 = 3.5$ $\qquad \mathrm{E}(X) = \sum x\mathrm{P}(X = x) = 3.5$

$\mathbf{2p + 4q = 1.3} \qquad \mathbf{(2)}$

b Solving simultaneously

Multiplying **(1)** by **(2)**

$2p + 2q = 0.8$

$2q = 0.5$

q= 0.25

subst in to **(1)** $p + 0.25 = 0.4$

p = 0.15, q = 0.25

c $\mathrm{E}(X^2) = 1^2 \times 0.10 + 2^2 \times 0.15 + 3^2 \times 0.2 + 4^2 \times 0.25 + 5^2 \times 0.30 = \mathbf{14}$ $\qquad$ Using $\Sigma x^2 p$

$\mathrm{Var}(X) = 14 - 3.5^2 = \mathbf{1.75}$

d $\mathrm{Var}(3 - 2X) = 4\mathrm{Var}(X) = 4 \times 1.75 = \mathbf{7.00}$ $\qquad$ Using $\mathrm{Var}(aX + b) = a^2\,\mathrm{Var}(X)$

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Review Exercise
Exercise A, Question 9

Question:

A manufacturer stores drums of chemicals. During storage, evaporation takes place. A random sample of 10 drums was taken and the time in storage, x weeks, and the evaporation loss, y ml, are shown in the table below.

x	3	5	6	8	10	12	13	15	16	18
y	36	50	53	61	69	79	82	90	88	96

a On graph paper, draw a scatter diagram to represent these data.

b Give a reason to support fitting a regression model of the form $y = a + bx$ to these data.

c Find, to two decimal places, the value of a and the value of b.

(You may use $\Sigma x^2 = 1352$, $\Sigma y^2 = 53112$ and $\Sigma xy = 8354$.)

d Give an interpretation of the value of b.

e Using your model, predict the amount of evaporation that would take place after

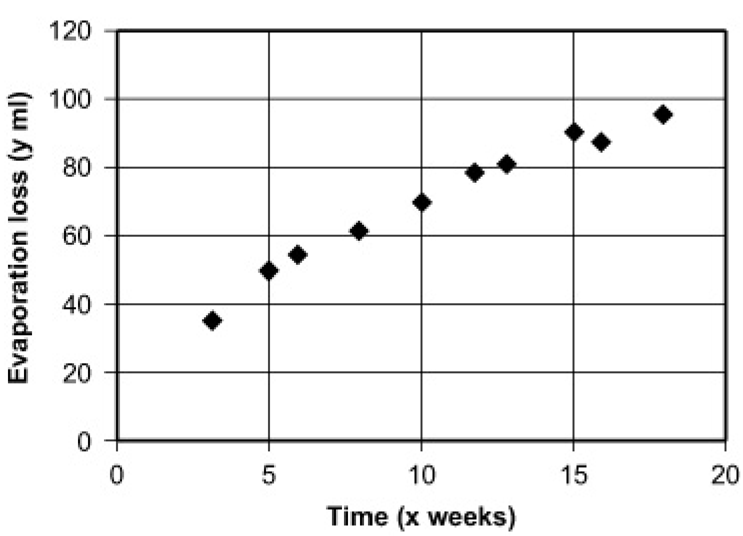
i 19 weeks,

ii 35 weeks.

f Comment, with a reason, on the reliability of each of your predictions.

Solution:

a



b Points lie close to a straight line

c $\sum x = 106, \sum y = 704, \sum xy = 8354$

$$S_{xy} = 8354 - \frac{106 \times 704}{10} = 891.6$$

$$S_{xx} = 1352 - \frac{106^2}{10} = 228.4$$

$$b = \frac{891.6}{228.4} = \underline{\mathbf{3.90}}$$

$$a = \frac{704}{10} - b\frac{106}{10} = \underline{\mathbf{29.02}} \text{ (2 dp required)}$$

d For every extra week in storage, another 3.90 ml of chemical evaporates

Interpretation must be done in the context of the question

e (i) $y = 29.0 + 3.90 \times 19 =$ **103.1** ml (ii) $y = 29.0 + 3.90 \times 35 =$ **165.5** ml

f (i) Close to range of x**, so reasonably reliable.**

(ii) **Well outside range of** x**, could be unreliable since no evidence that model will continue to hold.**

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Review Exercise
Exercise A, Question 10

Question:

a Write down two reasons for using statistical models.

b Give an example of a random variable that could be modelled by

i a normal distribution,

ii a discrete uniform distribution.

Solution:

a To simplify a real world problem

To improve understanding / describe / analyse a real world problem

Quicker and cheaper than using real thing

To predict possible future outcomes

Refine model / change parameters possible

b (i) e.g. height, weight (ii) score on a face after tossing a fair die

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Review Exercise
Exercise A, Question 11

Question:

The heights of a group of athletes are modelled by a normal distribution with mean 180 cm and standard deviation 5.2 cm. The weights of this group of athletes are modelled by a normal distribution with mean 85 kg and standard deviation 7.1 kg.

Find the probability that a randomly chosen athlete,

a is taller than 188 cm,

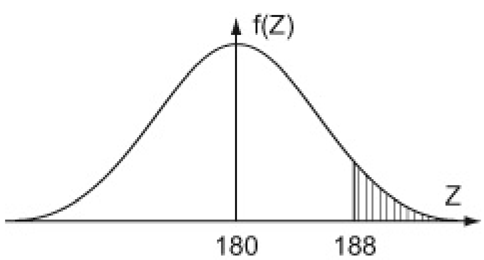
b weighs less than 97 kg.

c Assuming that for these athletes height and weight are independent, find the probability that a randomly chosen athlete is taller than 188 cm and weighs more than 97 kg.

d Comment on the assumption that height and weight are independent.

Solution:

a Let H be the random variable ~ height of athletes, so $H \sim N(180, 5.2^2)$



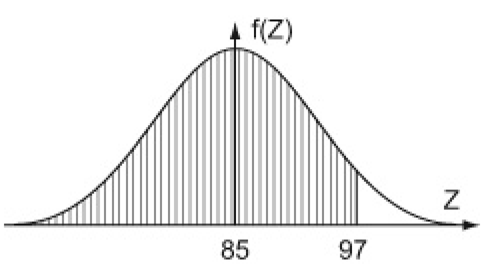
Drawing a diagram will help you to work out the correct area

$$\begin{aligned}P(H > 188) &= P\left(Z > \frac{188-180}{5.2}\right)\\ &= P(Z > 1.54)\\ &= 1 - 0.9382\\ &= \underline{\mathbf{0.0618}}\end{aligned}$$

Using $z = \frac{x-\mu}{\sigma}$. As 188 is to the right of 180 your z value should be positive

The tables give $P(Z < 1.54)$ so you want 1 − this probability.

b Let W be the random variable ~ weight of athletes, so $W \sim N(85, 7.1^2)$



$$\begin{aligned}P(W < 97) &= P\left(Z < \frac{97-85}{7.1}\right)\\ &= P(Z < 1.69)\\ &= \underline{\mathbf{0.9545}}\end{aligned}$$

Using $z = \frac{x-\mu}{\sigma}$. As 97 is to the right of 85 your z value should be positive

c

$$P(H > 188 \;\&\; W > 97) = 0.0618(1 - 0.9545) = 0.00281$$

$P(W > 97) = 1 - P(W < 97)$

d

Evidence suggests height and weight are positively correlated / linked

Use the context of the question when you are commenting

Assumption of independence is not sensible

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Review Exercise
Exercise A, Question 12

Question:

A metallurgist measured the length, l mm, of a copper rod at various temperatures, t°C, and recorded the following results.

t	l
20.4	2461.12
27.3	2461.41
32.1	2461.73
39.0	2461.88
42.9	2462.03
49.7	2462.37
58.3	2462.69
67.4	2463.05

The results were then coded such that $x = t$ and $y = l - 2460.00$.

a Calculate S_{xy} and S_{xx}.

(You may use $\Sigma x^2 = 15\ 965.01$ and $\Sigma xy = 757.467$)

b Find the equation of the regression line of y on x in the form $y = a + bx$.

c Estimate the length of the rod at 40°C.

d Find the equation of the regression line of l on t.

e Estimate the length of the rod at 90°C.

f Comment on the reliability of your estimate in **e**.

Solution:

a

$\Sigma x = \Sigma t = 337.1$, $\Sigma y = 16.28$

Σy is found by subtracting 2460 form all the l values to get y.

$$S_{xy} = 757.467 - \frac{337.1 \times 16.28}{8} = \underline{\mathbf{71.4685}}$$

$$S_{xx} = 15965.01 - \frac{337.1^2}{8} = \underline{\mathbf{1760.45875}}$$

b $b = \frac{71.4685}{1760.45875} = 0.04059652$

$a = \frac{16.28}{8} - 0.04059652 \times \frac{337.1}{8} = 0.324364$

$y = 0.324 + 0.0406x$

Remember to write the equation of the line at the end.

c $t = 40$ therefore $x = 40$,

$y = 0.324 + 0.0406 \times 40 = 1.948$,

$l = 2460 + 1.948 = \underline{\mathbf{2461.948}}$ mm

Calculate the value of y and then decode.

d

$l - 2460 = 0.324 + 0.0406t$

Just substitute in the coding for for x and y.

$l = 2460.324 + 0.0406t$

e At $t = 90$, $l = 2463.978$ mm

f 90°C outside range of data therefore unlikely to be reliable

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Review Exercise
Exercise A, Question 13

Question:

The random variable X has the discrete uniform distribution

$$P(X = x) = \frac{1}{5}, \quad x = 1, 2, 3, 4, 5.$$

a Write down the value of E(X) and show that Var(X) = 2.

Find

b E(3X − 2),

c Var(4 − 3X).

Solution:

a E(X) = 3;

$$\mathrm{Var}(X) = \frac{25-1}{12} = \mathbf{2}$$

This method is using the formulae for the uniform distribution

$$\text{Or } \mathrm{Var}(X) = 1 \times \frac{1}{5} + 2^2 \times \frac{1}{5} + 3^2 \times \frac{1}{5} + 4^2 \times \frac{1}{5} + 5^2 \times \frac{1}{5} - \mathrm{E}(X)^2$$

$$= 2$$

b

$$\mathrm{E}(3X - 2) = 3\mathrm{E}(X) - 2 = \mathbf{7}$$

Using E $(aX + b) = a\mathrm{E}(X) + b$

c

$$\mathrm{Var}(4 - 3\mathrm{X}) = 9\mathrm{Var}(X) = 18$$

Using Var $(aX + b) = a^2\mathrm{Var}(X)$

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Review Exercise
Exercise A, Question 14

Question:

From experience a high jumper knows that he can clear a height of at least 1.78 m once in five attempts. He also knows that he can clear a height of at least 1.65 m on seven out of 10 attempts.

Assuming that the heights the high jumper can reach follow a Normal distribution,

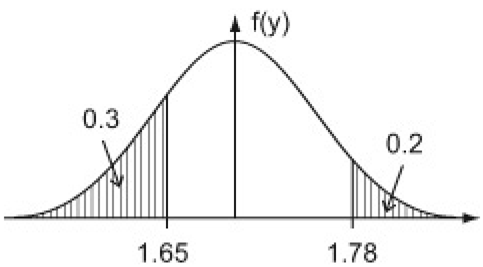
a draw a sketch to illustrate the above information,

b find, to three decimal places, the mean and the standard deviation of the heights the high jumper can reach,

c calculate the probability that he can jump at least 1.74 m.

Solution:

a



b

$$\begin{aligned} P(Z > a) &= 0.2 \\ a &= 0.8416 \\ P(Z < b) &= 0.3 \\ b &= -0.5244 \end{aligned}$$

Use the table of percentage points or calculator to find z. You must use at least the 4 decimal places given in the table. 0.5244 is negative since 1.65 is to the left of the centre. 0.8416 is positive as 1.78 is to the right of the centre.

Using $z = \frac{x-\mu}{\sigma}$.

$$\frac{1.78-\mu}{\sigma} = 0.8416 \Rightarrow 1.78 - \mu = 0.8416\sigma \quad \textbf{(1)}$$

$$\frac{1.65-\mu}{\sigma} = -0.5244 \Rightarrow 1.65 - \mu = -0.5244\sigma \quad \textbf{(2)}$$

Solving simultaneously **(1)–(2)**

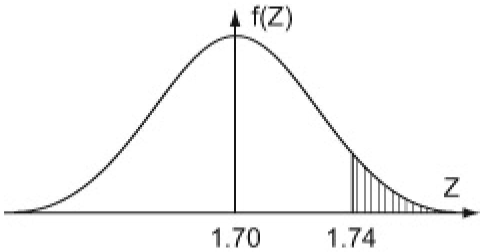
$0.13 = 1.366\sigma$

$\underline{\boldsymbol{\sigma = 0.095,}}$ metres

subst in $1.78 - \mu = 0.8416 \times 0.095$

$\underline{\boldsymbol{\mu = 1.70}}$ metres

c



P(height ≥ 1.74) $= P\left(z > \frac{1.74 - 1.70}{0.095}\right)$

$= P(z > 0.42)$

$= 1 - 0.6628$

$= 0.3372$ **(calc gives 0.3369)**

Using $z = \frac{x - \mu}{\sigma}$.

The tables give $P(Z < 0.42)$ so you want 1 − this probability.

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Review Exercise
Exercise A, Question 15

Question:

A young family were looking for a new three bedroom semi-detached house.

A local survey recorded the price x, in £1000s, and the distance y, in miles, from the station, of such houses. The following summary statistics were provided

$S_{xx} = 113573$, $S_{yy} = 8.657$,

$S_{xy} = -808.917$

a Use these values to calculate the product moment correlation coefficient.

b Give an interpretation of your answer to **a**.

Another family asked for the distances to be measured in km rather than miles.

c State the value of the product moment correlation coefficient in this case.

Solution:

a $r = \frac{S_{xy}}{\sqrt{S_{xx}S_{yy}}} = \frac{-808.917}{\sqrt{113573 \times 8.657}}$

$= -0.816 \ldots$

b

Houses are cheaper further away from the town centre or equivalent statement	Use the context of the question when you are asked to interpret

c

-0.816	To change miles to km you multiply by $\frac{8}{5}$. Coding makes no difference to the product moment correlation

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Review Exercise
Exercise A, Question 16

Question:

A student is investigating the relationship between the price (y pence) of 100 g of chocolate and the percentage (x%) of cocoa solids in the chocolate.

The following data are obtained

Chocolate brand	x (% cocoa)	y (pence)
A	10	35
B	20	55
C	30	40
D	35	100
E	40	60
F	50	90
G	60	110
H	70	130

(You may use: $\Sigma x = 315$, $\Sigma x^2 = 15\ 225$, $\Sigma y = 620$, $\Sigma y^2 = 56550$, $\Sigma xy = 28\ 750$)

a Draw a scatter diagram to represent these data.

b Show that $S_{xy} = 4337.5$ and find S_{xx}.

The student believes that a linear relationship of the form $y = a + bx$ could be used to describe these data.

c Use linear regression to find the value of a and the value of b, giving your answers to one decimal place.

d Draw the regression line on your diagram.

The student believes that one brand of chocolate is overpriced.

e Use the scatter diagram to

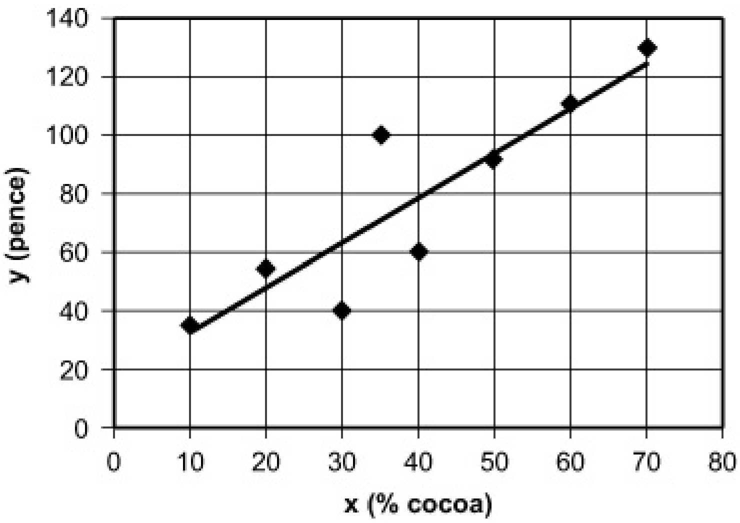
i state which brand is overpriced,

ii suggest a fair price for this brand.

Give reasons for both your answers.

Solution:

a, d



b $S_{xy} = 28750 - \frac{315 \times 620}{8} = 4337.5$

$S_{xx} = 15225 - \frac{315^2}{8} = 2821.875$

c $b = \frac{S_{xy}}{S_{xx}} = 1.537\ldots. = 1.5$

$a = \overline{y} - b\overline{x} = \frac{620}{8} - b\frac{315}{8} = 16.97\ldots = 17.0$

d on graph draw the line y = 17.0 + 1.5x

e i Brand D, since a long way above the line

ii Using the line: $y = 17 + 35 \times 1.5 = 69.5$ pence

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Review Exercise
Exercise A, Question 17

Question:

The random variable X has a normal distribution with mean 20 and standard deviation 4.

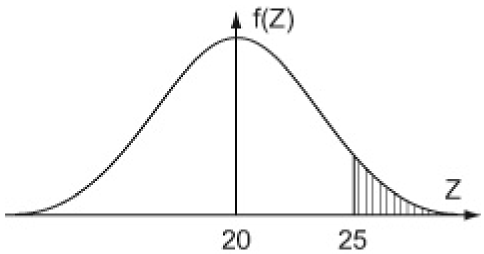
a Find $P(X > 25)$.

b Find the value of d such that

$P(20 < X < d) = 0.4641$.

Solution:

a



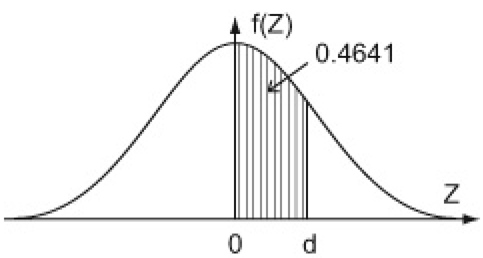
Drawing a diagram will help you to work out the correct area

$$\begin{aligned} P(X > 25) &= P\left(z > \frac{25-20}{4}\right) \\ &= P(Z > 1.25) \\ &= 1 - 0.8944 \\ &= \underline{\mathbf{0.1056}} \end{aligned}$$

Using $z = \frac{x-\mu}{\sigma}$. As 25 is to the right of 20 your z value should be positive.

The tables give $P(Z < 1.25)$ so you want 1 – this probability.

b



$P(X < 20) = 0.5$ so $P(X < d) = 0.5 + 0.4641 = 0.9641$

The area to the right of d is 0.0359. This is not in the table so you need to work out the area to the left of d.

$P(Z < z) = 0.9641,\ z = 1.80$

$$\frac{d-20}{4} = 1.80$$

$$\underline{\mathbf{d = 27.2}}$$

Use the first table or calculator to find the z value. It is positive as d is to the right of 0

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Review Exercise
Exercise A, Question 18

Question:

The random variable X has probability distribution

x	1	3	5	7	9
P(*X* = *x*)	0.2	p	0.2	q	0.15

a Given that E(X) = 4.5, write down two equations involving p and q.

Find

b the value of p and the value of q,

c $P(4 < X \leq 7)$.

Given that $E(X^2) = 27.4$, find

d Var(X),

e E(19 − 4X),

f Var(19 − 4X).

Solution:

a $0.2 + p + 0.2 + q + 0.15 = 1$ Sum of the probabilities = 1

$\underline{p + q = 0.45}$ **(1)**

$1 \times 0.2 + 3 \times p + 5 \times 0.2 + 7 \times q + 9 \times 0.15 = 4.5$ $E(X) = \Sigma x P(X = x) = 4.5$

$3p + 7q = 1.95$ **(2)**

b Solving the two equations simultaneously

$$\begin{aligned} 3p + 7q &= 1.95 \\ 3p + 3q &= 1.35 \quad \textbf{(1)} \times 3 \\ 4q &= 0.6 \end{aligned}$$

q= $\underline{0.15}$

subst $3p + 7 \times 0.15 = 1.95$

p= $\underline{0.3}$

c $P(4 < X \leq 7) = P(5) + P(7)$

$$\begin{aligned} &= 0.2 + q \\ &= \underline{\mathbf{0.35}} \end{aligned}$$

d $\text{Var}(X) = \text{E}(X^2) - [\text{E}(X)]^2 = 27.4 - 4.5^2$

$= \underline{\mathbf{7.15}}$

e $\text{E}(19 - 4X) = 19 - 4 \times 4.5$

$= \underline{\mathbf{1}}$

Using $\text{E}(aX + b) = a\text{E}(X) + b$

f $\text{Var}(19 - 4X) = 16\text{Var}(X)$

$= 16 \times 7.15$

$= \underline{\mathbf{114.4}}$

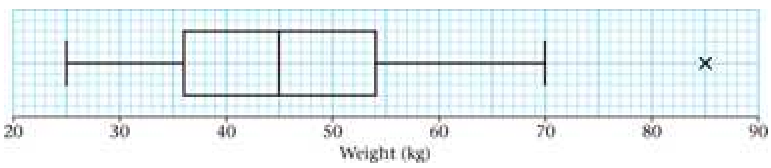
Using $\text{Var}(aX + b) = a^2\ \text{Var}(X)$

Solutionbank S1
Edexcel AS and A Level Modular Mathematics

Review Exercise
Exercise A, Question 19

Question:

The box plot in Figure 1 shows a summary of the weights of the luggage, in kg, for each musician in an orchestra on an overseas tour.



The airline's recommended weight limit for each musicians' luggage was 45 kg.

Given that none of the musician's luggage weighed exactly 45 kg,

a state the proportion of the musicians whose luggage was below the recommended weight limit.

A quarter of the musicians had to pay a charge for taking heavy luggage.

b State the smallest weight for which the charge was made.

c Explain what you understand by the × on the box plot in Figure 1, and suggest an instrument that the owner of this luggage might play.

d Describe the skewness of this distribution. Give a reason for your answer.

One musician in the orchestra suggests that the weights of the luggage, in kg, can be modelled by a normal distribution with quartiles as given in Figure 1.

e Find the standard deviation of this normal distribution. ***E***

Solution:

a $\frac{1}{2}$

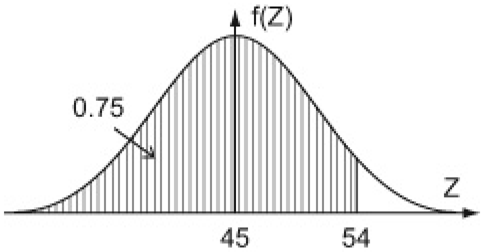
b <u>54</u> kg

c + is an 'outlier' or 'extreme' value

Any heavy musical instrument such as double-bass, drums.

d $Q_3 - Q_2 (9) = Q_2 - Q_1 (9)$ so symmetrical or no skew

e $P(W < 54) = 0.75$ or $P(W > 54) = 0.25$



$$\frac{54-45}{\sigma} = 0.67$$

σ = **13.4 (calc gives 13.3)**

Using the normal distribution table

$\Phi(0.67) = 0.7486$

$\Phi(0.68) = 0.7517$

0.7486 is closer to 0.75 therefore use $z = 0.67$.

or use the calculator to get 13.343…