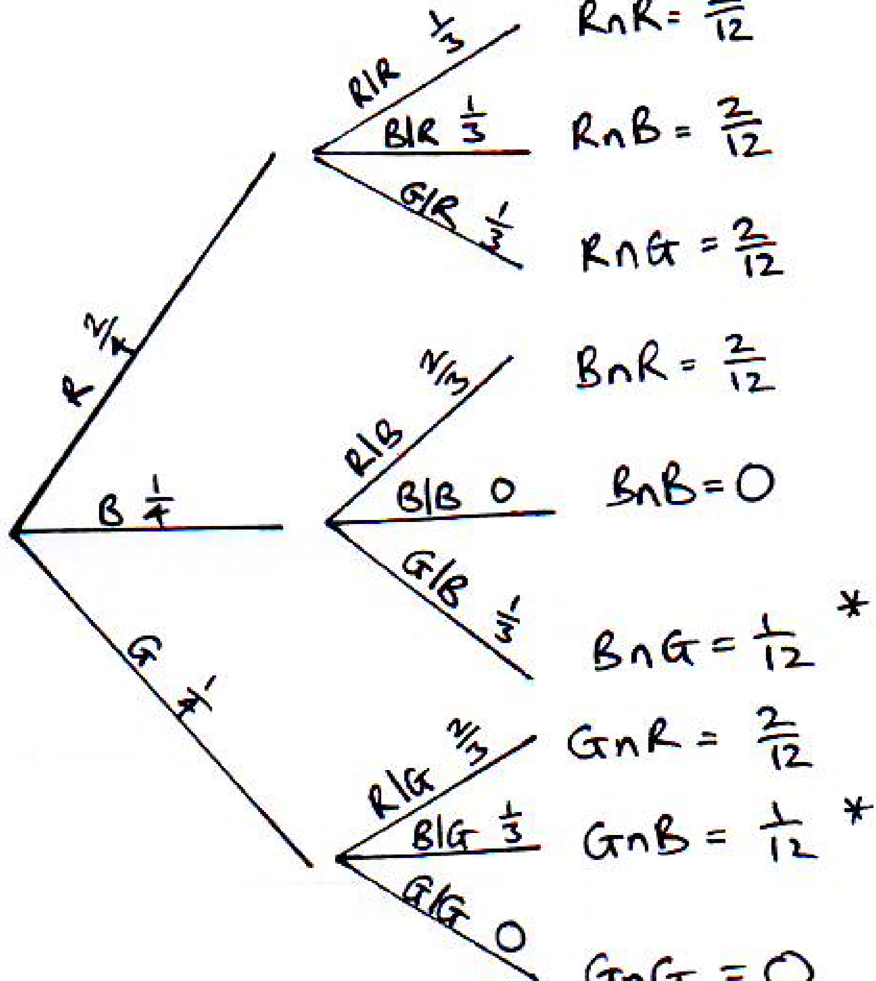
S1 JAN 10

1. A jar contains 2 red, 1 blue and 1 green bead. Two beads are drawn at random from the jar without replacement.

(a) In the space below, draw a tree diagram to illustrate all the possible outcomes and associated probabilities. State your probabilities clearly. **(3)**

(b) Find the probability that a blue bead and a green bead are drawn from the jar. **(2)**



b) $\frac{1}{12} + \frac{1}{12} = \frac{2}{12} = \underline{\frac{1}{6}}$

2. The 19 employees of a company take an aptitude test. The scores out of 40 are illustrated in the stem and leaf diagram below.

	2|6 means a score of 26	
0	7	(1)
1	88	(2)
2	4468	(4)
3	2333459	(7)
4	00000	(5)

Find

(a) the median score, $\frac{1}{2}n = 9.5 \quad Q_2 = x_{10} = \underline{33}$ **(1)**

(b) the interquartile range. $\frac{1}{4}n = 4.75 \quad Q_1 = x_5 = 24$ **(3)**

$\frac{3}{4}n = 14.25 \quad Q_3 = x_{15} = 40$

$IQR = 16$

The company director decides that any employees whose scores are so low that they are outliers will undergo retraining.

An outlier is an observation whose value is less than the lower quartile minus 1.0 times the interquartile range.

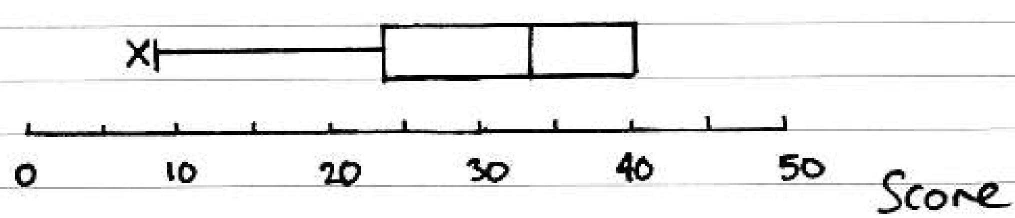
(c) Explain why there is only one employee who will undergo retraining. **(2)**

(d) On the graph paper on page 5, draw a box plot to illustrate the employees' scores. **(3)**

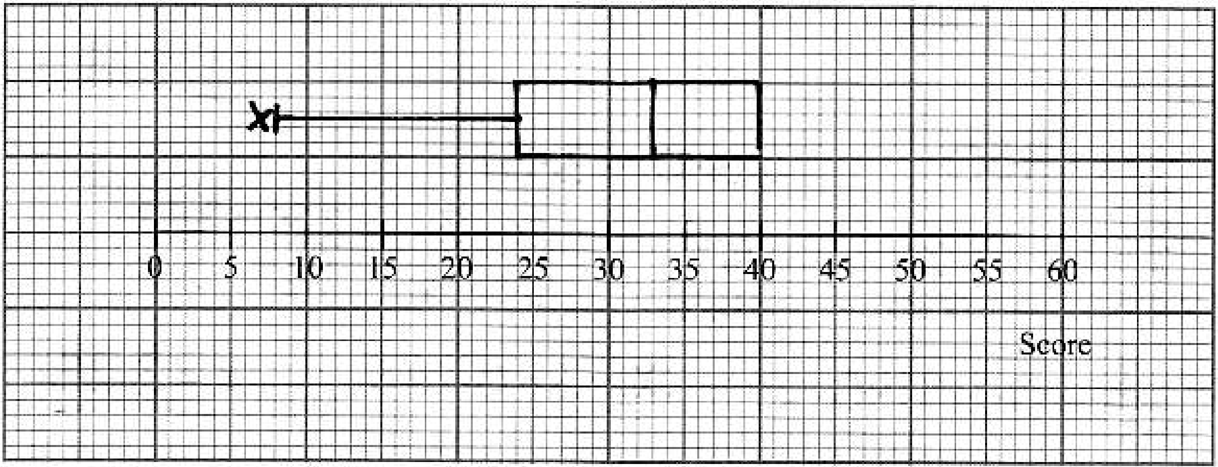
(c) $LQ = 24 \quad 24 - 1 \times 16 = 8$

only 1 employee scored less than 8.

d) $UQ + 1 \times IQR = 40 + 16 = 56$, no upper outliers.



(should be on graph paper)



Q2

(Total 9 marks)

3. The birth weights, in kg, of 1500 babies are summarised in the table below.

Weight (kg)	Midpoint, $x$kg	Frequency, f		fx	fx^2
			0		
0.0 – 1.0	0.50	1	1	0.5	
1.0 – 2.0	1.50	6	7	9	
2.0 – 2.5	2.25	60	67	135	
2.5 – 3.0	2.75	280	347		
3.0 – 3.5	3.25	820	1167		
3.5 – 4.0	3.75	320	1487		
4.0 – 5.0	4.50	10	1497		
5.0 – 6.0	5.50	3	1500		

[You may use $\sum fx = 4841$ and $\sum fx^2 = 15\,889.5$] ← no need!

(a) Write down the missing midpoints in the table above.
(2)

(b) Calculate an estimate of the mean birth weight. $\frac{\Sigma fx}{n} = \frac{4841}{1500} = \underline{3.23}$
(2)

(c) Calculate an estimate of the standard deviation of the birth weight.
(3)

(d) Use interpolation to estimate the median birth weight.
(2)

(e) Describe the skewness of the distribution. Give a reason for your answer.
(2)

(c) $\text{Var} = \frac{\Sigma fx^2}{n} - \bar{x}^2 = 0.1773\ldots \quad sd = \sqrt{\text{Var}} = \underline{0.421}$

(d) $\frac{1}{2}n = 750 \quad (3.0 - 3.5)$

3	x	3.5
347	750	1167

$\frac{x-3}{3.5-3} = \frac{750-347}{1167-347}$

$\frac{x-3}{0.5} = 0.49146\ldots \quad x - 3 = 0.245\ldots \quad \underline{x = 3.25}$

e) $\frac{3(\text{mean} - \text{median})}{\text{standard dev.}} = \frac{3(3.23-3.25)}{0.421} = -0.14$ Slight negative skew

4. There are 180 students at a college following a general course in computing. Students on this course can choose to take up to three extra options.

112 take systems support, S
70 take developing software, D
81 take networking, N
35 take developing software and systems support,
28 take networking and developing software,
40 take systems support and networking,
4 take all three extra options.

(a) In the space below, draw a Venn diagram to represent this information. **(5)**

A student from the course is chosen at random.

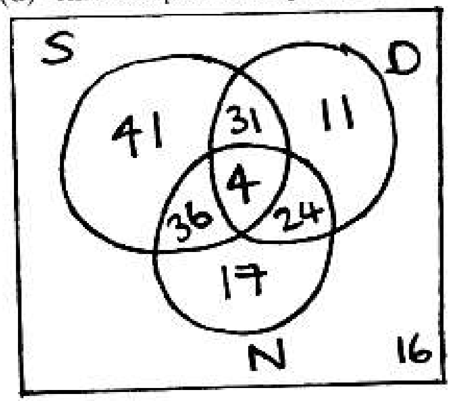
Find the probability that this student takes

(b) none of the three extra options, $= \frac{16}{180} = \frac{4}{45}$ **(1)**

(c) networking only. $= \frac{17}{180}$ **(1)**

Students who want to become technicians take systems support and networking. Given that a randomly chosen student wants to become a technician,

(d) find the probability that this student takes all three extra options. **(2)**



180

(d)

$= \frac{4}{40} = \frac{1}{10}$

5. The probability function of a discrete random variable X is given by

$$p(x) = kx^2 \qquad x = 1, 2, 3$$

where k is a positive constant.

x	1	2	3
$p(x)$	k	$4k$	$9k$

(a) Show that $k = \frac{1}{14}$ **(2)**

$\Sigma P(x) = 1 \Rightarrow 14k = 1$

$k = \frac{1}{14}$

Find

(b) $P(X \geqslant 2)$ **(2)**

(c) $E(X)$ **(2)**

(d) $Var(1 - X)$ **(4)**

x	1	2	3
P	$\frac{1}{14}$	$\frac{4}{14}$	$\frac{9}{14}$

$E(x) = \frac{1}{14} + \frac{8}{14} + \frac{27}{14} = \frac{36}{14}$

$E(x^2) = \frac{1}{14} + \frac{16}{14} + \frac{81}{14} = \frac{98}{14}$

b) $P(x \geqslant 2) = \frac{4}{14} + \frac{9}{14} = \frac{13}{14}$

c) $E(x) = \frac{36}{14}$

d) $V(x) = E(x^2) - E(x)^2 = \frac{98}{14} - \left(\frac{36}{14}\right)^2 = \frac{19}{49}$

$V(1-x) = (-1)^2 V(x) = \frac{19}{49}$

6. The blood pressures, p mmHg, and the ages, t years, of 7 hospital patients are shown in the table below.

$n=7$

Patient	A	B	C	D	E	F	G
t	42	74	48	35	56	26	60
p	98	130	120	88	182	80	135

$$[\sum t = 341, \sum p = 833, \sum t^2 = 18\,181, \sum p^2 = 106\,397, \sum tp = 42\,948]$$

(a) Find S_{pp}, S_{tp} and S_{tt} for these data. **(4)**

(b) Calculate the product moment correlation coefficient for these data. **(3)**

(c) Interpret the correlation coefficient. **(1)**

(d) On the graph paper on page 17, draw the scatter diagram of blood pressure against age for these 7 patients. **(2)**

$x \to t \quad y \to p$ $\qquad p = a + bt$

(e) Find the equation of the regression line of p on t. **(4)**

$b = \frac{S_{tp}}{S_{tt}} = 1.51$

$a = \bar{p} - b\bar{t}$

(f) Plot your regression line on your scatter diagram. **(2)**

$a = \frac{833}{7} - 1.51 \times \left(\frac{341}{7}\right)$

(g) Use your regression line to estimate the blood pressure of a 40 year old patient. **(2)**

$a = 45.5$

a) $S_{pp} = \sum p^2 - \frac{(\sum p^2)}{n} = \underline{7270}$

$S_{tp} = \underline{2369}$

$S_{tt} = \underline{1569.428571}\ldots$

b) $PMCC = \frac{2369}{\sqrt{7270 \times 1569.4..}} = 0.701$

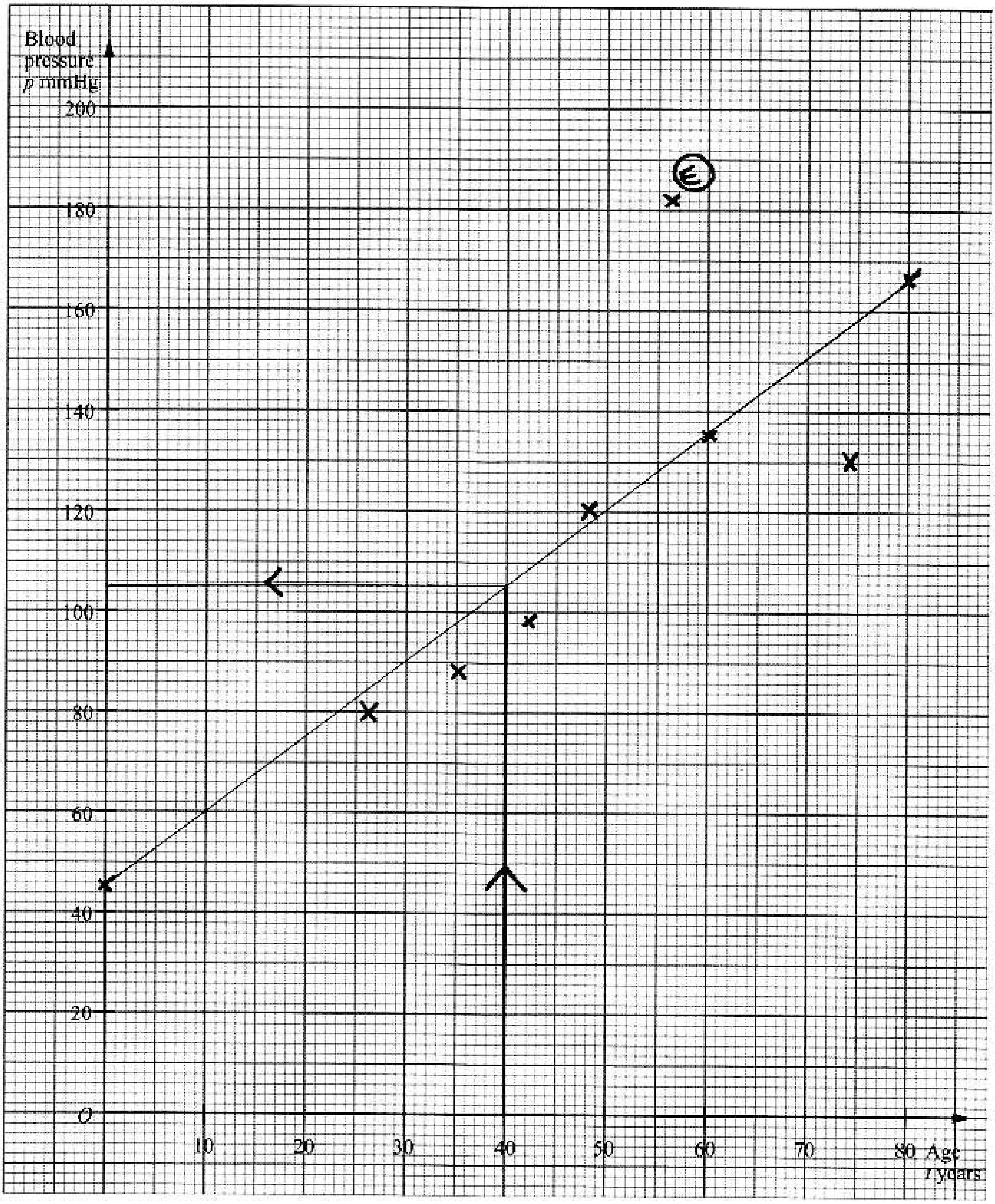
c) Some evidence to suggest positive correlation, the older the patient the higher their blood pressure.

Question 6 continued

6e) $p = 45.5 + 1.51t$

f) $t = 0 \quad p = 45.5$

$t = 80 \quad p = 166.3$



6f) $t = 40 \quad p = 45.5 + 1.51 \times 40 = 105.9$

7. The heights of a population of women are normally distributed with mean μ cm and standard deviation σ cm. It is known that 30% of the women are taller than 172 cm and 5% are shorter than 154 cm.

(a) Sketch a diagram to show the distribution of heights represented by this information.

(3)

(b) Show that $\mu = 154 + 1.6449\sigma$.

(3)

(c) Obtain a second equation and hence find the value of μ and the value of σ.

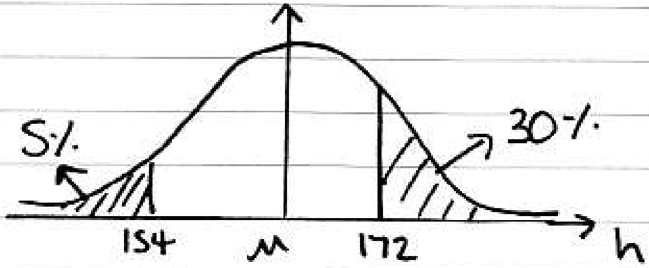
(4)

A woman is chosen at random from the population.

(d) Find the probability that she is taller than 160 cm.

(3)

a)



$h \sim N(\mu, \sigma^2)$

b) $P(h < 154) = 0.05$

$P\left(Z < \frac{154-\mu}{\sigma}\right) = 0.05 \Rightarrow P\left(Z > \frac{\mu-154}{\sigma}\right) = 0.05$

% points table $\Rightarrow$

$\Rightarrow \frac{\mu-154}{\sigma} = 1.6449 \Rightarrow \mu - 154 = 1.6449\sigma$

$\mu = 154 + 16449\sigma$

c) $P(h > 172) = 0.30$

$P\left(Z > \frac{172-\mu}{\sigma}\right) = 0.30 \Rightarrow \frac{172-\mu}{\sigma} = 0.5244$

$172 - \mu = 0.5244\sigma$

$\mu - 154 = 1.6449\sigma$ +

$18 = 2.1693\sigma$

$\sigma = 8.298$

$\mu = 154 + 1.6449(8.298)$

$\mu = \underline{167.6\text{ cm}}$

(d) $P(h > 160)$

$P\left(Z > \frac{160-167.6}{8.298}\right) = P(Z > -0.9218)$

$= P(Z < 0.9218)$

$= \Phi(0.9218) = \underline{0.8218}$