

Please check the examination details below before entering your candidate information

Candidate surname

Other names

Centre Number

Candidate Number

--	--	--	--	--

--	--	--	--	--

Pearson Edexcel International Advanced Level

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 6 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 ►

P72072A

©2023 Pearson Education Ltd.

J:1/1/1/

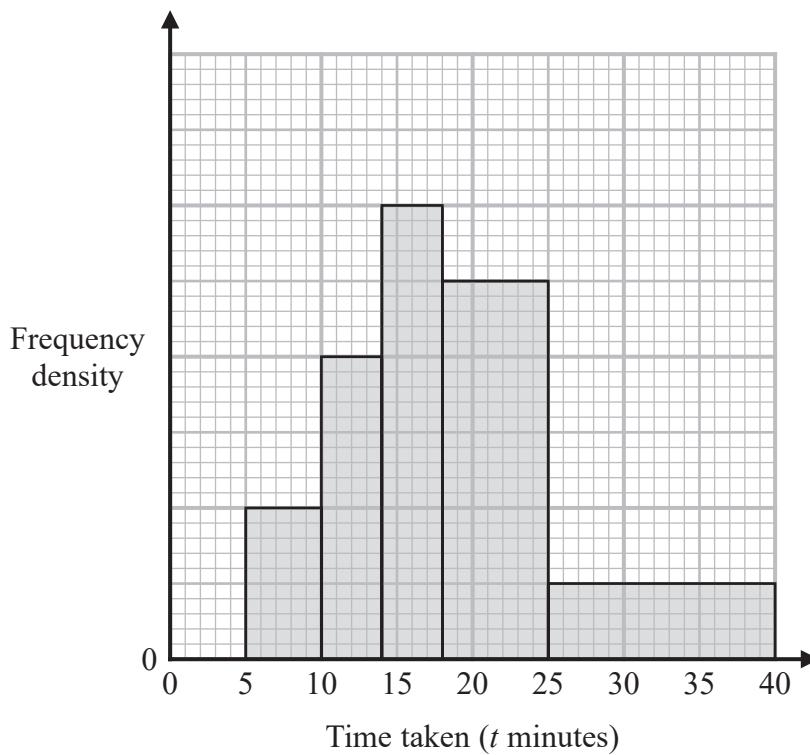


P 7 2 0 7 2 A 0 1 2 4



Pearson

1. The histogram shows the times taken, t minutes, by each of 100 people to swim 500 metres.



- (a) Use the histogram to complete the frequency table for the times taken by the 100 people to swim 500 metres.

Time taken (t minutes)	5 – 10	10 – 14	14 – 18	18 – 25	25 – 40
Frequency (f)	10	16	24		

(1)

- (b) Estimate the number of people who took less than 16 minutes to swim 500 metres.

(2)

- (c) Find an estimate for the mean time taken to swim 500 metres.

(2)

Given that $\sum ft^2 = 41033$

- (d) find an estimate for the standard deviation of the times taken to swim 500 metres.

(2)

Given that $Q_3 = 23$

- (e) use linear interpolation to estimate the interquartile range of the times taken to swim 500 metres.

(3)



Question 1 continued



P 7 2 0 7 2 A 0 3 2 4

Question 1 continued

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA



Question 1 continued

(Total for Question 1 is 10 marks)



2. Two bags, X and Y , each contain green marbles (G) and blue marbles (B) only.

- Bag X contains 5 green marbles and 4 blue marbles
- Bag Y contains 6 green marbles and 5 blue marbles

A marble is selected at random from bag X and placed in bag Y

A second marble is selected at random from bag X and placed in bag Y

A third marble is then selected, this time from bag Y

(a) Use this information to complete the tree diagram shown on page 7

(3)

(b) Find the probability that the 2 marbles selected from bag X are of different colours.

(2)

(c) Find the probability that all 3 marbles selected are the same colour.

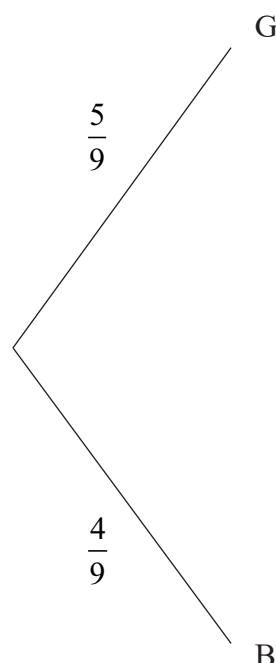
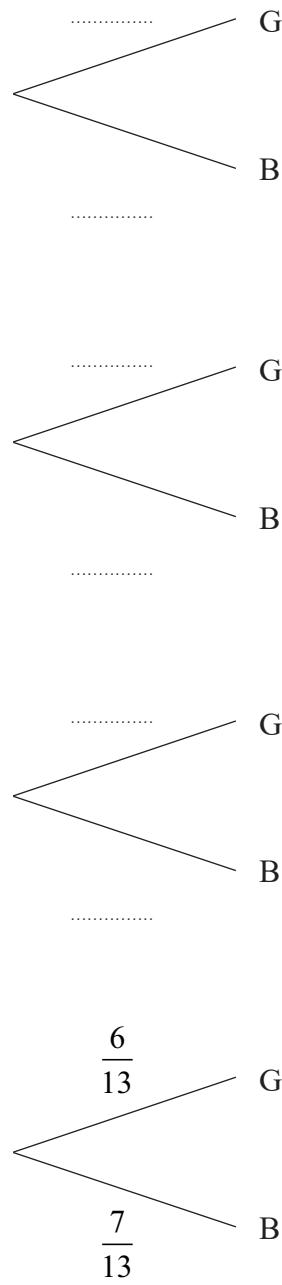
(2)

Given that all three marbles selected are the same colour,

(d) find the probability that they are all green.

(3)



Question 2 continued**1st Marble
(from bag X)****2nd Marble
(from bag X)****3rd Marble
(from bag Y)**

Question 2 continued

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA



Question 2 continued

(Total for Question 2 is 10 marks)



3. The probability distribution of the discrete random variable X is given by

x	2	3	4
$\text{P}(X = x)$	a	0.4	$0.6 - a$

where a is a constant.

- (a) Find, in terms of a , $E(X)$ (2)

- (b) Find the range of the possible values of $E(X)$ (3)

Given that $\text{Var}(X) = 0.56$

- (c) find the possible values of a (6)

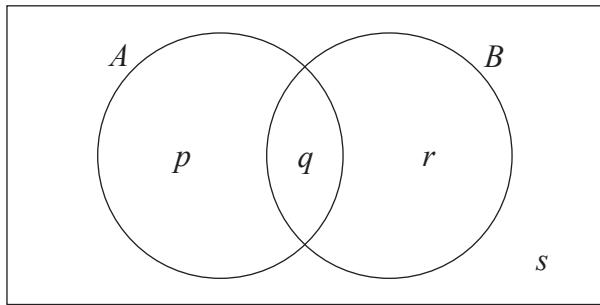


Question 3 continued

(Total for Question 3 is 11 marks)



4. (i) In the Venn diagram below, A and B represent events and p , q , r and s are probabilities.



$$P(A) = \frac{7}{25} \quad P(B) = \frac{1}{5} \quad P[(A \cap B') \cup (A' \cap B)] = \frac{8}{25}$$

- (a) Use algebra to show that $2p + 2q + 2r = \frac{4}{5}$ (4)

- (b) Find the value of p , the value of q , the value of r and the value of s (5)

- (ii) Two events, C and D , are such that

$$P(C) = \frac{x}{x+5} \quad P(D) = \frac{5}{x}$$

where x is a positive constant.

By considering $P(C) + P(D)$ show that C and D **cannot** be mutually exclusive.

(4)



Question 4 continued



Question 4 continued

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA



Question 4 continued

(Total for Question 4 is 13 marks)



5. The lengths, L mm, of housefly wings are normally distributed with $L \sim N(4.5, 0.4^2)$

- (a) Find the probability that a randomly selected housefly has a wing length of less than 3.86 mm.

(3)

- (b) Find

(i) the upper quartile (Q_3) of L

(ii) the lower quartile (Q_1) of L

(4)

A value that is greater than $Q_3 + 1.5 \times (Q_3 - Q_1)$ or smaller than $Q_1 - 1.5 \times (Q_3 - Q_1)$ is defined as an outlier.

- (c) Find these two outlier limits.

(3)

A housefly is selected at random.

- (d) Using standardisation, show that the probability that this housefly is **not** an outlier is 0.993 to 3 decimal places.

(3)

Given that this housefly is **not** an outlier,

- (e) showing your working, find the probability that the wing length of this housefly is greater than 5 mm.

(4)



Question 5 continued



Question 5 continued

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA



Question 5 continued

(Total for Question 5 is 17 marks)



6. A research student is investigating the maximum weight, y grams, of sugar that will dissolve in 100 grams of water at various temperatures, $x^{\circ}\text{C}$, where $10 \leqslant x \leqslant 80$

The research student calculated the regression line of y on x and found it to be

$$y = 151.2 + 2.72x$$

- (a) Give an interpretation of the gradient of the regression line.

(1)

- (b) Use the regression line to estimate the maximum weight of sugar that will dissolve in 100 grams of water when the temperature is 90°C .

(2)

- (c) Comment on the reliability of your estimate, giving a reason for your answer.

(2)

Using the regression line of y on x and the following summary statistics

$$\sum y = 3119 \quad \sum y^2 = 851\,093 \quad \sum x^2 = 24\,500 \quad n = 12$$

- (d) show that the product moment correlation coefficient for these data is 0.988 to 3 decimal places.

(7)

The research student's supervisor plotted the original data on a scatter diagram, shown on page 23

With reference to both the scatter diagram and the correlation coefficient,

- (e) discuss the suitability of a linear regression model to describe the relationship between x and y .

(2)



Question 6 continued



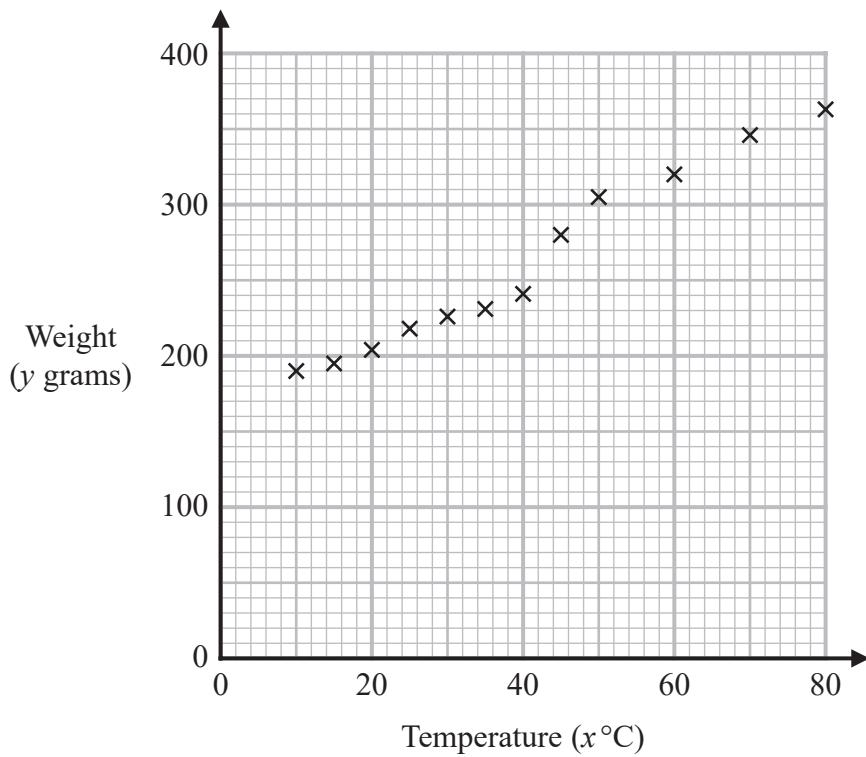
Question 6 continued

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA





P 7 2 0 7 2 A 0 2 3 2 4

Question 6 continued

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

(Total for Question 6 is 14 marks)

TOTAL FOR PAPER IS 75 MARKS

