Pearson Edexcel International Advanced Level

Tuesday 19 May 2026

Morning (Time: 1 hour 30 minutes)

Paper reference **WST01/01A**

Mathematics

International Advanced Subsidiary/Advanced Level
Statistics S1
Question paper

You must have:
Answer book (sent separately).
Do not return this question paper with the answer book.

Turn over

P81356A

P:1/1/1/

P 8 1 3 5 6 A

Pearson

1. Jeff plays a game which consists of throwing a bean bag into a bucket.

Jeff has a maximum of three attempts to get the bean bag into the bucket.

If he gets the bean bag into the bucket, he wins the game and stops throwing.

- On his first attempt, the probability that Jeff gets the bean bag into the bucket is p
- On his second attempt, the probability that Jeff gets the bean bag into the bucket is 0.35
- On his third attempt, the probability that Jeff gets the bean bag into the bucket is 0.5

(a) Draw a tree diagram to represent this information.
Label the probabilities, in terms of p where necessary.

(2)

The probability that Jeff wins the game is 0.74

(b) Find the value of p

(3)

(c) Find the probability that Jeff does **not** win the game.

(1)

(d) Given that Jeff wins the game, find the probability that he does this on his second attempt.

(2)

(Total for Question 1 is 8 marks)

2. The ages, in years, of a sample of 23 members of a tennis club are summarised by the stem and leaf diagram below.

Age		Totals
1	6 7 8 8	(4)
2	2 2 3 3 5 5 6 7	(8)
3	3 4 4 8 9 9	(6)
4	1 2 5	(3)
5	5	(1)
6	6	(1)

Key: 3 | 8 represents 38 years

(a) Find

(i) the median age of these members,

(ii) the interquartile range of the ages of these members.

(3)

For this dataset an outlier is defined as an age that is

either more than 1.5 × interquartile range above the upper quartile

or more than 1.5 × interquartile range below the lower quartile

(b) On the grid provided on page 7 of the answer book draw a box plot to represent the ages of these members, indicating clearly any outliers.

(5)

A record was kept of the time, x hours, spent by these 23 members at the club in a year.

The results are summarised by

$$\sum(x-200)=115 \qquad \sum(x-200)^2=4623$$

(c) Calculate

(i) the mean of x

(ii) the standard deviation of x

(5)

(Total for Question 2 is 13 marks)

3. The discrete random variable X has the following probability distribution,

x	-4	-2	0	2	6
$P(X = x)$	r	s	t	s	r

where r, s and t are probabilities.

Given that $E(X) = 0.36$

(a) show that $r = 0.18$

(2)

Given also that $\text{Var}(X) = 10.8304$

(b) find the value of s **and** the value of t

(5)

(c) Find

(i) $E(7 - 5X)$

(ii) $\text{Var}(7 - 5X)$

(3)

(d) Find $P(2X^2 > 7 - 5X)$

(4)

(Total for Question 3 is 14 marks)

4. The table shows information about the time, t minutes, taken by 250 people to complete a 10 km race.

Time (t)	**Frequency (f)**
$33 \leqslant t < 35$	3
$35 \leqslant t < 39$	16
$39 \leqslant t < 43$	49
$43 \leqslant t < 45$	40
$45 \leqslant t < 47$	50
$47 \leqslant t < 50$	92

A histogram is drawn (not shown) to represent these data.
The bar representing the class $35 \leqslant t < 39$ is 6 cm wide and 1.12 cm high.

(a) Calculate the width and the height of the bar representing the class $43 \leqslant t < 45$

(3)

(b) Use linear interpolation to estimate the median of t

(2)

For these data $\sum \mathrm{f}t = 11\,225$ and $\sum \mathrm{f}t^2 = 507\,388$

(c) Calculate an estimate of

(i) the mean of t

(ii) the standard deviation of t

(3)

Two of these 250 people are chosen at random.

(d) Estimate the probability that **both** of their times are between 38 minutes and 44 minutes.

(3)

It was found that the fastest time had been incorrectly recorded. The actual finishing time of the fastest person was 32.4 minutes.

(e) State what effect, if any, this would have on

(i) the median of t

(ii) the interquartile range of t

(2)

(Total for Question 4 is 13 marks)

5. The events A and B are independent.
Given that

- $\mathrm{P}(A) = \mathrm{P}(B) = x$
- $\mathrm{P}(A \cup B) = \dfrac{9}{25}$

(a) find the value of x

(2)

The events C and D are such that

- $\mathrm{P}(C) = \dfrac{11}{20}$
- $\mathrm{P}(C \mid D) = \dfrac{6}{13}$
- $\mathrm{P}(C \cup D) = 3\mathrm{P}(C \cap D)$

(b) Find

(i) $\mathrm{P}(D)$

(4)

(ii) $\mathrm{P}(D \mid C')$

(3)

(Total for Question 5 is 9 marks)

6. The weights of newborn babies are normally distributed with mean 3200 grams and standard deviation 200 grams.

(a) Find, using standardisation, the probability that a randomly selected newborn baby has a weight greater than 3500 grams.

(2)

A newborn baby whose weight is greater than 3500 grams is classified as high.

A newborn baby is chosen at random.

Given that it is classified as high,

(b) find, using standardisation, the probability that this newborn baby has a weight greater than 3600 grams.

(4)

Given that 75.6% of newborn babies classified as high have a weight greater than w grams,

(c) find, using standardisation, the value of w

(4)

(Total for Question 6 is 10 marks)

7. Based on 10 bivariate observations of the variables x and y, the following summary statistics were calculated.

	x	y
Mean	10.45	11.36
Standard deviation	9.3	9.0

The least squares regression line of $\boldsymbol{x}$ **on** $\boldsymbol{y}$ has gradient 0.8

(a) Find the exact value of $\sum xy$

(4)

(b) Find the equation of the least squares regression line of $\boldsymbol{y}$ **on** $\boldsymbol{x}$, giving your answer in the form $y = a + bx$

(4)

(Total for Question 7 is 8 marks)

TOTAL FOR PAPER IS 75 MARKS