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 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 ▶

P72154A

©2022 Pearson Education Ltd.

B:1/1/

P 7 2 1 5 4 A 0 1 2 8

Pearson

1. The stem lengths of a sample of 120 tulips are recorded in the grouped frequency table below.

Stem length (cm)	Frequency
$40 \leqslant x < 42$	12
$42 \leqslant x < 45$	18
$45 \leqslant x < 50$	23
$50 \leqslant x < 55$	35
$55 \leqslant x < 58$	24
$58 \leqslant x < 60$	8

A histogram is drawn to represent these data.

The area of the bar representing the $40 \leqslant x < 42$ class is $16.5\,\text{cm}^2$

(a) Calculate the exact area of the bar representing the $42 \leqslant x < 45$ class.

(2)

The height of the tallest bar in the histogram is 10 cm.

(b) Find the exact height of the second tallest bar.

(3)

Q_1 for these data is 45 cm.

(c) Use linear interpolation to find an estimate for

(i) Q_2

(ii) the interquartile range.

(4)

One measure of skewness is given by

$$\frac{Q_3 - 2Q_2 + Q_1}{Q_3 - Q_1}$$

(d) By calculating this measure, describe the skewness of these data.

(2)

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 11 marks)

2. The production cost, £c million, of a film and the total ticket sales, £t million, earned by the film are recorded for a sample of 40 films.

Some summary statistics are given below.

$$\sum c = 1634 \qquad \sum t = 1361 \qquad \sum t^2 = 82\,873 \qquad \sum ct = 83\,634 \qquad S_{cc} = 28\,732.1$$

(a) Find the exact value of S_{tt} and the exact value of S_{ct}

(3)

(b) Calculate the value of the product moment correlation coefficient for these data.

(2)

(c) Give an interpretation of your answer to part (b)

(1)

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

$$t = -5.84 + 0.976c$$

where the values of the intercept and gradient are given to 3 significant figures.

(3)

(e) Find the expected total ticket sales for a film with a production cost of £90 million.

(2)

Using the regression line in part (d)

(f) find the range of values of the production cost of a film for which the total ticket sales are less than 80% of its production cost.

(2)

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. Morgan is investigating the body length, b centimetres, of squirrels.

A random sample of 8 squirrels is taken and the data for each squirrel is coded using

$$x = \frac{b - 21}{2}$$

The results for the coded data are summarised below

$$\sum x = -1.2 \qquad \sum x^2 = 5.1$$

(a) Find the mean of b

(3)

(b) Find the standard deviation of b

(3)

A 9th squirrel is added to the sample. Given that for all 9 squirrels $\sum x = 0$

(c) find

(i) the body length of the 9th squirrel,

(2)

(ii) the standard deviation of x for all 9 squirrels.

(2)

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 10 marks)

4. The cumulative distribution function of the discrete random variable W, which takes only the values 6, 7 and 8, is given by

$$F(W) = \frac{(w + 3)(w - 1)}{77} \quad \text{for } w = 6, 7, 8$$

Find $E(W)$

(4)

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 4 continued

(Total for Question 4 is 4 marks)

5. The weights, W grams, of kiwi fruit grown on a farm are normally distributed with mean 80 grams and standard deviation 8 grams.

The table shows the classifications of the kiwi fruit by their weight, where k is a positive constant.

Small		**Large**		
Tiny	**Petite**	**Extra**	**Jumbo**	**Mega**
$w < 66$	$66 \leqslant w < 70$	$70 \leqslant w < 80$	$80 \leqslant w < k$	$w \geqslant k$

One kiwi fruit is selected at random from those grown on the farm.

(a) Find the probability that this kiwi fruit is Large.

(3)

35% of the kiwi fruit are Jumbo.

(b) Find the value of k to one decimal place.

(4)

75% of Tiny kiwi fruit weigh more than y grams.

(c) Find the value of y giving your answer to one decimal place.

(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 5 continued

Question 5 continued

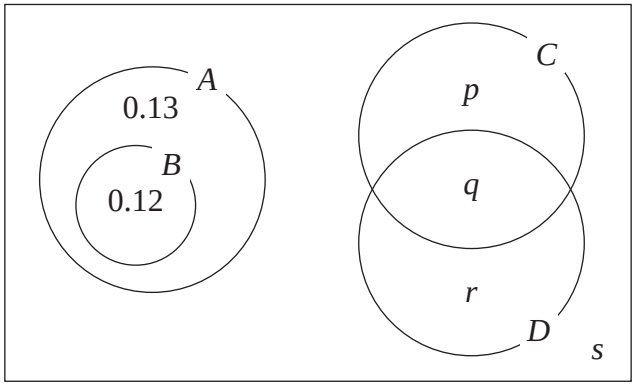
DO NOT WRITE IN THIS AREA
DO NOT WRITE IN THIS AREA
DO NOT WRITE IN THIS AREA

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 12 marks)

6. The Venn diagram shows the events A, B, C and D, where p, q, r and s are probabilities.



(a) Write down the value of

(i) $P(A)$

(ii) $P(A \mid B)$

(iii) $P(A \mid C)$

(3)

Given that $P(B' \cap D') = \frac{7}{10}$ and $P(C \mid D) = \frac{3}{5}$

(b) find the exact value of q and the exact value of r

(6)

Given also that $P(B \cup C') = \frac{5}{8}$

(c) find the exact value of s

(2)

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

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 11 marks)

7. Adana selects one number at random from the distribution of X which has the following probability distribution.

x	0	5	10
$P(X = x)$	0.1	0.2	0.7

(a) Given that the number selected by Adana is not 5, write down the probability it is 0

(1)

(b) Show that $E(X^2) = 75$

(1)

(c) Find $\text{Var}(X)$

(3)

(d) Find $\text{Var}(4 - 3X)$

(2)

Bruno and Charlie each independently select one number at random from the distribution of X

(e) Find the probability that the number Bruno selects is greater than the number Charlie selects.

(3)

Devika multiplies Bruno's number by Charlie's number to obtain a product, D

(f) Determine the probability distribution of D

(4)

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: 75 MARKS

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