Answer **all** questions.

1 At a certain small restaurant, the waiting time is defined as the time between sitting down at a table and a waiter first arriving at the table. This waiting time is dependent upon the number of other customers already seated in the restaurant.

Alex is a customer who visited the restaurant on 10 separate days. The table shows, for each of these days, the number, x, of customers already seated and his waiting time, y minutes.

x	9	3	4	10	8	12	7	11	2	6
y	11	6	5	11	9	13	9	12	4	7

(a) Calculate the equation of the least squares regression line of y on x in the form $y = a + bx$. *(4 marks)*

(b) Give an interpretation, in context, for each of your values of a and b. *(2 marks)*

(c) Use your regression equation to estimate Alex's waiting time when the number of customers already seated in the restaurant is:

(i) 5;

(ii) 25. *(2 marks)*

(d) Comment on the likely reliability of **each** of your estimates in part (c), given that, for the regression line calculated in part (a), the values of the 10 residuals lie between +1.1 minutes and −1.1 minutes. *(3 marks)*

2 Xavier, Yuri and Zara attend a sports centre for their judo club's practice sessions. The probabilities of them arriving late are, independently, 0.3, 0.4 and 0.2 respectively.

(a) Calculate the probability that for a particular practice session:

(i) all three arrive late; *(1 mark)*

(ii) none of the three arrives late; *(2 marks)*

(iii) exactly one of the three arrives late. *(3 marks)*

(b) Zara's friend, Wei, also attends the club's practice sessions. The probability that Wei arrives late is 0.9 when Zara arrives late, and is 0.25 when Zara does not arrive late.

Calculate the probability that for a particular practice session:

(i) both Zara and Wei arrive late; *(2 marks)*

(ii) either Zara or Wei, but not both, arrives late. *(3 marks)*

3 When an alarm is raised at a market town's fire station, the fire engine cannot leave until at least five fire-fighters arrive at the station. The call-out time, X minutes, is the time between an alarm being raised and the fire engine leaving the station.

The value of X was recorded on a random sample of 50 occasions. The results are summarised below, where $\bar{x}$ denotes the sample mean.

$$\sum x = 286.5 \qquad \sum (x - \bar{x})^2 = 45.16$$

(a) Find values for the mean and standard deviation of this sample of 50 call-out times. *(2 marks)*

(b) Hence construct a 99% confidence interval for the mean call-out time. *(4 marks)*

(c) The fire and rescue service claims that the station's mean call-out time is less than 5 minutes, whereas a parish councillor suggests that it is more than $6\frac{1}{2}$ minutes.

Comment on **each** of these claims. *(2 marks)*

4 [**Figure 1**, printed on the insert, is provided for use in this question.]

The table shows the times, in seconds, taken by a random sample of 10 boys from a junior swimming club to swim 50 metres freestyle and 50 metres backstroke.

Boy	**A**	**B**	**C**	**D**	**E**	**F**	**G**	**H**	**I**	**J**
Freestyle (x seconds)	30.2	32.8	25.1	31.8	31.2	35.6	32.4	38.0	36.1	34.1
Backstroke (y seconds)	33.5	35.4	37.4	27.2	34.7	38.2	37.7	41.4	42.3	38.4

(a) On **Figure 1**, complete the scatter diagram for these data. *(2 marks)*

(b) Hence:

(i) give **two** distinct comments on what your scatter diagram reveals; *(2 marks)*

(ii) state, **without calculation**, which of the following 3 values is most likely to be the value of the product moment correlation coefficient for the data in your scatter diagram.

0.912 0.088 0.462 *(1 mark)*

(c) In the sample of 10 boys, one boy is a junior-champion freestyle swimmer and one boy is a junior-champion backstroke swimmer.

Identify the **two** most likely boys. *(2 marks)*

(d) **Removing the data for the two boys whom you identified in part (c):**

(i) calculate the value of the product moment correlation coefficient for the remaining 8 pairs of values of x and y; *(3 marks)*

(ii) comment, in context, on the value that you obtain. *(1 mark)*

5 (a) The baggage loading time, X minutes, of a chartered aircraft at its UK airport may be modelled by a normal random variable with mean 55 and standard deviation 8.

Determine:

(i) $P(X < 60)$; *(3 marks)*

(ii) $P(55 < X < 60)$. *(2 marks)*

(b) The baggage loading time, Y minutes, of a chartered aircraft at its overseas airport may be modelled by a normal random variable with mean μ and standard deviation 16.

Given that $P(Y < 90) = 0.95$, find the value of μ. *(4 marks)*

6 The table shows, for a particular population, the proportion of people in each of the four main blood groups.

Blood group	**O**	**A**	**B**	**AB**
Proportion	0.40	0.28	0.20	0.12

(a) A random sample of 20 people is selected from this population.

Determine the probability that the sample contains:

(i) at most 10 people with blood group O; *(2 marks)*

(ii) exactly 3 people with blood group A; *(3 marks)*

(iii) more than 4 but fewer than 8 people with blood group B. *(3 marks)*

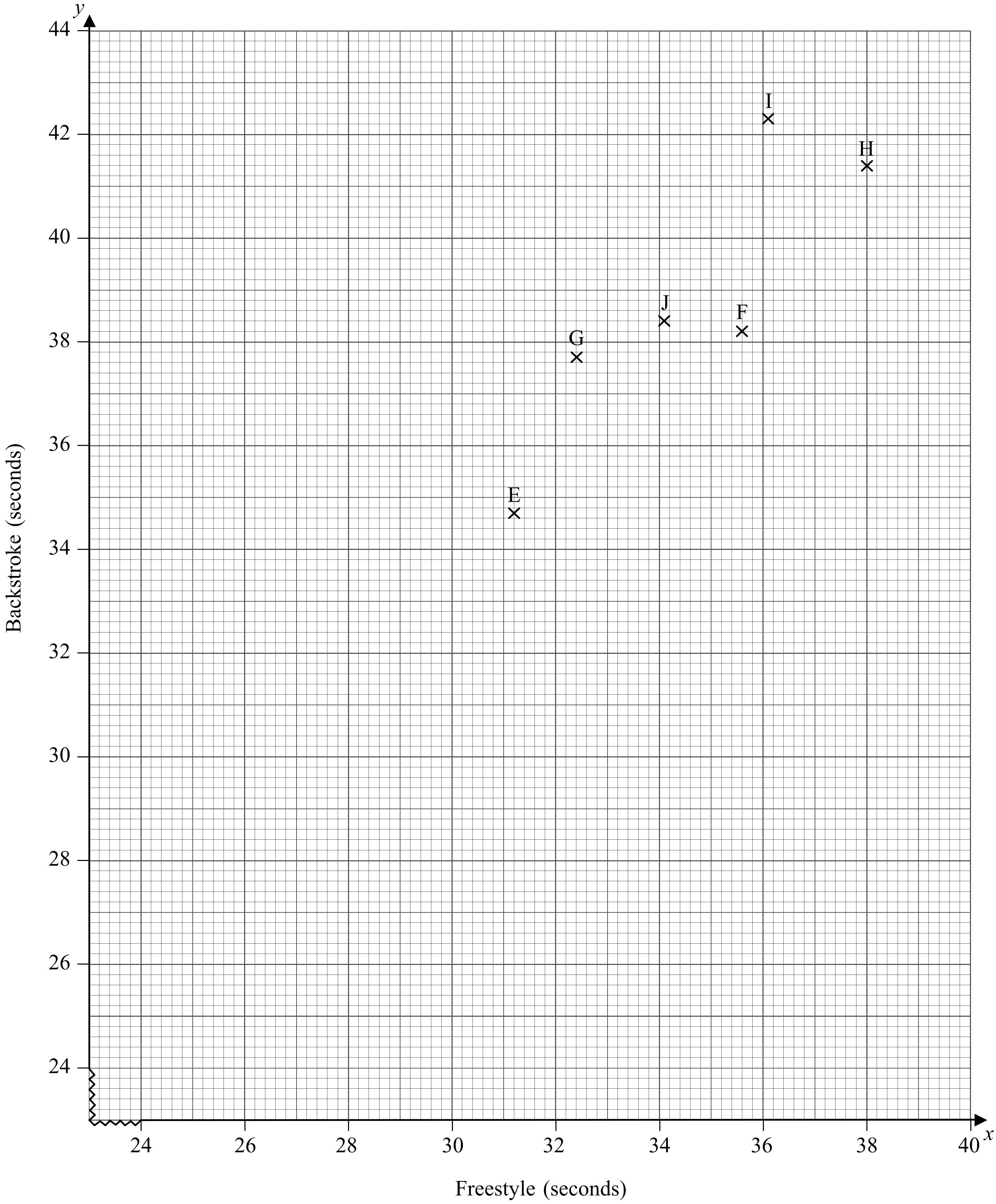
(b) A random sample of 500 people is selected from this population.

Find values for the mean and variance of the number of people in the sample with blood group AB. *(2 marks)*

END OF QUESTIONS

Figure 1 (for use in Question 4)

Scatter Diagram for Freestyle and Backstroke Swimming Times



1 **(i)** The letters A, B, C, D and E are arranged in a straight line.

(a) How many different arrangements are possible? [2]

(b) In how many of these arrangements are the letters A and B next to each other? [3]

(ii) From the letters A, B, C, D and E, two different letters are selected at random. Find the probability that these two letters are A and B. [2]

2 A random variable T has the distribution $\text{Geo}\left(\frac{1}{5}\right)$. Find

(i) $\text{P}(T = 4)$, [2]

(ii) $\text{P}(T > 4)$, [2]

(iii) $\text{E}(T)$. [1]

3 A sample of bivariate data was taken and the results were summarised as follows.

$$n = 5 \qquad \Sigma x = 24 \qquad \Sigma x^2 = 130 \qquad \Sigma y = 39 \qquad \Sigma y^2 = 361 \qquad \Sigma xy = 212$$

(i) Show that the value of the product moment correlation coefficient r is 0.855, correct to 3 significant figures. [2]

(ii) The ranks of the data were found. One student calculated Spearman's rank correlation coefficient r_s, and found that $r_s = 0.7$. Another student calculated the product moment coefficient, R, of these ranks. State which one of the following statements is true, and explain your answer briefly.

(A) $R = 0.855$
(B) $R = 0.7$
(C) It is impossible to give the value of R without carrying out a calculation using the original data.

[2]

(iii) All the values of x are now multiplied by a scaling factor of 2. State the new values of r and r_s. [2]

4 A supermarket has a large stock of eggs. 40% of the stock are from a firm called Eggzact. 12% of the stock are brown eggs from Eggzact.

An egg is chosen at random from the stock. Calculate the probability that

(i) this egg is brown, given that it is from Eggzact, [2]

(ii) this egg is from Eggzact and is not brown. [2]

5 **(i)** 20% of people in the large town of Carnley support the Residents' Party. 12 people from Carnley are selected at random. Out of these 12 people, the number who support the Residents' Party is denoted by U.

Find

(a) $\mathrm{P}(U \leqslant 5)$, [2]

(b) $\mathrm{P}(U \geqslant 3)$. [3]

(ii) 30% of people in Carnley support the Commerce Party. 15 people from Carnley are selected at random. Out of these 15 people, the number who support the Commerce Party is denoted by V.

Find $\mathrm{P}(V = 4)$. [3]

6 The probability distribution for a random variable Y is shown in the table.

y	1	2	3
$\mathrm{P}(Y = y)$	0.2	0.3	0.5

(i) Calculate $\mathrm{E}(Y)$ and $\mathrm{Var}(Y)$. [5]

Another random variable, Z, is independent of Y. The probability distribution for Z is shown in the table.

z	1	2	3
$\mathrm{P}(Z = z)$	0.1	0.25	0.65

One value of Y and one value of Z are chosen at random. Find the probability that

(ii) $Y + Z = 3$, [3]

(iii) $Y \times Z$ is even. [3]

7 **(i)** Andrew plays 10 tennis matches. In each match he either wins or loses.

(a) State, in this context, two conditions needed for a binomial distribution to arise. [2]

(b) Assuming these conditions are satisfied, define a variable in this context which has a binomial distribution. [1]

(ii) The random variable X has the distribution $\mathrm{B}(21, p)$, where $0 < p < 1$.

Given that $\mathrm{P}(X = 10) = \mathrm{P}(X = 9)$, find the value of p. [5]

[Turn over

8 The stem-and-leaf diagram shows the age in completed years of the members of a sports club.

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

Key: 1 | 4 | 0 represents a male aged 41 and a female aged 40.

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

(ii) The median and interquartile range for the females are 27 and 15 respectively. Make two comparisons between the ages of the males and the ages of the females. [2]

(iii) The mean age of the males is 30.7 and the mean age of the females is 27.5, each correct to 1 decimal place. Give one advantage of using the median rather than the mean to compare the ages of the males with the ages of the females. [1]

A record was kept of the number of hours, X, spent by each member at the club in a year. The results were summarised by

$$n = 49, \quad \Sigma(x - 200) = 245, \quad \Sigma(x - 200)^2 = 9849.$$

(iv) Calculate the mean and standard deviation of X. [6]

9 It is thought that the pH value of sand (a measure of the sand's acidity) may affect the extent to which a particular species of plant will grow in that sand. A botanist wished to determine whether there was any correlation between the pH value of the sand on certain sand dunes, and the amount of each of two plant species growing there. She chose random sections of equal area on each of eight sand dunes and measured the pH values. She then measured the area within each section that was covered by each of the two species. The results were as follows.

	Dune	A	B	C	D	E	F	G	H
	pH value, x	8.5	8.5	9.5	8.5	6.5	7.5	8.5	9.0
Area, $y\,\text{cm}^2$, covered	Species P	150	150	575	330	45	15	340	330
	Species Q	170	15	80	230	75	25	0	0

The results for species P can be summarised by

$$n = 8, \quad \Sigma x = 66.5, \quad \Sigma x^2 = 558.75, \quad \Sigma y = 1935, \quad \Sigma y^2 = 711\,275, \quad \Sigma xy = 17\,082.5.$$

(i) Give a reason why it might be appropriate to calculate the equation of the regression line of y on x rather than x on y in this situation. [1]

(ii) Calculate the equation of the regression line of y on x for species P, in the form $y = a + bx$, giving the values of a and b correct to 3 significant figures. [4]

(iii) Estimate the value of y for species P on sand where the pH value is 7.0. [2]

The values of the product moment correlation coefficient between x and y for species P and Q are $r_P = 0.828$ and $r_Q = 0.0302$.

(iv) Describe the relationship between the area covered by species Q and the pH value. [1]

(v) State, with a reason, whether the regression line of y on x for species P will provide a reliable estimate of the value of y when the pH value is

(a) 8, [1]

(b) 4. [1]

(vi) Assume that the equation of the regression line of y on x for species Q is also known. State, with a reason, whether this line will provide a reliable estimate of the value of y when the pH value is 8. [1]

Answer **all** questions.

1 The table shows, for each of a random sample of 8 paperback fiction books, the number of pages, x, and the recommended retail price, £y, to the nearest 10p.

x	223	276	374	433	564	612	704	766
y	6.50	4.00	5.50	8.00	4.50	5.00	8.00	5.50

(a) Calculate the value of the product moment correlation coefficient between x and y. *(3 marks)*

(b) Interpret your value in the context of this question. *(2 marks)*

(c) Suggest one other variable, in addition to the number of pages, which may affect the recommended retail price of a paperback fiction book. *(1 mark)*

2 A new car tyre is fitted to a wheel. The tyre is inflated to its recommended pressure of 265 kPa and the wheel left unused. At 3-month intervals thereafter, the tyre pressure is measured with the following results:

Time after fitting (x months)	0	3	6	9	12	15	18	21	24
Tyre pressure (y kPa)	265	250	240	235	225	215	210	195	180

(a) Calculate the equation of the least squares regression line of y on x. *(4 marks)*

(b) Interpret in context the value for the gradient of your line. *(2 marks)*

(c) Comment on the value for the intercept with the y-axis of your line. *(2 marks)*

3 Kirk and Les regularly play each other at darts.

(a) The probability that Kirk wins any game is 0.3, and the outcome of each game is independent of the outcome of every other game.

Find the probability that, in a match of 15 games, Kirk wins:

(i) fewer than half of the games; *(4 marks)*

(ii) more than 2 but fewer than 7 games. *(3 marks)*

(b) Kirk attends darts coaching sessions for three months. He then claims that he has a probability of 0.4 of winning any game, and that the outcome of each game is independent of the outcome of every other game.

(i) Assuming this claim to be true, calculate the mean and standard deviation for the number of games won by Kirk in a match of 15 games. *(3 marks)*

(ii) To assess Kirk's claim, Les keeps a record of the number of games won by Kirk in a series of 10 matches, each of 15 games, with the following results:

8	5	6	3	9	12	4	2	6	5

Calculate the mean and standard deviation of these values. *(2 marks)*

(iii) Hence comment on the validity of Kirk's claim. *(3 marks)*

Turn over for the next question

4 A housing estate consists of 320 houses: 120 detached and 200 semi-detached. The numbers of children living in these houses are shown in the table.

	Number of children				
	None	**One**	**Two**	**At least three**	**Total**
Detached house	24	32	41	23	120
Semi-detached house	40	37	88	35	200
Total	64	69	129	58	320

A house on the estate is selected at random.

D denotes the event 'the house is detached'.
R denotes the event 'no children live in the house'.
S denotes the event 'one child lives in the house'.
T denotes the event 'two children live in the house'.
(D' denotes the event 'not D'.)

(a) Find:

(i) $\mathrm{P}(D)$; *(1 mark)*

(ii) $\mathrm{P}(D \cap R)$; *(1 mark)*

(iii) $\mathrm{P}(D \mid R)$; *(2 marks)*

(iv) $\mathrm{P}(R \mid D')$. *(3 marks)*

(b) (i) Name two of the events D, R, S and T that are mutually exclusive. *(1 mark)*

(ii) Determine whether the events D and R are independent. Justify your answer. *(2 marks)*

(c) Define, in the context of this question, the event:

(i) $D' \cup T$; *(2 marks)*

(ii) $D \cap (R \cup S)$. *(2 marks)*

5 Currants are sold in 1000-gram packets and in 500-gram packets.

(a) The weights of 1000-gram packets may be assumed to be normally distributed with a mean of 1012 grams and a standard deviation of 5 grams.

Determine the probability that a randomly selected packet weighs:

(i) less than 1015 grams; *(3 marks)*

(ii) more than 1005 grams; *(3 marks)*

(iii) between 1005 grams and 1015 grams. *(2 marks)*

(b) The weight, y grams, of each of a random sample of fifty 500-gram packets of currants was recorded with the following results, where $\bar{y}$ denotes the sample mean:

$$n = 50 \qquad \sum y = 25\,142.5 \qquad \sum (y - \bar{y})^2 = 2519.0361$$

Number of packets weighing less than 500 grams = 6

(i) Construct a 98% confidence interval for the mean weight of 500-gram packets of currants, giving the limits to two decimal places. *(6 marks)*

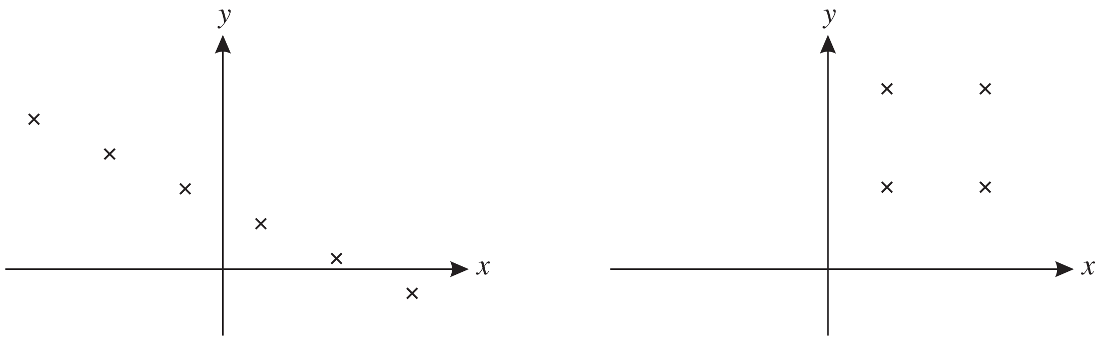
(ii) On each packet it states ‘Contents 500 grams’.

Comment on this statement using **both** the given information **and** your confidence interval. *(3 marks)*

END OF QUESTIONS

1 **(i)** State the value of the product moment correlation coefficient for each of the following scatter diagrams. [2]

(a) **(b)**



(ii) Calculate the value of Spearman's rank correlation coefficient for the following data. [5]

x	3.8	4.1	4.5	5.3
y	1.4	0.8	0.7	1.2

2 A class consists of 7 students from Ashville and 8 from Bewton. A committee of 5 students is chosen at random from the class.

(i) Find the probability that 2 students from Ashville and 3 from Bewton are chosen. [3]

(ii) In fact 2 students from Ashville and 3 from Bewton are chosen. In order to watch a video, all 5 committee members sit in a row. In how many different orders can they sit if no two students from Bewton sit next to each other? [2]

3 **(i)** A random variable X has the distribution $\mathrm{B}(8, 0.55)$. Find

(a) $\mathrm{P}(X < 7)$, [1]

(b) $\mathrm{P}(X = 5)$, [2]

(c) $\mathrm{P}(3 \leqslant X < 6)$. [3]

(ii) A random variable Y has the distribution $\mathrm{B}\left(10, \frac{5}{12}\right)$. Find

(a) $\mathrm{P}(Y = 2)$, [2]

(b) $\mathrm{Var}(Y)$. [1]

4 At a fairground stall, on each turn a player receives prize money with the following probabilities.

Prize money	£0.00	£0.50	£5.00
Probability	$\frac{17}{20}$	$\frac{1}{10}$	$\frac{1}{20}$

(i) Find the probability that a player who has two turns will receive a total of £5.50 in prize money. [3]

(ii) The stall-holder wishes to make a profit of 20p per turn on average. Calculate the amount the stall-holder should charge for each turn. [4]

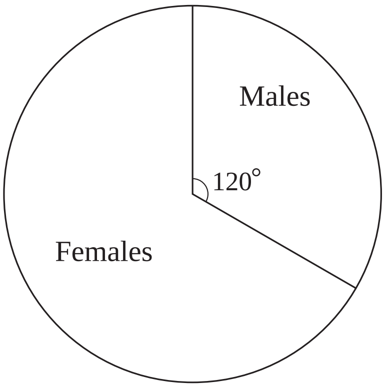
5 **(i)** A bag contains 12 red discs and 10 black discs. Two discs are removed at random, without replacement. Find the probability that both discs are red. [2]

(ii) Another bag contains 7 green discs and 8 blue discs. Three discs are removed at random, without replacement. Find the probability that exactly two of these discs are green. [3]

(iii) A third bag contains 45 discs, each of which is either yellow or brown. Two discs are removed at random, without replacement. The probability that both discs are yellow is $\frac{1}{15}$. Find the number of yellow discs which were in the bag at first. [4]

[Turn over

6 (i) The numbers of males and females in Year 12 at a school are illustrated in the pie chart. The number of males in Year 12 is 128.

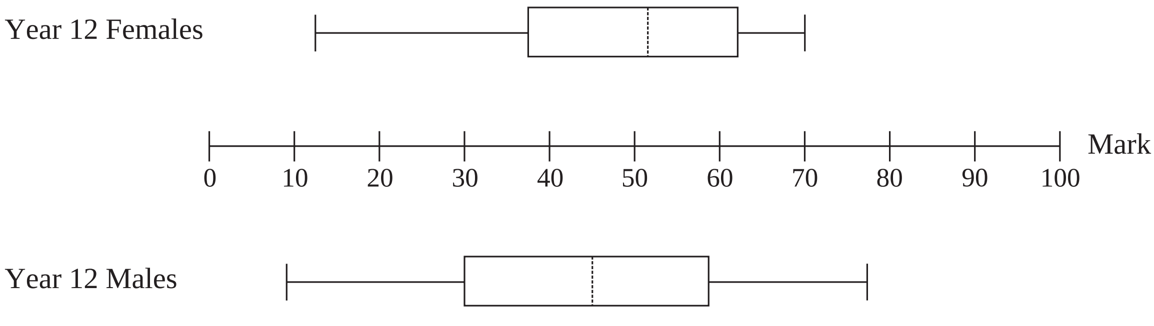


Year 12

(a) Find the number of females in Year 12. [1]

(b) On a corresponding pie chart for Year 13, the angle of the sector representing males is 150°. Explain why this does not necessarily mean that the number of males in Year 13 is more than 128. [1]

(ii) All the Year 12 students took a General Studies examination. The results are illustrated in the box-and-whisker plots.



(a) One student said "The Year 12 pie chart shows that there are more females than males, but the box-and-whisker plots show that there are more males than females."

Comment on this statement. [1]

(b) Give two comparisons between the overall performance of the females and the males in the General Studies examination. [2]

(c) Give one advantage and one disadvantage of using box-and-whisker plots rather than histograms to display the results. [2]

(iii) The mean mark for 102 of the male students was 51. The mean mark for the remaining 26 male students was 59. Calculate the mean mark for all 128 male students. [3]

7 Once each year, Paula enters a lottery for a place in an annual marathon. Each time she enters the lottery, the probability of her obtaining a place is 0.3. Find the probability that

(i) the first time she obtains a place is on her 4th attempt, [3]

(ii) she does not obtain a place on any of her first 6 attempts, [2]

(iii) she needs fewer than 10 attempts to obtain a place, [3]

(iv) she obtains a place exactly twice in her first 5 attempts. [3]

8 A city council attempted to reduce traffic congestion by introducing a congestion charge. The charge was set at £4.00 for the first year and was then increased by £2.00 each year. For each of the first eight years, the council recorded the average number of vehicles entering the city centre per day. The results are shown in the table.

Charge, £x	4	6	8	10	12	14	16	18
Average number of vehicles per day, y million	2.4	2.5	2.2	2.3	2.0	1.8	1.7	1.5

[$n = 8$, $\Sigma x = 88$, $\Sigma y = 16.4$, $\Sigma x^2 = 1136$, $\Sigma y^2 = 34.52$, $\Sigma xy = 168.6$.]

(i) Calculate the product moment correlation coefficient for these data. [3]

(ii) Explain why x is the independent variable. [1]

(iii) Calculate the equation of the regression line of y on x. [4]

(iv) **(a)** Use your equation to estimate the average number of vehicles which will enter the city centre per day when the congestion charge is raised to £20.00. [2]

(b) Comment on the reliability of your estimate. [2]

(v) The council wishes to estimate the congestion charge required to reduce the average number of vehicles entering the city per day to 1.0 million. Assuming that a reliable estimate can be made by extrapolation, state whether they should use the regression line of y on x or the regression line of x on y. Give a reason for your answer. [2]

Answer **all** questions.

1 The times, in seconds, taken by 20 people to solve a simple numerical puzzle were

17	19	22	26	28	31	34	36	38	39
41	42	43	47	50	51	53	55	57	58

(a) Calculate the mean and the standard deviation of these times. *(3 marks)*

(b) In fact, 23 people solved the puzzle. However, 3 of them failed to solve it within the allotted time of 60 seconds.

Calculate the median and the interquartile range of the times taken by all 23 people. *(4 marks)*

(c) For the times taken by all 23 people, explain why:

(i) the mode is **not** an appropriate numerical measure;

(ii) the range is **not** an appropriate numerical measure. *(2 marks)*

2 The post office in a market town is located within a small supermarket.

The probability that an individual customer entering the supermarket requires a service from:

- the post office only is 0.48;
- the supermarket only is 0.30;
- both the post office and the supermarket is 0.22 .

It may be assumed that the service required is independent from customer to customer.

(a) For a random sample of 12 individual customers, calculate the probability that exactly 5 of them require a service from the post office only. *(3 marks)*

(b) For a random sample of 40 individual customers, determine the probability that more than 10 but fewer than 15 of them require a service from the supermarket only. *(3 marks)*

(c) For a random sample of 100 individual customers, calculate the mean and the standard deviation for the number of them requiring a service from both the post office and the supermarket. *(3 marks)*

3 A very popular play has been performed at a London theatre on each of 6 evenings per week for about a year. Over the past 13 weeks (78 performances), records have been kept of the proceeds from the sales of programmes at each performance. An analysis of these records has found that the mean was £184 and the standard deviation was £32.

(a) Assuming that the 78 performances may be considered to be a random sample, construct a 90% confidence interval for the mean proceeds from the sales of programmes at an evening performance of this play. *(4 marks)*

(b) Comment on the likely validity of the assumption in part (a) when constructing a confidence interval for the mean proceeds from the sales of programmes at an evening performance of:

(i) this particular play;

(ii) any play. *(3 marks)*

4 Rea, Suki and Tora take part in a shooting competition. The final round of the competition requires each of them to try to hit the centre of a target, placed at 100 metres, with a single shot. The independent probabilities that Rea, Suki and Tora hit the centre of this target with a single shot are 0.7, 0.6 and 0.8 respectively.

Find the probability that, in the final round of the competition, the centre of the target will be hit by:

(a) Tora only; *(2 marks)*

(b) exactly one of the three competitors; *(3 marks)*

(c) at least one of the three competitors. *(3 marks)*

Turn over for the next question

5 When Monica walks to work from home, she uses either route A or route B.

(a) Her journey time, X minutes, by route A may be assumed to be normally distributed with a mean of 37 and a standard deviation of 8.

Determine:

(i) $\mathrm{P}(X < 45)$; *(3 marks)*

(ii) $\mathrm{P}(30 < X < 45)$. *(3 marks)*

(b) Her journey time, Y minutes, by route B may be assumed to be normally distributed with a mean of 40 and a standard deviation of σ.

Given that $\mathrm{P}(Y > 45) = 0.12$, calculate the value of σ. *(4 marks)*

(c) If Monica leaves home at 8.15 am to walk to work hoping to arrive by 9.00 am, state, with a reason, which route she should take. *(2 marks)*

6 [**Figure 1**, printed on the insert, is provided for use in this question.]

Stan is a retired academic who supplements his pension by mowing lawns for customers who live nearby.

As part of a review of his charges for this work, he measures the areas, x m^2, of a random sample of eight of his customers' lawns and notes the times, y minutes, that it takes him to mow these lawns. His results are shown in the table.

Customer	**A**	**B**	**C**	**D**	**E**	**F**	**G**	**H**
x	360	140	860	600	1180	540	260	480
y	50	25	135	70	140	90	55	70

(a) On **Figure 1**, plot a scatter diagram of these data. *(2 marks)*

(b) Calculate the equation of the least squares regression line of y on x. Draw your line on **Figure 1**. *(6 marks)*

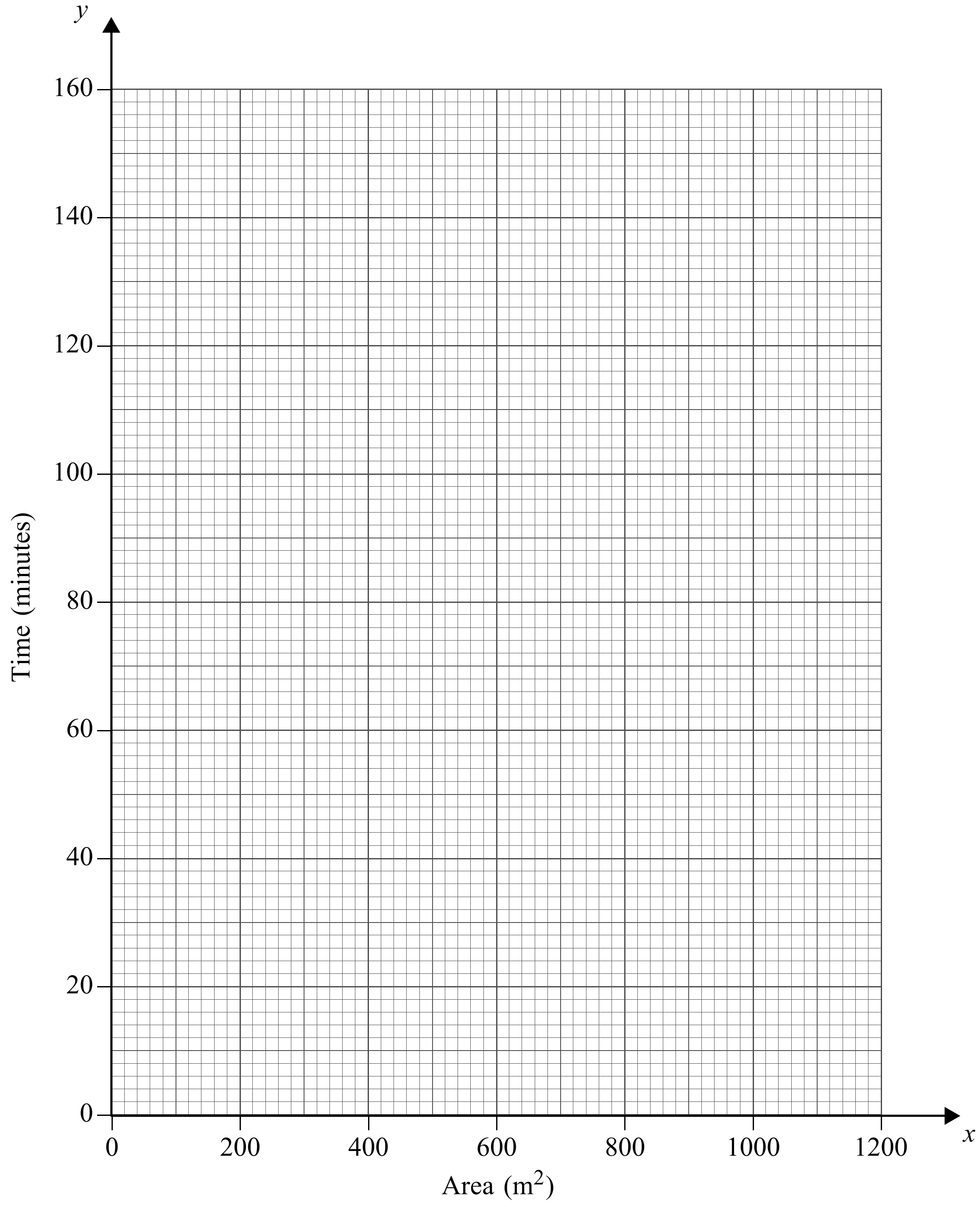
(c) Calculate the value of the residual for Customer H and indicate how your value is confirmed by your scatter diagram. *(3 marks)*

(d) Given that Stan charges £12 per hour, estimate the charge for mowing a customer's lawn that has an area of 560 m^2. *(4 marks)*

END OF QUESTIONS

Figure 1 (for use in Question 6)

Lawn Areas and Mowing Times



Answer **all** questions.

1 The table shows the length, in centimetres, and maximum diameter, in centimetres, of each of 10 honeydew melons selected at random from those on display at a market stall.

Length	24	25	19	28	27	21	35	23	32	26
Maximum diameter	18	14	16	11	13	14	12	16	15	14

(a) Calculate the value of the product moment correlation coefficient. *(3 marks)*

(b) Interpret your value in the context of this question. *(2 marks)*

2 The British and Irish Lions 2005 rugby squad contained 50 players. The nationalities and playing positions of these players are shown in the table.

		Nationality			
		English	**Welsh**	**Scottish**	**Irish**
Playing position	**Forward**	14	5	2	6
	Back	8	7	2	6

(a) A player was selected at random from the squad for a radio interview. Calculate the probability that the player was:

(i) English; *(2 marks)*

(ii) Irish, given that the player was a back; *(2 marks)*

(iii) a forward, given that the player was not Scottish. *(2 marks)*

(b) Four players were selected at random from the squad to visit a school. Calculate the probability that all four players were English. *(3 marks)*

3 Payton has a pay-as-you-go internet account. To save money, he reads his e-mail messages just after 6 pm each day. The probability that he has no e-mail messages to read at this time is 0.45, and the number of e-mail messages he receives is independent from day to day.

(a) Calculate the probability that Payton has no e-mail messages to read on exactly 3 days during a 7-day period. *(3 marks)*

(b) Determine the probability that, during June (30 days), Payton has no e-mail messages to read:

(i) on fewer than 15 days; *(2 marks)*

(ii) on more than 10 days; *(2 marks)*

(iii) on at least 12 but at most 18 days. *(3 marks)*

4 A library allows each member to have up to 15 books on loan at any one time.

The table shows the numbers of books currently on loan to a random sample of 95 members of the library.

Number of books on loan	0	1	2	3	4	5–9	10–14	15
Number of members	4	13	24	17	15	11	5	6

(a) For these data:

(i) state values for the mode and range; *(2 marks)*

(ii) determine values for the median and interquartile range; *(4 marks)*

(iii) calculate estimates of the mean and standard deviation. *(4 marks)*

(b) Making reference to your answers to part (a), give a reason for preferring:

(i) the median and interquartile range to the mean and standard deviation for summarising the given data; *(1 mark)*

(ii) the mean and standard deviation to the mode and range for summarising the given data. *(1 mark)*

Turn over for the next question

5 Bob, a gardener, measures the time taken, y minutes, for 60 grams of weedkiller pellets to dissolve in 10 litres of water at different set temperatures, x °C. His results are shown in the table.

x	16	20	24	28	32	36	40	44	48	52	56
y	4.7	4.3	3.8	3.5	3.0	2.7	2.4	2.0	1.8	1.6	1.1

(a) Calculate the equation of the least squares regression line $y = a + bx$. *(4 marks)*

(b) (i) Interpret, in the context of this question, your value for b. *(2 marks)*

(ii) Explain why no sensible practical interpretation can be given for your value of a. *(2 marks)*

6 (a) The length, X centimetres, of adult male eels in a river may be assumed to be normally distributed with a mean of 38 and a standard deviation of 5.

Determine:

(i) $P(X < 40)$; *(2 marks)*

(ii) $P(30 < X < 40)$; *(3 marks)*

(iii) the length exceeded by 75% of adult male eels in the river. *(4 marks)*

(b) A sample of 40 adult female eels was taken at random from the river and the length of each eel was measured.

The mean and standard deviation of these lengths were found to be 107 cm and 19.1 cm respectively.

(i) Construct a 98% confidence interval for the mean length of adult female eels in the river. *(4 marks)*

(ii) Hence comment on a claim that, in this river, the average length of adult female eels is more than $2\frac{1}{2}$ times that of adult male eels. *(3 marks)*

END OF QUESTIONS