Please check the examination details below before entering your candidate information

Candidate surname	Other names

Centre Number | Candidate Number

Pearson Edexcel International Advanced Level

Wednesday 8 January 2025

Afternoon (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 ▶

P76199A

©2025 Pearson Education Ltd.
H:1/1/1/

P 7 6 1 9 9 A 0 1 2 4

Pearson

1. Jen has one fair 4-sided red die and one fair 4-sided blue die.

- The red die has sides numbered 1, 2, 3 and 4
- The blue die has sides numbered 1, 3, 5 and 7

The discrete random variable R represents the score from one roll of the red die.

The discrete random variable B represents the score from one roll of the blue die.

(a) Write down the name of the distribution of R

(1)

(b) Find $P(R < 3)$

(1)

(c) Write down the value of

(i) $E(R)$

(ii) $E(B)$

(2)

(d) Showing your working, find $Var(B)$

(3)

Jen rolls each die once.

(e) Find $P(R + B \leqslant 5)$

(2)

(f) Find $P(R^2 < B)$

(3)

The random variable D is defined as the magnitude of the difference between the score on the red die and the score on the blue die.

The table below shows the cumulative distribution function of D

d	0	1	2	3	4	5	6
$F(d)$	$\frac{1}{8}$	$\frac{3}{8}$	$\frac{9}{16}$	$\frac{3}{4}$	p	$\frac{15}{16}$	1

(g) Showing your working, find the value of p

(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 15 marks)

2. As part of an investigation, Bobby collects a sample of 47 observations, x

The results are shown in the following stem and leaf diagram, where a is a constant.

Stem	Leaf										
2	1	2	5	7	7						(5)
3	0	2	2	3	4	5	5	5	9	9	(10)
4	0	0	1	4	4	5	7	8	8	9	9 (11)
5	3	3	5	6	7	9					(6)
6	0	2	a	a	a	7	8				(7)
7	1	2	3	6	8						(5)
8	0	6	7								(3)

Key: 3 | 2 means 0.32

(a) Find the range of these observations.

(1)

(b) Find the value of the median of these observations.

(1)

Given that the interquartile range of these observations is 0.31

(c) find the value of a

(3)

Bobby calculates the following statistics from these observations

$$\sum x = 23.72 \qquad \sum x^2 = 13.4228$$

(d) Show that the standard deviation of these observations is 0.176 to 3 significant figures.

(2)

Bobby now collects 18 more observations, y, from the same investigation.

(e) Using all 65 observations, the sample mean is 0.502 and the sample standard deviation is 0.204

(i) Show that $\sum y = 8.91$

(2)

(ii) Showing your working, calculate $\sum y^2$

(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 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

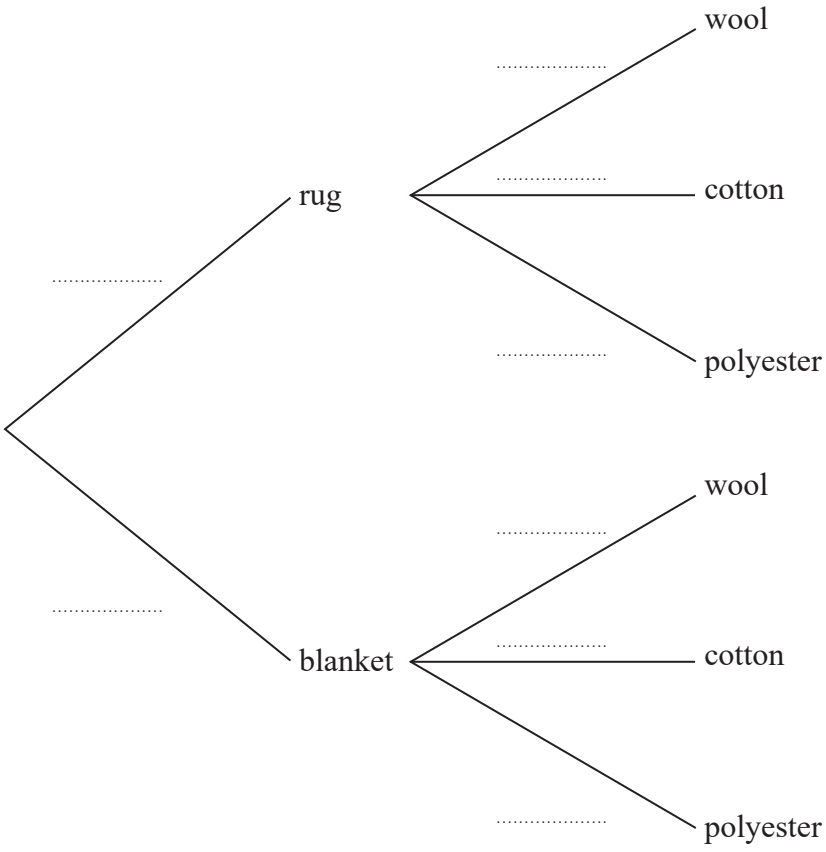
(Total for Question 2 is 12 marks)

3. A factory makes rugs and blankets from wool **or** cotton **or** polyester.

- 30% of the items made are rugs and 70% of the items made are blankets
- 60% of the rugs are made from wool
- 35% of the blankets are made from wool
- 10% of the rugs are made from cotton
- 20% of the blankets are made from cotton
- the rest of the rugs and blankets are made from polyester

An item made by the factory is selected at random.

(a) Complete the tree diagram below to illustrate this information.



(2)

(b) (i) Find the probability that the item selected is **not** made from wool.

(2)

(ii) Given that this item is **not** made from wool, find the probability that it is a blanket.

(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

(Total for Question 3 is 6 marks)

4. A manufacturer of pet food gives samples of three new flavours of cat food, A, B and C, to each of 100 customers who each own one cat. These customers are surveyed to find out whether or not their cat likes each flavour with the following results

40 cats like A

45 cats like B

50 cats like C

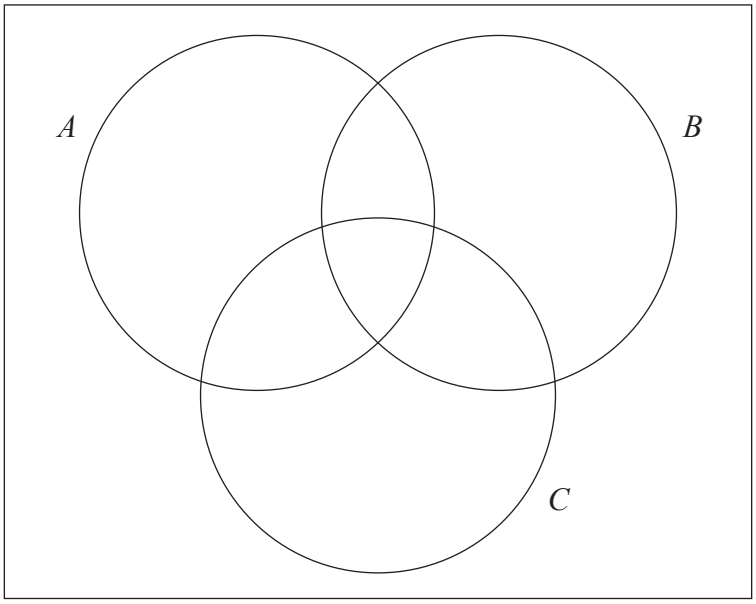
19 cats like both A and B

20 cats like both A and C

22 cats like both B and C

12 cats like all three of these flavours.

(a) Using these results complete the Venn diagram below.



(4)

For these cats,

(b) determine whether or not a cat liking flavour A is independent of it liking flavour C

(2)

(c) One of these 100 cats is chosen at random.

(i) Find the probability that it likes only one of the three flavours.

(1)

(ii) Given that it likes flavour C, find the probability that it also likes flavour B

(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 4 continued

(Total for Question 4 is 9 marks)

5. A recycling centre measures the weight of glass deposited by the public each day.

The weight of glass, S kg, deposited at the recycling centre in a day during the summer can be modelled by $S \sim \mathrm{N}(700, 50^2)$

(a) Using standardisation and showing your working, find the probability that, in one randomly selected day during the summer,

(i) more than 640 kg of glass is deposited at the recycling centre,

(2)

(ii) 700 kg of glass, correct to the nearest 50 kg, is deposited at the recycling centre.

(5)

The weight of glass, W kg, deposited at the recycling centre in a day during the winter can be modelled by $W \sim \mathrm{N}(\mu, \sigma^2)$

(b) Given that $\mathrm{P}(W > 680) = 0.0668$ and $\mathrm{P}(W < 599) = 0.3$

(i) find **two** equations in terms of μ and σ

(3)

(ii) Hence, showing your working, find the value of μ and the value of σ

(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 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 13 marks)

6. As a part of a selection process, applicants for a television game show have to take a test, and then complete a task as quickly as possible. The test score, w, and the time taken to complete the task, t minutes, are recorded for each applicant.

The summary statistics below represent the data for a random sample of 30 applicants.

$$S_{wt} = -1648.83 \qquad S_{ww} = 2396.97 \qquad \sum w = 839 \qquad \sum t = 635 \qquad \sum t^2 = 14\,837$$

(a) Show that the product moment correlation coefficient for these data is −0.901 to 3 significant figures.

(2)

A scatter diagram of t against w is plotted for these data.

(b) State **two** features of the graph that you would expect to see, given the correlation coefficient in part (a)

(2)

(c) Calculate the equation of the regression line of t on w in the form

$$t = a + bw$$

Give the values of the constants a and b to 3 significant figures.

(4)

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

(1)

The test score, w, was a score out of 50
The manager of the selection process now decides to double all the values of w to make them into percentages.
The manager then recalculates the product moment correlation coefficient and the equation of the regression line.

(e) State, for **each** of the following, whether the value would increase **or** decrease **or** stay the same as a result of applying this change.

(i) The product moment correlation coefficient

(ii) The magnitude of the gradient of the regression line

(iii) The t intercept of the regression line

(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 6 continued

Question 6 continued

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

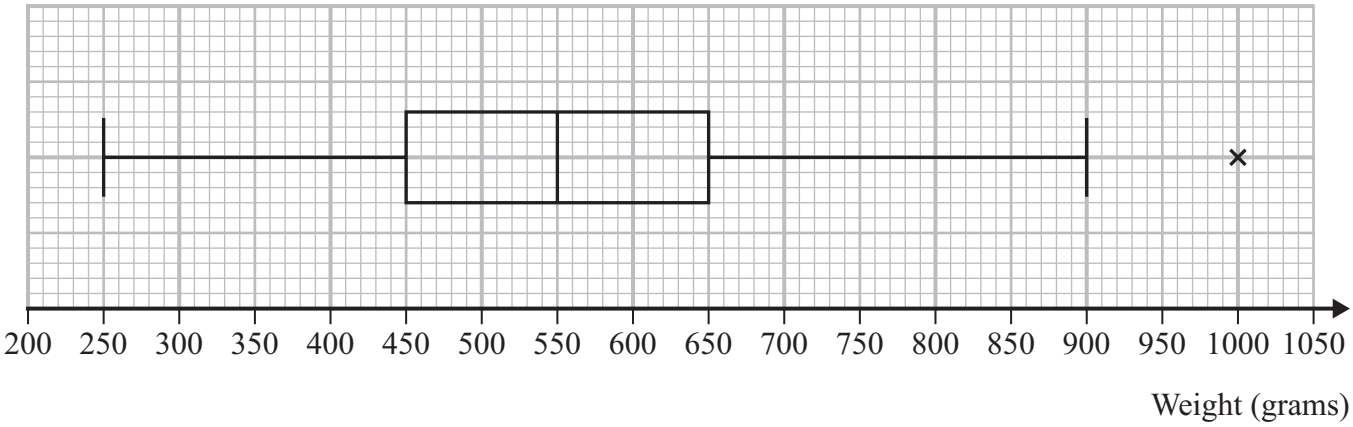
Question 6 continued

(Total for Question 6 is 12 marks)

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

7. A farmer supplies cabbages to a local supermarket. The supermarket manager takes a random sample of 200 of these cabbages and the weight of each cabbage is recorded.

The weights of the cabbages in the sample are summarised in the following box plot.



(a) Use linear interpolation to estimate the probability that a randomly chosen cabbage from this sample has a weight of less than 570 grams.

(2)

The supermarket manager takes the quartiles from this sample and now decides to classify any cabbage whose weight is at least $Q_3 + 1.5 \times (Q_3 - Q_1)$ as extra-large.

(b) Find the minimum weight of a cabbage that the supermarket manager would classify as extra-large.

(1)

The supermarket manager assumes that the weights of cabbages supplied to the supermarket may be modelled by a normal distribution.

(c) Explain whether or not the box plot supports this assumption.

(1)

The farmer records the weight of every cabbage harvested from a field and classifies any cabbage whose weight is more than 2 standard deviations above the mean as an outlier.

Assuming that the weights of the cabbages from this field may be modelled by a normal distribution,

(d) find the probability that a randomly selected cabbage harvested from the field would be classified as an outlier using the farmer's method.

(2)

The mean weight of the cabbages harvested from the field is 560 grams.

Given that a cabbage with a weight of 1000 grams is classified as an outlier by the farmer,

(e) find the maximum possible value of the standard deviation of the weights of cabbages harvested from the field.

(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 7 continued

Question 7 continued

(Total for Question 7 is 8 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