Please check the examination details below before entering your candidate information

Candidate surname	Other names

Centre Number | Candidate Number

Pearson Edexcel International Advanced Level

Thursday 11 May 2023

Morning (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 your 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 7 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 ▶

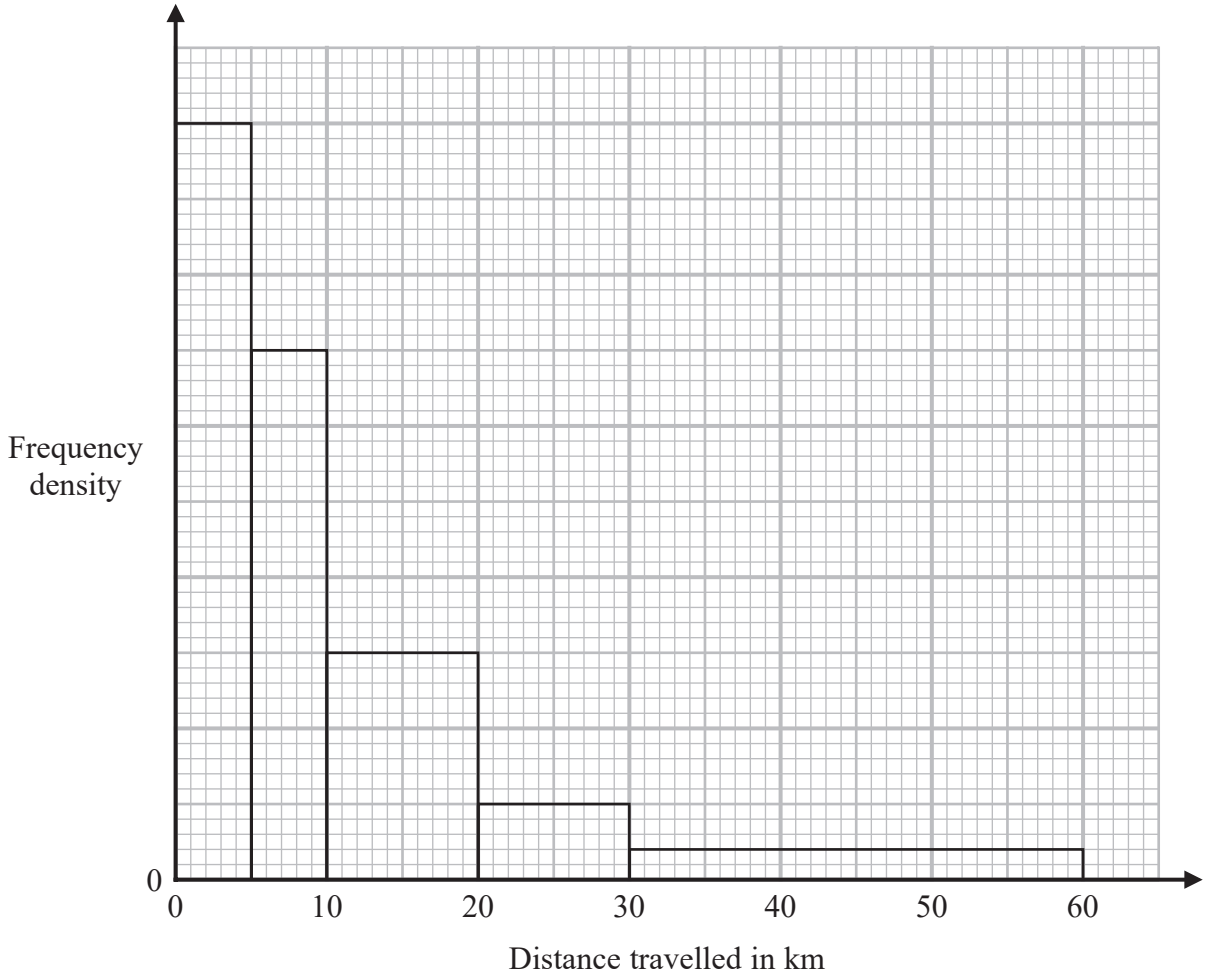
P72904A

©2023 Pearson Education Ltd.
N:1/1/1/

P 7 2 9 0 4 A 0 1 2 8

Pearson

1. The histogram shows the distances, in km, that 274 people travel to work.



Given that 60 of these people travel between 10 km and 20 km to work, estimate

(a) the number of people who travel between 22 km and 45 km to work,

(3)

(b) the median distance travelled to work by these 274 people,

(2)

(c) the mean distance travelled to work by these 274 people.

(3)

DO NOT WRITE IN THIS AREA
DO NOT WRITE IN THIS AREA
DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA
DO NOT WRITE IN THIS AREA
DO NOT WRITE IN THIS AREA

Question 1 continued

Question 1 continued

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

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 8 marks)

2. Two students, Olive and Shan, collect data on the weight, w grams, and the tail length, t cm, of 15 mice.

Olive summarised the data as follows

$S_{tt} = 5.3173$ $\quad \sum w^2 = 6089.12$ $\quad \sum tw = 2304.53$ $\quad \sum w = 297.8$ $\quad \sum t = 114.8$

(a) Calculate the value of S_{tw} and the value of S_{ww}

(3)

(b) Calculate the value of the product moment correlation coefficient between w and t

(2)

(c) Show that the equation of the regression line of w on t can be written as

$$w = -16.7 + 4.77t$$

(3)

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

(1)

(e) Explain why it would not be appropriate to use the regression line in part (c) to estimate the weight of a mouse with a tail length of 2 cm.

(2)

Shan decided to code the data using $x = t - 6$ and $y = \frac{w}{2} - 5$

(f) Write down the value of the product moment correlation coefficient between x and y

(1)

(g) Write down an equation of the regression line of y on x
You do not need to simplify your equation.

(1)

DO NOT WRITE IN THIS AREA
DO NOT WRITE IN THIS AREA
DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA
DO NOT WRITE IN THIS AREA
DO NOT WRITE IN THIS AREA

Question 2 continued

Question 2 continued

DO NOT WRITE IN THIS AREA
DO NOT WRITE IN THIS AREA
DO NOT WRITE IN THIS AREA

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 13 marks)

3. Jim records the length, l mm, of 81 salmon. The data are coded using $x = l - 600$ and the following summary statistics are obtained.

$$n = 81 \qquad \sum x = 3711 \qquad \sum x^2 = 475\,181$$

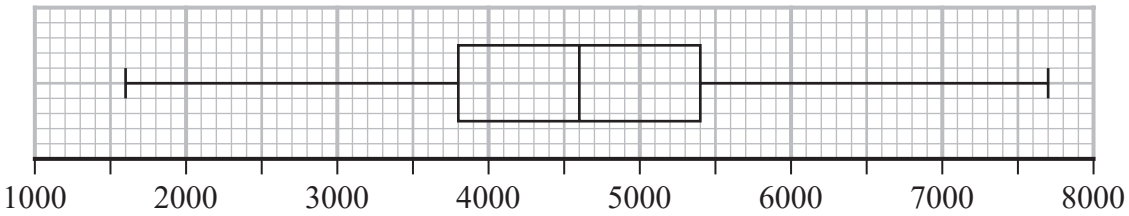
(a) Find the mean length of these salmon.

(3)

(b) Find the variance of the lengths of these salmon.

(2)

The weight, w grams, of each of the 81 salmon is recorded to the nearest gram.
The recorded results for the 81 salmon are summarised in the box plot below.



Weight of salmon (w grams)

(c) Find the maximum number of salmon that have weights in the interval

$$4600 < w \leqslant 7700$$

(1)

Raj says that the box plot is incorrect as Jim has not included outliers.

For these data an outlier is defined as a value that is more than

$1.5 \times$ IQR above the upper quartile or $1.5 \times$ IQR below the lower quartile

(d) Show that there are no outliers.

(3)

DO NOT WRITE IN THIS AREA
DO NOT WRITE IN THIS AREA
DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA
DO NOT WRITE IN THIS AREA
DO NOT WRITE IN THIS AREA

Question 3 continued

Question 3 continued

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA
DO NOT WRITE IN THIS AREA
DO NOT WRITE IN THIS AREA

Question 3 continued

(Total for Question 3 is 9 marks)

4. A bag contains a large number of coloured counters. Each counter is labelled A, B or C

30% of the counters are labelled A
45% of the counters are labelled B
The rest of the counters are labelled C

It is known that

2% of the counters labelled A are red
4% of the counters labelled B are red
6% of the counters labelled C are red

One counter is selected at random from the bag.

(a) Complete the tree diagram on the opposite page to illustrate this information.

(2)

(b) Calculate the probability that the counter is labelled A and is not red.

(2)

(c) Calculate the probability that the counter is red.

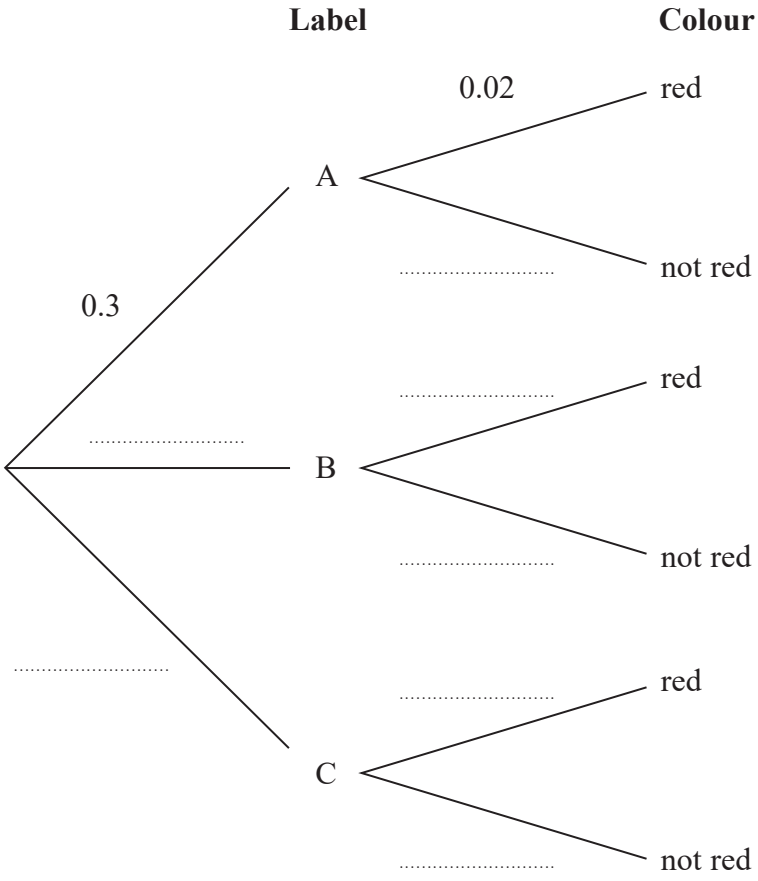
(2)

(d) Given that the counter is red, find the probability that it is labelled C

(3)

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 9 marks)

DO NOT WRITE IN THIS AREA
DO NOT WRITE IN THIS AREA
DO NOT WRITE IN THIS AREA

5. A discrete random variable Y has probability function

$$\mathrm{P}(Y = y) = \begin{cases} k(3 - y) & y = 1, 2 \\ k(y^2 - 8) & y = 3, 4, 5 \\ k & y = 6 \\ 0 & \text{otherwise} \end{cases}$$

where k is a constant.

(a) Show that $k = \dfrac{1}{30}$

(2)

Find the exact value of

(b) $\mathrm{P}(1 < Y \leqslant 4)$

(2)

(c) $\mathrm{E}(Y)$

(2)

The random variable $X = 15 - 2Y$

(d) Calculate $\mathrm{P}(Y \geqslant X)$

(3)

(e) Calculate $\mathrm{Var}(X)$

(4)

DO NOT WRITE IN THIS AREA
DO NOT WRITE IN THIS AREA
DO NOT WRITE IN THIS AREA

Question 5 continued

DO NOT WRITE IN THIS AREA
DO NOT WRITE IN THIS AREA
DO NOT WRITE IN THIS AREA

Question 5 continued

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

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 13 marks)

6. Three events A, B and C are such that

$$\mathrm{P}(A) = 0.1 \qquad \mathrm{P}(B \mid A) = 0.3 \qquad \mathrm{P}(A \cup B) = 0.25 \qquad \mathrm{P}(C) = 0.5$$

Given that A and C are mutually exclusive

(a) find $\mathrm{P}(A \cup C)$

(1)

(b) Show that $\mathrm{P}(B) = 0.18$

(3)

Given also that B and C are independent,

(c) draw a Venn diagram to represent the events A, B and C and the probabilities associated with each region.

(5)

DO NOT WRITE IN THIS AREA
DO NOT WRITE IN THIS AREA
DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA
DO NOT WRITE IN THIS AREA
DO NOT WRITE IN THIS AREA

Question 6 continued

Question 6 continued

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

Question 6 continued

(Total for Question 6 is 9 marks)

DO NOT WRITE IN THIS AREA
DO NOT WRITE IN THIS AREA
DO NOT WRITE IN THIS AREA

7. A machine squeezes apples to extract their juice. The volume of juice, J ml, extracted from 1 kg of apples is modelled by a normal distribution with mean μ and standard deviation σ

Given that $\mu = 500$ and $\sigma = 25$ use standardisation to

(a) (i) show that $P(J > 510) = 0.3446$

(2)

(ii) calculate the value of d such that $P(J > d) = 0.9192$

(3)

Zen randomly selects 5 bags each containing 1 kg of apples and records the volume of juice extracted from each bag of apples.

(b) Calculate the probability that each of the 5 bags of apples produce less than 510 ml of juice.

(2)

Following adjustments to the machine, the volume of juice, R ml, extracted from 1 kg of apples is such that $\mu = 520$ and $\sigma = k$

Given that $P(R < r) = 0.15$ and $P(R > 3r - 800) = 0.005$

(c) find the value of r and the value of k

(7)

DO NOT WRITE IN THIS AREA
DO NOT WRITE IN THIS AREA
DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA
DO NOT WRITE IN THIS AREA
DO NOT WRITE IN THIS AREA

Question 7 continued

Question 7 continued

DO NOT WRITE IN THIS AREA
DO NOT WRITE IN THIS AREA
DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA
DO NOT WRITE IN THIS AREA
DO NOT WRITE IN THIS AREA

Question 7 continued

Question 7 continued

(Total for Question 7 is 14 marks)

TOTAL FOR PAPER IS 75 MARKS

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA