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 ▶

P71284A

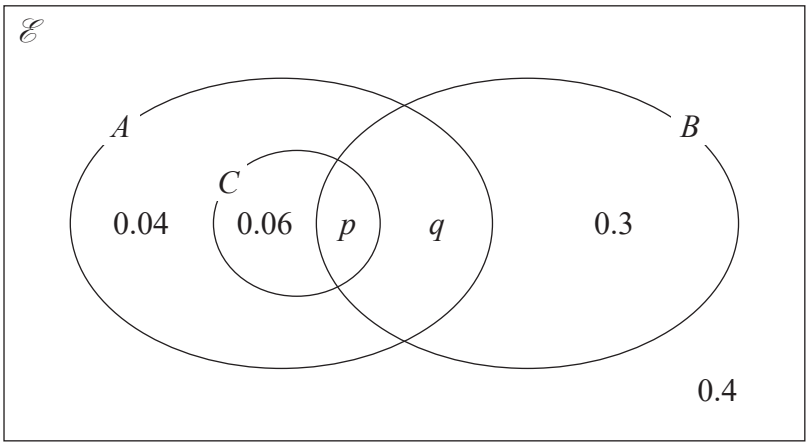
E:1/1/1/

P 7 1 2 8 4 A 0 1 2 4

Pearson

Leave blank

1. The Venn diagram shows the events A, B and C and their associated probabilities, where p and q are probabilities.



(a) Find $\mathrm{P}(B)$

(1)

(b) Determine whether or not A and B are independent.

(2)

Given that $\mathrm{P}(C \mid B) = \mathrm{P}(C)$

(c) find the value of p and the value of q

(3)

The event D is such that

- A and D are mutually exclusive
- $\mathrm{P}(B \cap D) > 0$

(d) On the Venn diagram show a possible position for the event D

(1)

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

Leave blank

Question 1 continued

Q1

(Total 7 marks)

Leave blank

2. A large company is analysing how much money it spends on paper in its offices each year. The number of employees in the office, x, and the amount spent on paper in a year, p (\$ hundreds), in each of 12 randomly selected offices were recorded.

The results are summarised in the following statistics.

$$\sum x = 93 \qquad S_{xx} = 148.25 \qquad \sum p = 273 \qquad \sum p^2 = 6602.72 \qquad \sum xp = 2347$$

(a) Show that $S_{xp} = 231.25$

(1)

(b) Find the product moment correlation coefficient for these data.

(3)

(c) Find the equation of the regression line of p on x in the form $p = a + bx$

(4)

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

(1)

The director of the company wants to reduce the amount spent on paper each year.

He wants each office to aim for a model of the form $p = \frac{4}{5}a + \frac{1}{2}bx$, where a and b are the values found in part (c).

Using the data for the 93 employees from the 12 offices,

(e) estimate the percentage saving in the amount spent on paper each year by the company using the director's model.

(3)

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

Leave blank

Question 2 continued

Leave blank

Question 2 continued

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

Leave blank

Question 2 continued

(Total 12 marks)

Q2

Leave blank

3. The stem and leaf diagram shows the ages of the 35 male passengers on a cruise.

Age **Key:** 1 | 3 represents an age of 13 years

1	3	(1)
2	7 9	(2)
3	1 2 8 8	(4)
4	5 5 6 7 8 8 9	(7)
5	2 2 3 3 4 4 5 6 6 8	(10)
6	0 1 1 4 4 4 7	(7)
7	3 6	(2)
8	7 8	(2)

(a) Find the median age of the male passengers.

(1)

(b) Show that the interquartile range (IQR) of these ages is 16

(2)

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

(c) Show that there are 3 outliers amongst these ages.

(3)

(d) On the grid in Figure 1 on page 9, draw a box plot for the ages of the male passengers on the cruise.

(4)

Figure 1 on page 9 also shows a box plot for the ages of the female passengers on the cruise.

(e) Comment on any difference in the distributions of ages of male and female passengers on the cruise.
State the values of any statistics you have used to support your comment.

(1)

Anja, along with her 2 daughters and a granddaughter, now join the cruise.
Anja's granddaughter is younger than both of Anja's daughters.
Anja had her 23rd birthday on the day her eldest daughter was born.
When their 4 ages are included with the other female passengers on the cruise, the box plot does not change.

(f) State, giving reasons, what you can say about
(i) the granddaughter's age
(ii) Anja's age.

(3)

Leave blank

Question 3 continued

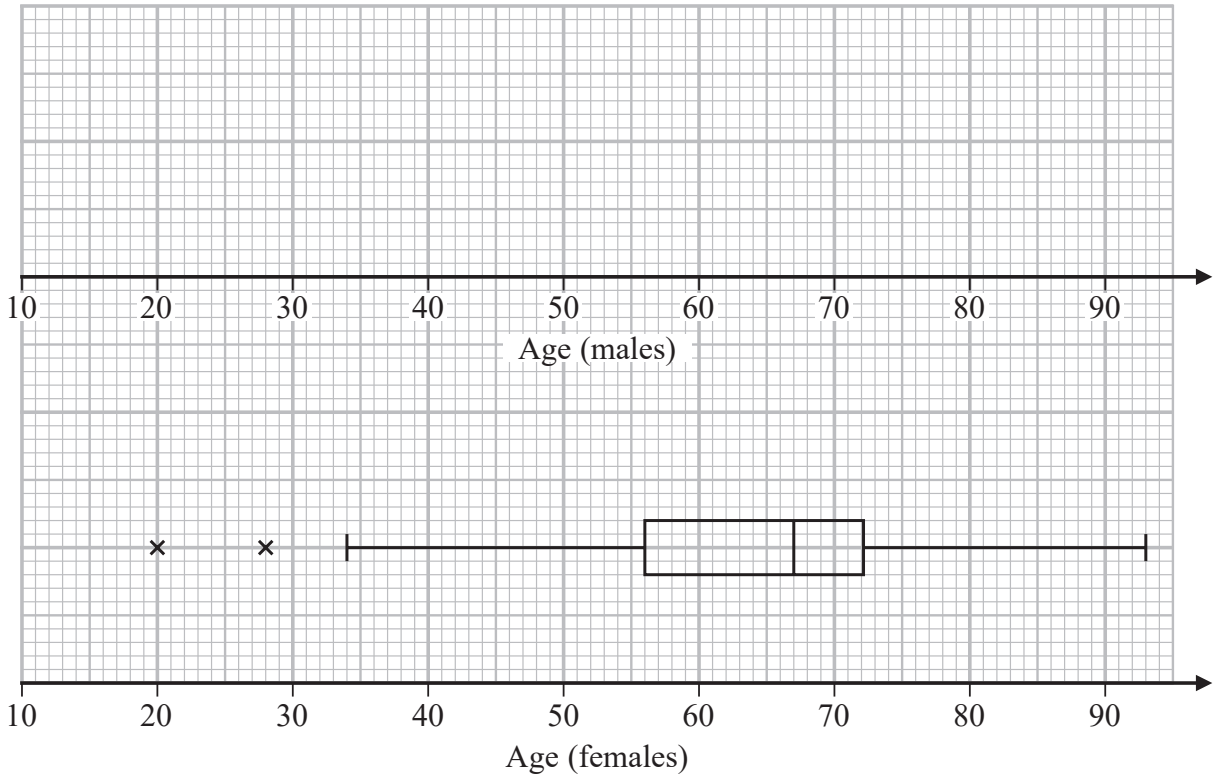


Figure 1

Turn over for a spare grid if you need to redraw your box plot.

Leave blank

Question 3 continued

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

Leave blank

Question 3 continued

Only use this grid if you need to redraw your box plot.

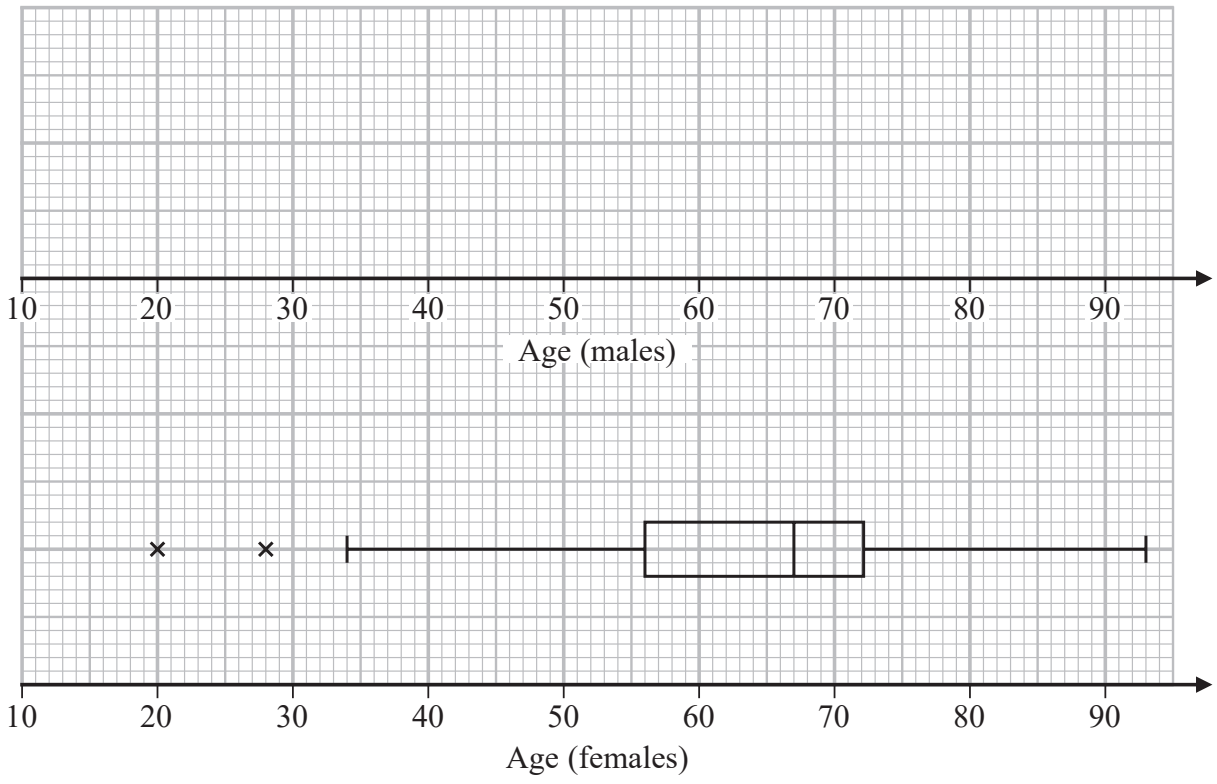


Figure 1

(Total 14 marks)

Q3

Leave blank

4. Three bags **A**, **B** and **C** each contain coloured balls.

Bag **A** contains 4 red balls and 2 yellow balls only.
Bag **B** contains 4 red balls and 1 yellow ball only.
Bag **C** contains 6 red balls only.

In a game
Mike takes a ball at random from bag **A**, records the colour and places it in bag **C**.
He then takes a ball at random from bag **B**, records the colour and places it in bag **C**.
Finally, Mike takes a ball at random from bag **C** and records the colour.

(a) Complete the tree diagram on the page opposite, to illustrate the game by adding the remaining branches and all probabilities.

(3)

(b) Show that the probability that Mike records a yellow ball exactly twice is $\frac{1}{10}$

(3)

Given that Mike records exactly 2 yellow balls,

(c) find the probability that the ball drawn from bag **A** is red.

(2)

Mike plays this game a large number of times, each time starting with the bags containing balls as described above. The random variable X represents the number of yellow balls recorded in a single game.

(d) Find the probability distribution of X

(3)

(e) Find $\mathrm{E}(X)$

(2)

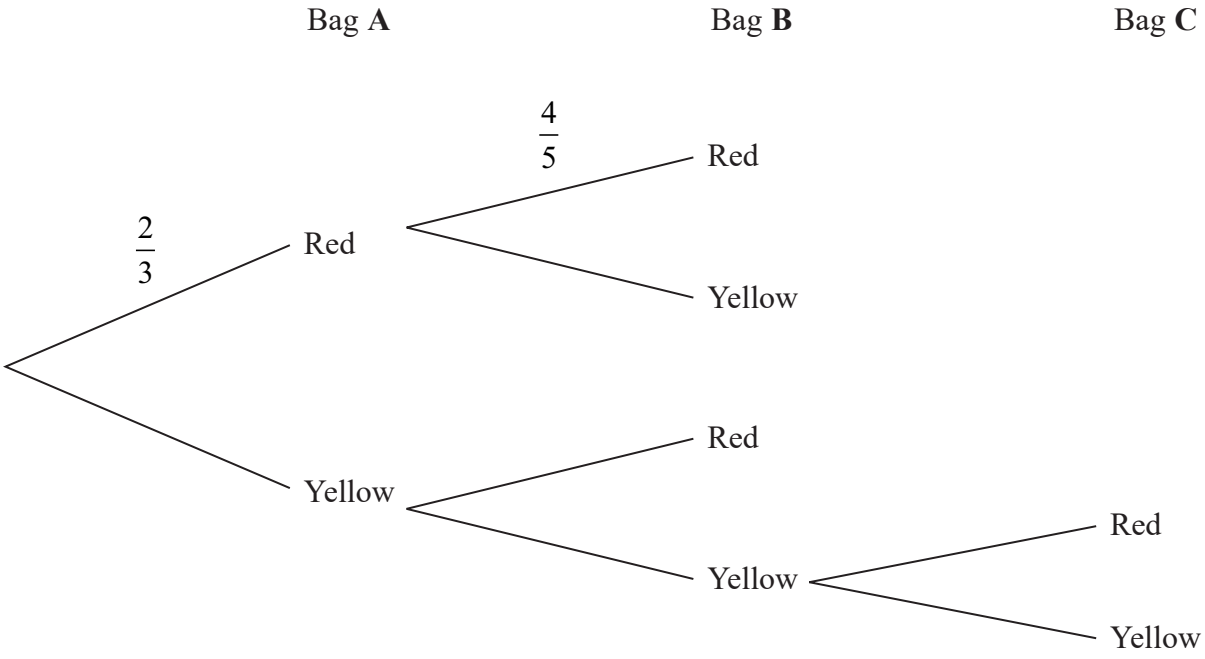
DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

Leave blank

Question 4 continued



Leave blank

Question 4 continued

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

Leave blank

Question 4 continued

Q4

(Total 13 marks)

Leave blank

5. The discrete random variable Y has the following probability distribution

y	-9	-5	0	5	9
$P(Y = y)$	q	r	u	r	q

where q, r and u are probabilities.

(a) Write down the value of $E(Y)$

(1)

The cumulative distribution function of Y is $F(y)$

Given that $F(0) = \frac{19}{30}$

(b) show that the value of u is $\frac{4}{15}$

(3)

Given also that $\text{Var}(Y) = 37$

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

(4)

The coordinates of a point P are $(12, Y)$

The random variable D represents the length of OP

(d) Find the probability distribution of D

(6)

Leave blank

Question 5 continued

Leave blank

Question 5 continued

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

Leave blank

Question 5 continued

(Total 14 marks)

Q5

Leave blank

6. Xiang is designing shelves for a bookshop. The height, H cm, of books is modelled by the normal distribution with mean 25.1 cm and standard deviation 5.5 cm

(a) Show that $P(H > 30.8) = 0.15$

(3)

Xiang decided that the smallest 5% of books and books taller than 30.8 cm would not be placed on the shelves. All the other books will be placed on the shelves.

(b) Find the range of heights of books that will be placed on the shelves.

(3)

The books that will be placed on the shelves have heights classified as small, medium or large.
The numbers of small, medium and large books are in the ratios 2 : 3 : 3

(c) The medium books have heights x cm where $m < x < d$

(i) Show that $d = 25.8$ to 1 decimal place.

(3)

(ii) Find the value of m

(4)

Xiang wants 2 shelves for small books, 3 shelves for medium books and 3 shelves for large books.
These shelves will be placed one above another and made of wood that is 1 cm thick.

(d) Work out the minimum total height needed.

(2)

Leave blank

Question 6 continued

Leave blank

Question 6 continued

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

Leave blank

Question 6 continued

Leave blank

Question 6 continued

Q6

(Total 15 marks)

TOTAL FOR PAPER: 75 MARKS

END

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA