1. The histogram in Figure 1 shows the time, to the nearest minute, that a random sample of 100 motorists were delayed by roadworks on a stretch of motorway.

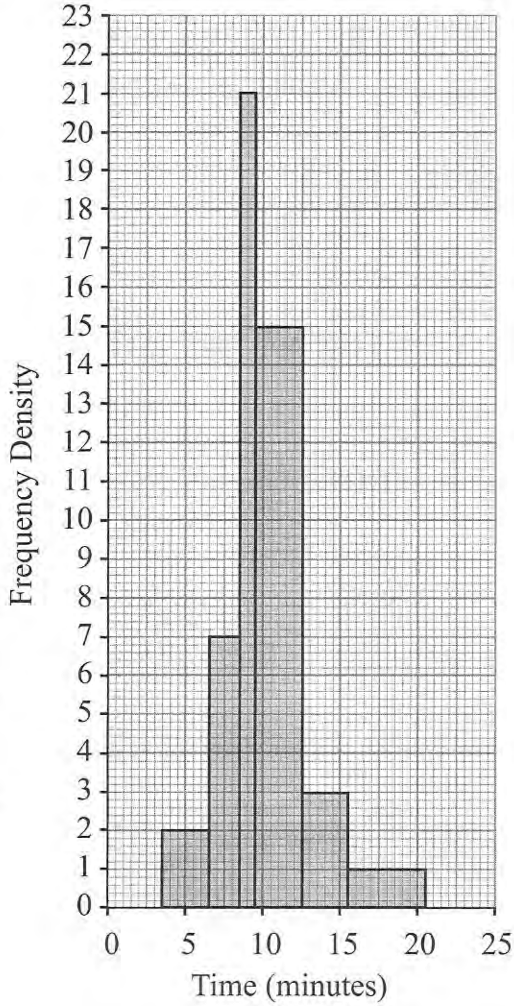


Figure 1

3

Area = 3×1 = 3

12·5 13·5

1

8·5-9·5 = 21
9·5-12·5 = 45 +
12·5-13·5 = 3
69

(a) Complete the table.

Delay (minutes)	cw	fd	Number of motorists
3·5 4 – 6 6·5	3	2	6
6·5 7 – 8 8·5	2	7	2×7 = 14
8·5 9 9·5	1	21	21
9·5 10 – 12 12·5	3	15	45
12·5 13 – 15 15·5	3	3	9
15·5 16 – 20 20·5	5	1	5×1 = 5

(2)

(b) Estimate the number of motorists who were delayed between 8.5 and 13.5 minutes by the roadworks.

69

(2)

2. (a) State in words the relationship between two events R and S when $P(R \cap S) = 0$ **(1)**

The events A and B are independent with $P(A) = \frac{1}{4}$ and $P(A \cup B) = \frac{2}{3}$

Find

(b) $P(B)$ **(4)**

(c) $P(A' \cap B)$ **(2)**

(d) $P(B' | A)$ **(2)**

a) Mutually exclusive

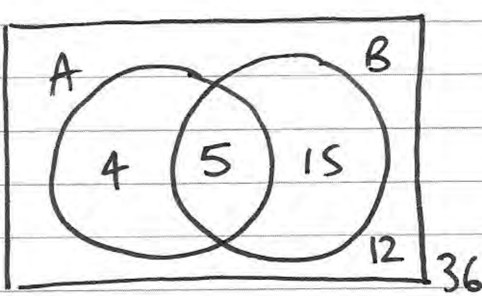
b) $P(A \cap B) = P(A) + P(B) - P(A \cup B)$

$\Rightarrow P(A) \times P(B) = P(A) + P(B) - P(A \cup B)$ Since independent.

$\Rightarrow \frac{1}{4} P(B) = \frac{1}{4} + P(B) - \frac{2}{3}$

$\Rightarrow -\frac{3}{4} P(B) = -\frac{5}{12} \Rightarrow P(B) = \frac{20}{36} = \frac{5}{9}$

c) $P(A \cap B) = \frac{1}{4} \times \frac{5}{9} = \frac{5}{36}$ $P(A) = \frac{1}{4} = \frac{9}{36}$ $P(B) = \frac{20}{36}$



$P(A' \cap B) = \frac{5}{12}$

d) $P(B' | A) = \frac{4}{9}$

3. The discrete random variable X can take only the values 2, 3, 4 or 6. For these values the probability distribution function is given by

x	2	3	4	6
$P(X = x)$	$\frac{5}{21}$	$\frac{2k}{21}$ $\frac{6}{21}$	$\frac{7}{21}$	$\frac{k}{21}$ $\frac{3}{21}$

where k is a positive integer.

(a) Show that $k = 3$ $\quad \Sigma P = 1 \quad \frac{5+2k+7+k}{21} = 1 \Rightarrow 3k+12=21$

$3k = 9$

$k = 3$ **(2)**

\#

Find

(b) F(3) $= P(2)+P(3) = \frac{5+6}{21} = \frac{11}{21}$

(1)

(c) E(X) $= 2\left(\frac{5}{21}\right)+3\left(\frac{6}{21}\right)+4\left(\frac{7}{21}\right)+6\left(\frac{3}{21}\right) = \frac{74}{21}$

(2)

(d) E(X^2) $= 2^2\left(\frac{5}{21}\right)+3^2\left(\frac{6}{21}\right)+4^2\left(\frac{7}{21}\right)+6^2\left(\frac{3}{21}\right) = 14$

(2)

(e) Var($7X - 5$)

(4)

e) $V(x) = E(x^2) - E(x)^2 = 14 - \left(\frac{74}{21}\right)^2 = \frac{698}{441}$

$V(7x-5) = 7^2 \times V(x) = 77\frac{5}{9}$

4. The marks, x, of 45 students randomly selected from those students who sat a mathematics examination are shown in the stem and leaf diagram below.

Mark		Totals	
3	6 9 9	(3)	3
4	0 1 2 2 3 4	(6)	9
4	5 6 6 6 8	(5)	14
5	0 2 3 3 4 4	(6)	20
5	5 5 6 7 7 9	(6)	26
6	0 0 0 0 1 3 4 4 4	(9)	35
6	5 5 6 7 8 9	(6)	41
7	1 2 3 3	(4)	45

Key	(3|6 means 36)

(a) Write down the modal mark of these students. mode $= \underline{60}$

(1)

(b) Find the values of the lower quartile, the median and the upper quartile.

(3)

For these students $\sum x = 2497$ and $\sum x^2 = 143\,369$

(c) Find the mean and the standard deviation of the marks of these students.

(3)

(d) Describe the skewness of the marks of these students, giving a reason for your answer.

(2)

The mean and standard deviation of the marks of all the students who sat the examination were 55 and 10 respectively. The examiners decided that the total mark of each student should be scaled by subtracting 5 marks and then reducing the mark by a further 10 %.

(e) Find the mean and standard deviation of the scaled marks of all the students.

(4)

b) $\frac{1}{4}n = 11.25 \Rightarrow LQ = x_{12} = 46$

$\frac{2}{4}n = 22.5 \Rightarrow \text{median} = x_{23} = 56$

$\frac{3}{4}n = 33.75 \Rightarrow UQ = x_{34} = 64$

c) $S_{xx} = \Sigma x^2 - \frac{(\Sigma x)^2}{n} = 4813.244..$

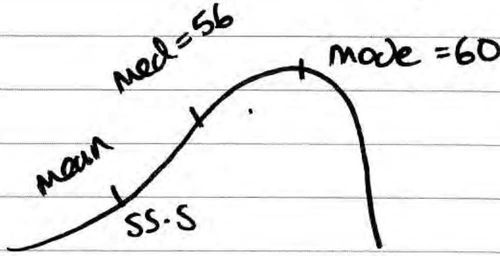
$\text{s.d.} = \sqrt{\frac{S_{xx}}{n}} = 10.3 \ (3sf)$

Question 4 continued

$\bar{x} = \dfrac{\Sigma x}{n} = 55.5$ (3sf)

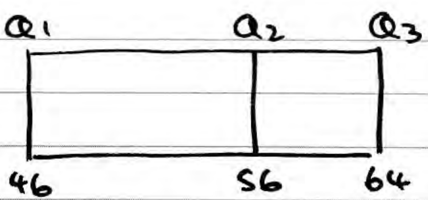
alt sd $\quad s.d = \sqrt{\dfrac{\Sigma x^2}{n} - \bar{x}^2} = 10.3$(3sf)

d)



mean < median < mode

alt



$\therefore$ negative skew

$$\underset{10}{Q_2 - Q_1} > \underset{8}{Q_3 - Q_2}$$

e) mark $\Rightarrow -5 \Rightarrow \overset{-10\%}{\times 0.9} \Rightarrow$ scaled mark

scaled mean $= (\bar{x} - 5) \times 0.9 = 45$

Scaled s.d. $= sd_x \times 0.9 = 9$

5. The age, t years, and weight, w grams, of each of 10 coins were recorded. These data are summarised below.

$$\sum t^2 = 2688 \quad \sum tw = 1760.62 \quad \sum t = 158 \quad \sum w = 111.75 \quad S_{ww} = 0.16$$

(a) Find S_{tt} and S_{tw} for these data.

(3)

(b) Calculate, to 3 significant figures, the product moment correlation coefficient between t and w.

(2)

(c) Find the equation of the regression line of w on t in the form $w = a + bt$

(4)

(d) State, with a reason, which variable is the explanatory variable.

(2)

(e) Using this model, estimate
  (i) the weight of a coin which is 5 years old,
  (ii) the effect of an increase of 4 years in age on the weight of a coin.

(2)

It was discovered that a coin in the original sample, which was 5 years old and weighed 20 grams, was a fake.

(f) State, without any further calculations, whether the exclusion of this coin would increase or decrease the value of the product moment correlation coefficient. Give a reason for your answer.

(2)

a) $Stt = \Sigma t^2 - \dfrac{(\Sigma t)^2}{n} = 191.6$

$Stw = \Sigma tw - \dfrac{(\Sigma t)(\Sigma w)}{n} = -5.03$

b) $PMCC = \dfrac{Stw}{\sqrt{Stt \times Sww}} = -0.908$ (3sf)

c) $\overset{\nearrow y}{w} = a + b\overset{\nearrow x}{t}$ $\quad b = \dfrac{Stw}{Stt} = -0.0262526...$

$a = \bar{w} - b\bar{t} = 11.5897...$

$w = 11.6 - 0.0263t$

d) Age is the explanatory variable as it is independent of the weight.

Weight will be the dependent variable as it will be determined from the age.

e) i) $w = 11.6 - 0.0263 \times 5 \simeq 11.5$ (3sf)

ii) Increase of 4 years will reduce the weight by $4 \times 0.0263...$
$\simeq$ reduction of 0.105g per 4 years

f) We would expect a 5yr old coin to weigh $\simeq 11.5$g, which would suggest the fake coin is an outlier. If excluded it should increase the correlation between age and weight.

$\therefore$ PMCC should get closer to -1, hence its value should decrease.

6. The following shows the results of a survey on the types of exercise taken by a group of 100 people.

65 run
48 swim
60 cycle
40 run and swim
30 swim and cycle
35 run and cycle
25 do all three

(a) Draw a Venn Diagram to represent these data. **(4)**

Find the probability that a randomly selected person from the survey

(b) takes none of these types of exercise, **(2)**

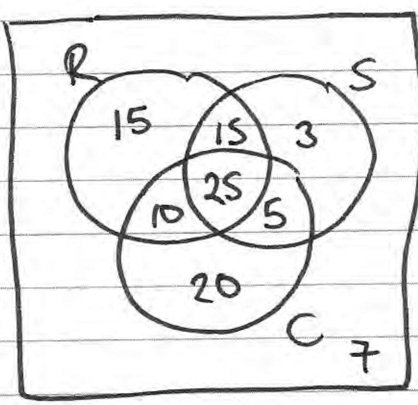
(c) swims but does not run, **(2)**

(d) takes at least two of these types of exercise. **(2)**

Jason is one of the above group.
Given that Jason runs,

(e) find the probability that he swims but does not cycle. **(3)**

a)



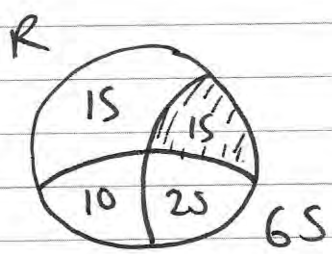
100

b) $\frac{7}{100}$

c) $\frac{8}{100}$ $\left(\frac{3+5}{100}\right)$

d) $\frac{55}{100}$ $\left(\frac{15+25+10+5}{100}\right)$

e) $\frac{15}{65}$



7. A manufacturer fills jars with coffee. The weight of coffee, W grams, in a jar can be modelled by a normal distribution with mean 232 grams and standard deviation 5 grams.

(a) Find $P(W < 224)$.

(3)

(b) Find the value of w such that $P(232 < W < w) = 0.20$

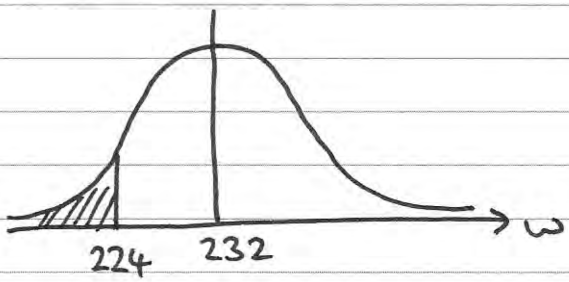
(4)

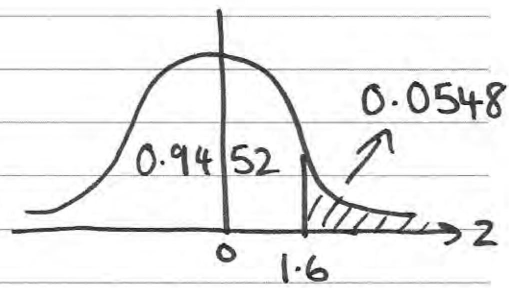
Two jars of coffee are selected at random.

(c) Find the probability that only one of the jars contains between 232 grams and w grams of coffee.

(3)

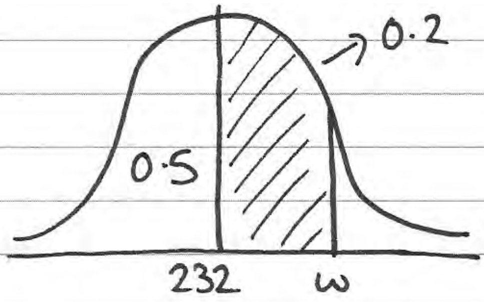
a) $\mu = 232 \quad \sigma = 5$



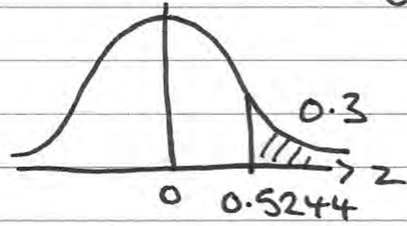


$z \to \frac{224-232}{5} = -1.6 \quad \Rightarrow P(W<224) = \Phi(-1.6) = 0.0548$

b)



$P(W<w) = 0.7 \Rightarrow P(W>w) = 0.3$



$z \to \quad \frac{w-232}{5} = 0.5244$

$\Rightarrow w = 234.6$

c) YN or NY $\Rightarrow 0.2 \times 0.8 \times 2 = 0.16 \times 2$

$= \underline{0.32}$