Please check the examination details below before entering your candidate information

Candidate surname	Other names

Centre Number | Candidate Number

Pearson Edexcel International Advanced Level

Wednesday 10 January 2024

Afternoon (Time: 1 hour 30 minutes)

Paper reference **WST01/01**

Mathematics

International Advanced Subsidiary/Advanced Level
Statistics S1

You must have:
Mathematical Formulae and Statistical Tables (Yellow), calculator

Total Marks

Candidates may use any calculator permitted by Pearson regulations. Calculators must not have the facility for symbolic algebra manipulation, differentiation and integration, or have retrievable mathematical formulae stored in them.

Instructions

- Use **black** ink or ball-point pen.
- If pencil is used for diagrams/sketches/graphs it must be dark (HB or B).
- **Fill in the boxes** at the top of this page with your name, centre number and candidate number.
- Answer **all** questions and ensure that your answers to parts of questions are clearly labelled.
- Answer the questions in the spaces provided
  – *there may be more space than you need.*
- You should show sufficient working to make your methods clear. Answers without working may not gain full credit.
- Values from the statistical tables should be quoted in full. If a calculator is used instead of the tables, the value should be given to an equivalent degree of accuracy.
- Inexact answers should be given to three significant figures unless otherwise stated.

Information

- A booklet 'Mathematical Formulae and Statistical Tables' is provided.
- There are 8 questions in this question paper. The total mark for this paper is 75.
- The marks for **each** question are shown in brackets
  – *use this as a guide as to how much time to spend on each question.*

Advice

- Read each question carefully before you start to answer it.
- Try to answer every question.
- Check your answers if you have time at the end.
- If you change your mind about an answer, cross it out and put your new answer and any working underneath.

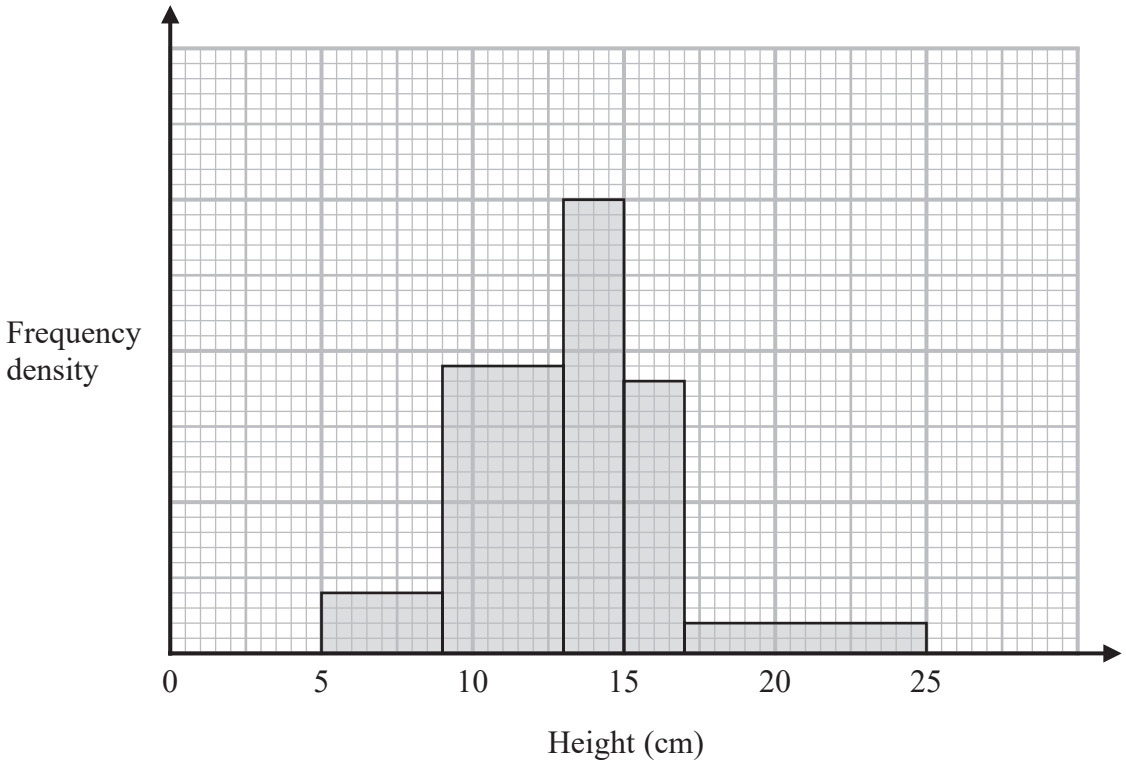
Turn over ▶

P74320A

P 7 4 3 2 0 A 0 1 2 8

Pearson

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)

Question 1 continued

Question 1 continued

Question 1 continued

(Total for Question 1 is 8 marks)

2. The average minimum monthly temperature, x degrees Fahrenheit (°F), and the average maximum monthly temperature, y degrees Fahrenheit (°F), in Kolkata were recorded for 12 months.

Some of the summary statistics are given below.

$$\sum x = 862 \quad \sum x^2 = 62\,802 \quad S_{yy} = 413.67 \quad S_{xy} = 512.67 \quad n = 12$$

(a) (i) Calculate the mean of the 12 values of the average **minimum** monthly temperature.

(ii) Show that the standard deviation of the 12 values of the average **minimum** monthly temperature is 8.57°F to 3 significant figures.

(3)

(b) Calculate the product moment correlation coefficient between x and y

(3)

For comparative purposes with a UK city, it was necessary to convert the temperatures from degrees Fahrenheit (°F) to degrees Celsius (°C).

The formula used was

$$c = \frac{5}{9}(f - 32)$$

where f is the temperature in °F and c is the temperature in °C

(c) Use this formula and the values from part (a) to calculate, in °C, the mean and the standard deviation of the 12 values of the average **minimum** monthly temperature in Kolkata.
Give your answers to 3 significant figures.

(4)

Given that

- u is the equivalent temperature in °C of x
- v is the equivalent temperature in °C of y

(d) state, giving a reason, the product moment correlation coefficient between u and v

(2)

Question 2 continued

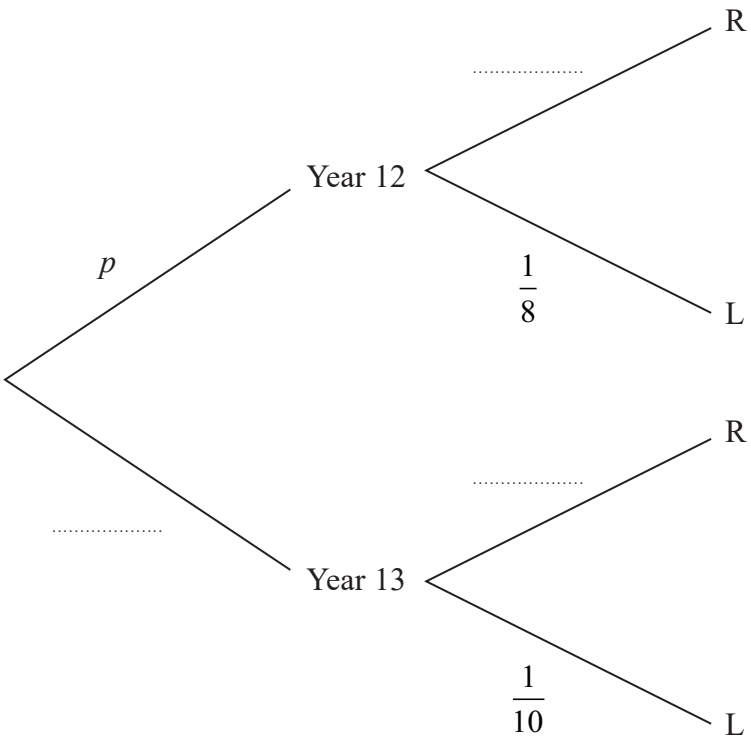
Question 2 continued

Question 2 continued

(Total for Question 2 is 12 marks)

3. In a sixth form college each student in Year 12 and Year 13 is either left-handed (L) or right-handed (R).

The partially completed tree diagram, where p is a probability, gives information about these students.



(a) Complete the tree diagram, in terms of p where necessary.

(1)

The probability that a student is left-handed is 0.11

(b) Find the value of p

(3)

(c) Find the probability that a student selected at random is in Year 12 and left-handed.

(2)

Given that a student is right-handed,

(d) find the probability that the student is in Year 12

(2)

Question 3 continued

(Total for Question 3 is 8 marks)

4. A French test and a Spanish test were sat by 11 students.

The table below shows their marks.

Student	A	B	C	D	E	F	G	H	I	J	K
French mark (f)	24	30	32	32	36	36	40	44	50	60	68
Spanish mark (s)	16	90	24	28	32	36	38	44	48	48	68

Greg says that if these points were plotted on a scatter diagram, then the point (30, 90) would be an outlier because 90 is an outlier for the Spanish marks.

An outlier is defined as a value that is

$$\text{greater than } Q_3 + 1.5 \times (Q_3 - Q_1) \text{ or smaller than } Q_1 - 1.5 \times (Q_3 - Q_1)$$

(a) Show that 90 is an outlier for the Spanish marks.

(3)

Ignoring the point (30, 90), Greg calculated the following summary statistics.

$$\sum f = 422 \qquad \sum s = 382 \qquad S_{ff} = 1667.6 \qquad S_{fs} = 1735.6$$

(b) Use these summary statistics to show that the equation of the least squares regression line of s on f for the remaining 10 students is

$$s = -5.72 + 1.04f$$

where the values of the intercept and gradient are given to 3 significant figures. You must show your working.

(3)

(c) Give an interpretation of the gradient of the regression line.

(1)

Two further students sat the French test but missed the Spanish test.

(d) Using the equation given in part (b), estimate

(i) a Spanish mark for the student who scored 55 marks in their French test,

(ii) a Spanish mark for the student who scored 18 marks in their French test.

(3)

(e) State, giving a reason, which of the two estimates found in part (d) would be the more reliable estimate.

(2)

Question 4 continued

Question 4 continued

Question 4 continued

(Total for Question 4 is 12 marks)

5. The distance an athlete can throw a discus is normally distributed with mean 40 m and standard deviation 4 m

(a) Using standardisation, show that the probability that this athlete throws the discus less than 38.8 m is 0.3821

(2)

This athlete enters a discus competition.
To qualify for the final, they have 3 attempts to throw the discus a distance of more than 38.8 m
Once they qualify, they do not use any of their remaining attempts.

Given that they qualified for the final and that throws are independent,

(b) find the probability that this athlete qualified for the final on their second throw with a distance of more than 44 m

(5)

Question 5 continued

(Total for Question 5 is 7 marks)

6. The events A and B satisfy

$$\mathrm{P}(A) = x \qquad \mathrm{P}(B) = y \qquad \mathrm{P}(A \cup B) = 0.65 \qquad \mathrm{P}(B \mid A) = 0.3$$

(a) Show that

$$14x + 20y = 13$$

(3)

The events B and C are mutually exclusive such that

$$\mathrm{P}(B \cup C) = 0.85 \qquad \mathrm{P}(C) = \frac{1}{2}x + y$$

(b) (i) Find a second equation in x and y

(ii) Hence find the value of x and the value of y

(4)

(c) Determine whether or not A and B are statistically independent.
You must show your working clearly.

(2)

Question 6 continued

Question 6 continued

Question 6 continued

(Total for Question 6 is 9 marks)

7. The cumulative distribution of a discrete random variable X is given by

x	1	2	3	4
$F(x)$	$\frac{1}{13}$	$\frac{2k-1}{26}$	$\frac{3(k+1)}{26}$	$\frac{k+4}{8}$

where k is a positive constant.

(a) Show that $k = 4$

(1)

(b) Find the probability distribution of the discrete random variable X

(3)

(c) Using your answer to part (b), write down the mode of X

(1)

(d) Calculate $\text{Var}(13X - 6)$

(5)

Question 7 continued

Question 7 continued

Question 7 continued

(Total for Question 7 is 10 marks)

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)

Question 8 continued

Question 8 continued

(Total for Question 8 is 9 marks)

TOTAL FOR PAPER IS 75 MARKS