1. The company *Seafield* requires contractors to record the number of hours they work each week. A random sample of 38 weeks is taken and the number of hours worked per week by contractor Kiana is summarised in the stem and leaf diagram below.

Stem	Leaf											
1	4	4	4	5	5	5	6	6	9	9	9	(11)
2	1	2	2	3	3	4	4	4	w	9		(10)
3	2	3	4	4	5	6	7	7	7	9		(10)
4	1	1	2	3								(4)
5	1	9										(2)
6	4											(1)

Key : 3|2 means 32

The quartiles for this distribution are summarised in the table below.

Q_1	Q_2	Q_3
x	26	y

(a) Find the values of w, x and y

(3)

Kiana is looking for outliers in the data. She decides to classify as outliers any observations greater than

$$Q_3 + 1.0 \times (Q_3 - Q_1)$$

(b) Showing your working clearly, identify any outliers that Kiana finds.

(2)

(c) Draw a box plot for these data in the space provided on the grid opposite.

(3)

(d) Use the formula

$$\text{skewness} = \frac{(Q_3 - Q_2) - (Q_2 - Q_1)}{(Q_3 - Q_1)}$$

to find the skewness of these data. Give your answer to 2 significant figures.

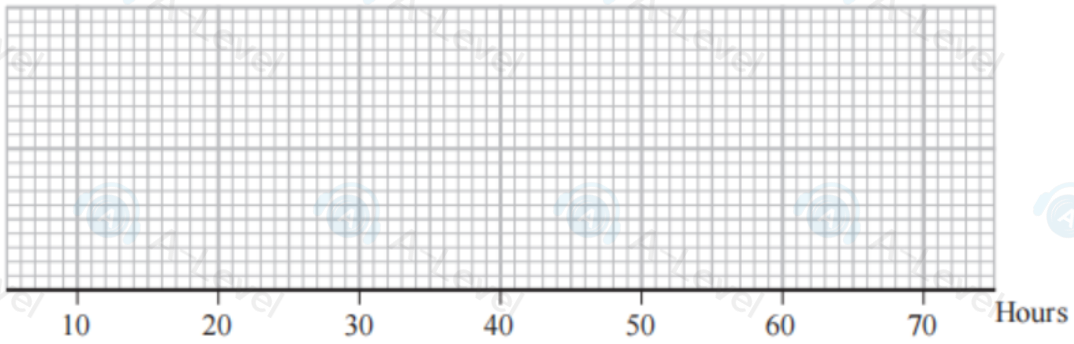
(2)

Kiana's new employer, *Landacre*, wishes to know the average number of hours per week she worked during her employment at *Seafield* to help calculate the cost of employing her.

(e) Explain why *Landacre* might prefer to know Kiana's mean, rather than median, number of hours worked per week.

(1)

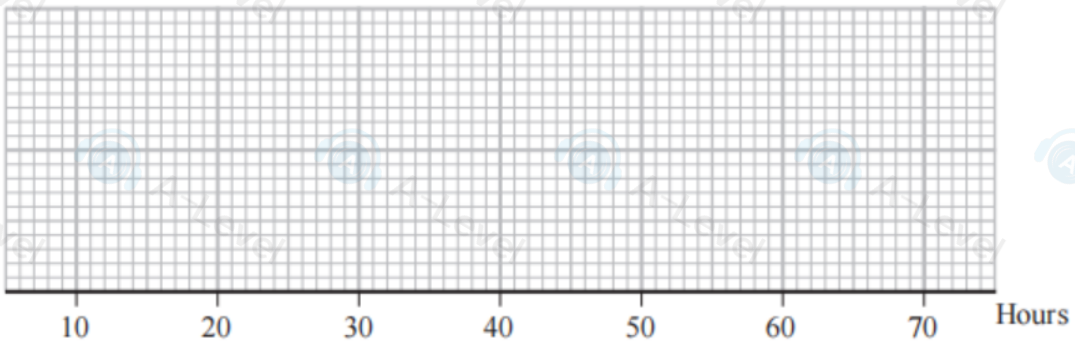
DO NOT WRITE IN THIS



Turn over for a spare grid if you need to redraw your box plot.

DO NOT WRITE IN THIS AREA

Only use this grid if you need to redraw your box plot.



(Total for Question 1 is 11 marks)

1. The discrete random variable X has the probability distribution given in the table below.

x	-2	1	3	4	6
$P(X = x)$	$\frac{1}{4}$	$\frac{1}{6}$	$\frac{1}{3}$	$\frac{1}{12}$	$\frac{1}{6}$

(a) Write down the value of $F(5)$

(1)

(b) Find $E(X)$

(2)

(c) Find $\text{Var}(X)$

(3)

The random variable $Y = 7 - 2X$

(d) Find

(i) $E(Y)$

(ii) $\text{Var}(Y)$

(iii) $P(Y > X)$

(6)

8. The random variable X is normally distributed with mean μ and variance 36

Given that

$$P(\mu - 2k < X < \mu + 2k) = 0.6$$

(a) find the value of k

(4)

The random variable Y is normally distributed with mean μ and standard deviation σ

Given that

$$2\mu = 3\sigma^2 \quad \text{and} \quad P\left(Y > \frac{3}{2}\mu\right) = 0.0668$$

(b) find the value of μ and the value of σ

(5)

DO NOT WRITE IN THIS AREA

Leave blank

2. A large company is analysing how much money it spends on paper in its offices each year. The number of employees in the office, x, and the amount spent on paper in a year, p (\$ hundreds), in each of 12 randomly selected offices were recorded.

The results are summarised in the following statistics.

$$\sum x = 93 \qquad S_{xx} = 148.25 \qquad \sum p = 273 \qquad \sum p^2 = 6602.72 \qquad \sum xp = 2347$$

(a) Show that $S_{xp} = 231.25$

(1)

(b) Find the product moment correlation coefficient for these data.

(3)

(c) Find the equation of the regression line of p on x in the form $p = a + bx$

(4)

(d) Give an interpretation of the gradient of your regression line.

(1)

The director of the company wants to reduce the amount spent on paper each year.

He wants each office to aim for a model of the form $p = \frac{4}{5}a + \frac{1}{2}bx$, where a and b are the values found in part (c).

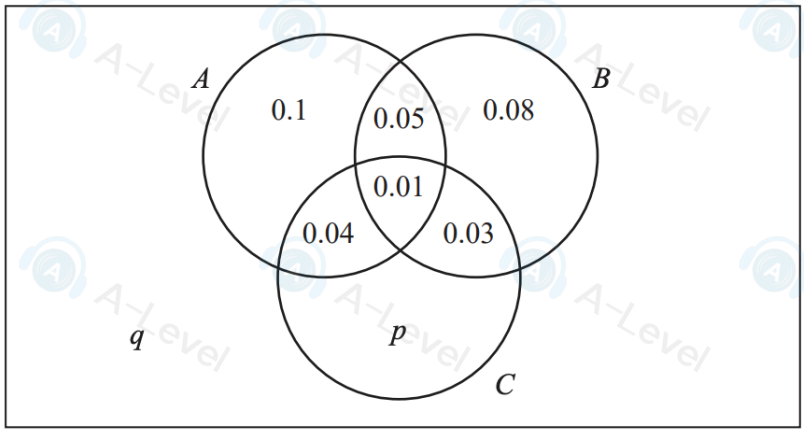
Using the data for the 93 employees from the 12 offices,

(e) estimate the percentage saving in the amount spent on paper each year by the company using the director's model.

(3)

4. Pieces of wood cladding are produced by a timber merchant. There are three types of fault, A, B and C, that can appear in each piece of wood cladding.

The Venn diagram shows the probabilities of a piece of wood cladding having the various types of fault.



A piece of wood cladding is chosen at random.

(a) Find the probability that the piece of wood cladding has more than one type of fault.

(1)

Fault types A and C occur independently.

(b) Find the probability that the piece of wood cladding has no faults.

(4)

Given that the piece of wood cladding has fault A,

(c) find the probability that it also has fault B but not fault C.

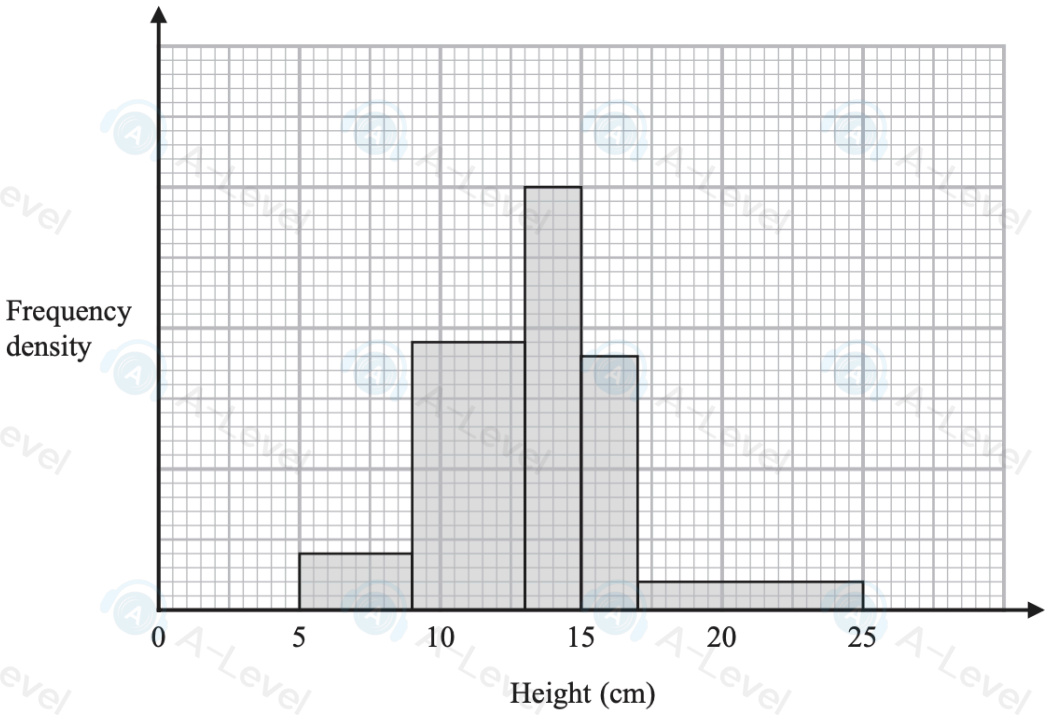
(2)

Two pieces of the wood cladding are selected at random.

(d) Find the probability that both have exactly 2 types of fault.

(3)

1. The histogram below shows the distribution of the heights, to the nearest cm, of 408 plants.



(a) Use the histogram to complete the following table.

Height (h cm)	$5 \leqslant h < 9$	$9 \leqslant h < 13$	$13 \leqslant h < 15$	$15 \leqslant h < 17$	$17 \leqslant h < 25$
Frequency	32	152	120		

(2)

(b) Use interpolation to estimate the median.

(2)

The mean height of these plants is 13.2 cm correct to one decimal place.

(c) Describe the skew of these data. Give a reason for your answer.

(1)

Two of these plants are chosen at random.

(d) Estimate the probability that both of their heights are between 8 cm and 14 cm

(3)

3. The table shows the price of a bottle of milk, m pence, and the price of a loaf of bread, b pence, for 8 different years.

m	29	29	35	39	41	43	44	46
b	75	83	91	121	120	126	119	126

(You may use $S_{bb} = 3083.875$ and $S_{mm} = 305.5$)

(a) Find the exact value of $\sum bm$

(1)

(b) Find S_{bm}

(3)

(c) Calculate the product moment correlation coefficient between b and m

(2)

(d) Interpret the value of the correlation coefficient.

(1)

A ninth year is added to the data set. In this year the price of the bottle of milk is 46 pence and the price of a loaf of bread is 175 pence.

(e) Without further calculation, state whether the value of the product moment correlation coefficient will increase, decrease or stay the same when all nine years are used. Give a reason for your answer.

(2)

5. The weights, X grams, of a particular variety of fruit are normally distributed with

$$X \sim \mathrm{N}(210, 25^2)$$

A fruit of this variety is selected at random.

(a) Show that the probability that the weight of this fruit is less than 240 grams is 0.8849

(2)

(b) Find the probability that the weight of this fruit is between 190 grams and 240 grams.

(2)

(c) Find the value of k such that $\mathrm{P}(210 - k < X < 210 + k) = 0.95$

(3)

A wholesaler buys large numbers of this variety of fruit and classifies the lightest 15% as small.

(d) Find the maximum weight of a fruit that is classified as small.
You must show your working clearly.

(3)

The weights, Y grams, of a second variety of this fruit are normally distributed with

$$Y \sim \mathrm{N}(\mu, \sigma^2)$$

Given that 5% of these fruit weigh less than 152 grams and 40% weigh more than 180 grams,

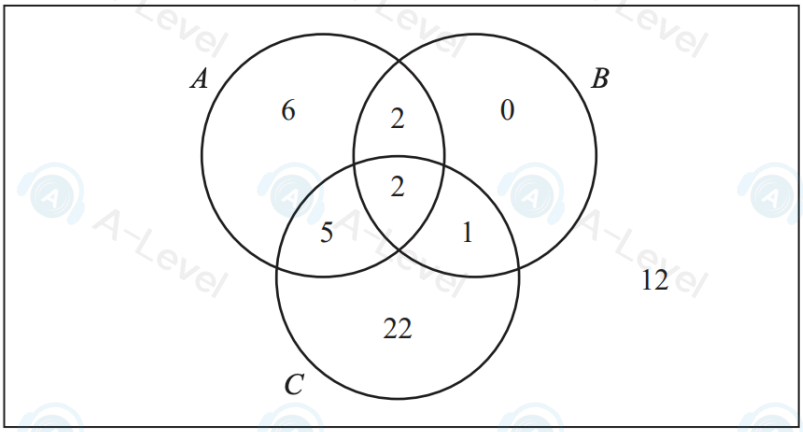
(e) calculate the mean and standard deviation of the weights of this variety of fruit.

(5)

2. An integer is selected at random from the integers 1 to 50 inclusive.

A is the event that the integer selected is prime.
B is the event that the integer selected ends in a 3
C is the event that the integer selected is greater than 20

The Venn diagram shows the number of integers in each region for the events A, B and C



(a) Describe in words the event $(A \cap B)$

(1)

(b) Write down the probability that the integer selected is prime.

(1)

(c) Find $P([A \cup B \cup C]')$

(1)

Given that the integer selected is greater than 20

(d) find the probability that it is prime.

(2)

Using your answers to (b) and (d),

(e) state, with a reason, whether or not the events A and C are statistically independent.

(2)

Given that the integer selected is greater than 20 and prime,

(f) find the probability that it ends in a 3

(2)