Please check the examination details below before entering your candidate information

Candidate surname StemUp Model Solution | Other names

Centre Number | Candidate Number

Pearson Edexcel International Advanced Level

Tuesday 20 May 2025

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 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 ▶

P76176A

©2025 Pearson Education Ltd.
Y:1/1/1/1/1/

P 7 6 1 7 6 A 0 1 2 4

Pearson

1: The discrete random variable X has probability distribution

x	1	2	3	4
$\mathbf{P}(X = x)$	0.3	0.2	0.35	0.15

(a) Find F(3.5)

(1)

(b) Find $P(3X - 3 < X + 2)$

(2)

(c) Find $E(X)$

(2)

(d) Show that $\text{Var}(X) = 1.1275$

(3)

(e) Find $\text{Var}(5 - 2X)$

(1)

(a) $F(3.5) = P(X \leq 3) = 0.3 + 0.2 + 0.35 = 0.85$

(b) $3X - 3 < X + 2 \Rightarrow 2X < 5 \Rightarrow X < 2.5$

$P(X < 2.5) = P(X = 1) + P(X = 2) = 0.3 + 0.2 = 0.5$

(c) $E(X) = 1 \times 0.3 + 2 \times 0.2 + 3 \times 0.35 + 4 \times 0.15$

$= 0.3 + 0.4 + 1.05 + 0.6 = 2.35$

(d) $E(X^2) = 1^2 \times 0.3 + 2^2 \times 0.2 + 3^2 \times 0.35 + 4^2 \times 0.15$

$= 0.3 + 0.8 + 3.15 + 2.4 = 6.65$

$\text{Var}(X) = E(X^2) - [E(X)]^2 = 6.65 - 2.35^2$

$= 6.65 - 5.5225 = 1.1275$ as required

(e) $\text{Var}(5 - 2X) = (-2)^2 \text{Var}(X) = 4 \times 1.1275 = 4.51$

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 9 marks)

stemup.co.uk

2: Students in class A and class B sit a statistics test.

There are 24 students in class A and 30 students in class B

The marks, x, for students in each class are summarised in the table below.

	n	$\bar{x}$	$\sum x^2$
Class A	24	47	66 876
Class B	30	b	73 826

The two classes are combined into one group of 54 students.
The mean mark for all 54 students is 45

(a) Show that $b = 43.4$

(2)

(b) Find the standard deviation of the marks for all 54 students.

(2)

Following moderation, each student in class B has their mark increased by 2

(c) Without further calculations state, giving a reason in each case, the effect this will have on

(i) the variance of the marks for class B

(ii) the mean mark for all 54 students

(iii) the standard deviation of the marks for all 54 students.

(3)

(a) Total for all 54 students = $54 \times 45 = 2430$

Total for class A = $24 \times 47 = 1128$

Total for class B = $30b = 2430 - 1128 = 1302$

$b = \frac{1302}{30} = 43.4$ as required

(b) $\Sigma x^2 = 66876 + 73826 = 140702$

$\sigma^2 = \frac{140702}{54} - 45^2 = 2605.592... - 2025 = 580.592...$

$\sigma = \sqrt{580.592...} = 24.0955... = 24.1$ (3 s.f.)

(c)(i) No effect: adding 2 to every mark in class B does not change the spread of the marks, so the variance stays the same.

(ii) Increase: the total of the marks for class B increases (by 60),

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

Question 2 continued

so the total for all S4 students increases and so does the mean.

(iii) Decrease: the mean mark of class B is now closer to the mean mark of class A, so the marks are closer together (less spread).

(Total for Question 2 is 7 marks)

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

3: Statistical models are a cheap and quick way to make predictions about real-world situations.

(a) Give one other reason why statistical models are used.

(1)

Madison wants to develop a model to describe the relationship between the average daily temperature, t °C, and a household's daily gas consumption, d m^3, in winter.

Madison takes a random sample of 12 days in winter and codes the daily gas consumption so that $w = 11.5d$

These data are summarised as follows

$$S_{tt} = 26.43 \quad S_{tw} = -91.55 \quad \sum w = 339.25 \quad \sum t = 9.1 \quad \sum w^2 = 10\,036.45$$

(b) Show that $S_{ww} = 445.57$ to 2 decimal places.

(1)

(c) Find the value of S_{td} and the value of S_{dd}

(3)

(d) Find the product moment correlation coefficient between d and t

(2)

(e) Give an interpretation, in context, of your product moment correlation coefficient.

(1)

(f) Show that the equation of the regression line of w on t is

$$w = -3.46t + 30.9$$

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

(3)

(g) Write down an equation of the regression line of d on t

(1)

(h) Using your equation in part (g)

(i) estimate the daily gas consumption in winter when the temperature is 2 °C

(ii) interpret the effect an increase of 1 °C in average daily temperature is expected to have on the daily gas consumption in winter.

(2)

(a) To improve understanding of the real-world situation, e.g. to show the relationship between the variables.

(b) $S_{ww} = 10036.45 - \frac{339.25^2}{12} = 10036.45 - 9590.88...$

$= 445.5698... = 445.57$ (2 d.p.) as required

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

stemup.co.uk

Question 3 continued

(c) $w = 11.5d \Rightarrow d = w \div 11.5$

$S_{td} = \frac{S_{tw}}{11.5} = \frac{-91.55}{11.5} = -7.9608... = -7.96$

$S_{dd} = \frac{S_{ww}}{11.5^2} = \frac{445.57}{132.25} = 3.3691... = 3.37$

(d) $r = \frac{S_{tw}}{\sqrt{(S_{tt}\, S_{ww})}} = \frac{-91.55}{\sqrt{(26.43 \times 445.57)}} = -0.84363... = -0.844$

(coding $w = 11.5d$ does not change r, so r for d and t is the same)

(e) As the average daily temperature increases, the daily gas consumption decreases.

(f) $b = \frac{S_{tw}}{S_{tt}} = \frac{-91.55}{26.43} = -3.46387...$

$a = \frac{339.25}{12} - (-3.46387...) \times \frac{9.1}{12} = 28.2708... + 2.62677... = 30.8976...$

$w = -3.46t + 30.9$ (3 s.f.) as required

(g) $11.5d = -3.46t + 30.9 \Rightarrow d = -0.301t + 2.69$

(h)(i) $t = 2$: $d = -0.30087 \times 2 + 2.68696 = 2.0852...$

Daily gas consumption $\approx 2.09\ m^3$

(ii) For each 1 °C increase in average daily temperature, the daily gas consumption decreases by $0.301\ m^3$ (about $0.3\ m^3$).

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

Question 3 continued

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

stemup.co.uk

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 14 marks)

stemup.co.uk

4: Two events C and D are such that

$$P(C \cup D) = 0.59 \qquad P(D) = 0.45 \qquad P(C \mid D) = 0.2$$

Find the value of

(a) $P(C \cap D)$

(2)

(b) $P(C)$

(2)

(a) $P(C \mid D) = \frac{P(C \cap D)}{P(D)} \Rightarrow 0.2 = \frac{P(C \cap D)}{0.45}$

$P(C \cap D) = 0.2 \times 0.45 = 0.09$

(b) $P(C \cup D) = P(C) + P(D) - P(C \cap D)$

$0.59 = P(C) + 0.45 - 0.09$

$P(C) = 0.59 - 0.45 + 0.09 = 0.23$

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

stemup.co.uk

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)

stemup.co.uk

5: The following grouped frequency distribution summarises the speeds in km/h that a random sample of 225 cars were doing on a road.

Speed (x km/h)	Number of cars
$20 \leqslant x < 25$	54
$25 \leqslant x < 30$	90
$30 \leqslant x < 40$	60
$40 \leqslant x < 55$	15
$55 \leqslant x < 70$	6

A histogram is drawn to represent these data.

The height of the tallest bar is 10 cm.

(a) Calculate the height of the second tallest bar.

(3)

(b) Estimate how many of the 225 cars were doing a speed between 35 km/h and 57 km/h. You must show your working clearly.

(2)

(c) Use linear interpolation to estimate the median speed in km/h for the 225 cars.

(2)

The lower quartile for these data is 25.14 km/h to 4 significant figures.

(d) Find an estimate for

(i) the upper quartile,

(ii) the interquartile range.

(4)

(e) Describe, giving a reason, the skewness of these data.

(2)

Ali suggests that a normal distribution is a suitable model for the speed that cars were doing on this road.

(f) With reference to your answer in part (e), comment on Ali's suggestion, giving a reason for your answer.

(1)

(a) Frequency density = frequency ÷ class width:

20-25: 54 ÷ 5 = 10.8 25-30: 90 ÷ 5 = 18 30-40: 60 ÷ 10 = 6

40-55: 15 ÷ 15 = 1 55-70: 6 ÷ 15 = 0.4

Tallest bar: 25 ≤ x < 30 (f.d. 18); 2nd tallest: 20 ≤ x < 25 (f.d. 10.8)

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

stemup.co.uk

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

Question 5 continued

Height of 2nd tallest bar $= \frac{10.8}{18} \times 10 = 6$ cm

(b) Number $\approx \frac{40 - 35}{10} \times 60 + 15 + \frac{57 - 55}{15} \times 6$

$= 30 + 15 + 0.8 = 45.8$, so about 46 cars

(c) Median: $225 \div 2 = 112.5$th value, in class $25 \leq x < 30$

$Q_2 = 25 + \frac{112.5 - 54}{90} \times 5 = 25 + 3.25 = 28.25$ km/h (28.3)

(d)(i) Q_3: $3 \times 225 \div 4 = 168.75$th value, in class $30 \leq x < 40$

$Q_3 = 30 + \frac{168.75 - 144}{60} \times 10 = 30 + 4.125 = 34.125$ km/h (34.1)

(ii) $IQR = Q_3 - Q_1 = 34.125 - 25.14 = 8.985 \approx 8.99$ km/h

(e) $Q_2 - Q_1 = 28.25 - 25.14 = 3.11$

$Q_3 - Q_2 = 34.125 - 28.25 = 5.875$

$Q_3 - Q_2 > Q_2 - Q_1$, so the data are positively skewed.

(f) Ali's suggestion is not suitable: part (e) shows the data are positively skewed (not symmetrical), but a normal distribution is symmetrical.

Question 5 continued

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

stemup.co.uk

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 14 marks)

stemup.co.uk

6: A company produces small bags of flour and large bags of flour.

The weight, X grams, of flour in a small bag is normally distributed with mean 502 and standard deviation 3

One of these small bags of flour is selected at random.

(a) (i) Using standardisation find $P(X > 508)$

(2)

(ii) Hence, find $P(496 < X < 508)$

(2)

A random sample of 4 small bags of flour is taken.

(b) Find the probability that exactly 2 of these small bags each contain more than 508 grams of flour.

(3)

The weight, Y grams, of flour in a large bag is normally distributed with mean 1024 and standard deviation σ

The 85th percentile for the weight of flour in a large bag is 1038.51 to 2 decimal places.

(c) Show that $\sigma = 14$ to the nearest gram.

(2)

Given that $P(X > k) = P(Y < 2k) = p$, where k and p are constants,

(d) (i) find the value of k

(ii) find the value of p

(5)

(a)(i) $P(X > 508) = P\left(Z > \frac{508 - 502}{3}\right) = P(Z > 2)$

$= 1 - 0.9772 = 0.0228$

(ii) $P(496 < X < 508) = 1 - 2 \times P(X > 508)$ (by symmetry)

$= 1 - 2 \times 0.0228 = 0.9544$

(b) Let B = number of the 4 bags with more than 508 g, $B \sim B(4, 0.0228)$

$P(B = 2) = {}^4C_2 \times 0.0228^2 \times (1 - 0.0228)^2$

$= 6 \times 0.0228^2 \times 0.9772^2 = 0.0029785... = 0.00298$

(c) $P(Y < 1038.51) = 0.85 \Rightarrow z = 1.0364$

$\frac{1038.51 - 1024}{\sigma} = 1.0364 \Rightarrow \sigma = \frac{14.51}{1.0364} = 14.0004...$

$\sigma = 14$ to the nearest gram as required

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

stemup.co.uk

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

Question 6 continued

(d)(i) $P(X > k) = P\left(Z > \frac{k - 502}{3}\right)$

$P(Y < 2k) = P\left(Z < \frac{2k - 1024}{14}\right) = P\left(Z > \frac{1024 - 2k}{14}\right)$

So $\frac{k - 502}{3} = \frac{1024 - 2k}{14}$

$14k - 7028 = 3072 - 6k \Rightarrow 20k = 10100 \Rightarrow k = 505$

(ii) $p = P(X > 505) = P\left(Z > \frac{505 - 502}{3}\right) = P(Z > 1)$

$= 1 - 0.8413 = 0.1587$

P 7 6 1 7 6 A 0 1 7 2 4
stemup.co.uk

Question 6 continued

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

stemup.co.uk

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 14 marks)



7: Riley has a bag containing 3 white beads and 6 yellow beads only.

Two beads are selected at random from the bag, one at a time without replacement.

(a) Complete the tree diagram on page 21

(3)

(b) Find the probability that exactly one of the beads selected is yellow.
Show your working.

(2)

(c) Find the probability that at least one of the beads selected is white.

(2)

Given that at least one of the beads selected is white,

(d) find the probability that both of the beads selected are white.

(2)

Riley adds some beads to the bag so it contains a total of n beads. The beads are either red or yellow or white such that

number of **red** beads : number of **yellow** beads : number of **white** beads = 4 : 3 : 1

Four beads are selected at random from the bag, one at a time without replacement.
The probability of getting 2 red beads, followed by a yellow bead, followed by a white bead is $\frac{3}{235}$

(e) Show that n satisfies the equation

$$21n^2 + bn + c = 0$$

where b and c are integers to be found.

(4)

(a) Tree diagram completed on page 21.

(b) P(exactly one yellow) = P(WY) + P(YW) = $\frac{3}{9} \times \frac{6}{8} + \frac{6}{9} \times \frac{3}{8}$

$= \frac{18}{72} + \frac{18}{72} = \frac{36}{72} = \frac{1}{2}$

(c) P(at least one white) = 1 − P(YY) = $1 - \frac{6}{9} \times \frac{5}{8}$

$= 1 - \frac{30}{72} = \frac{42}{72} = \frac{7}{12}$

DO NOT WRITE IN THIS AREA

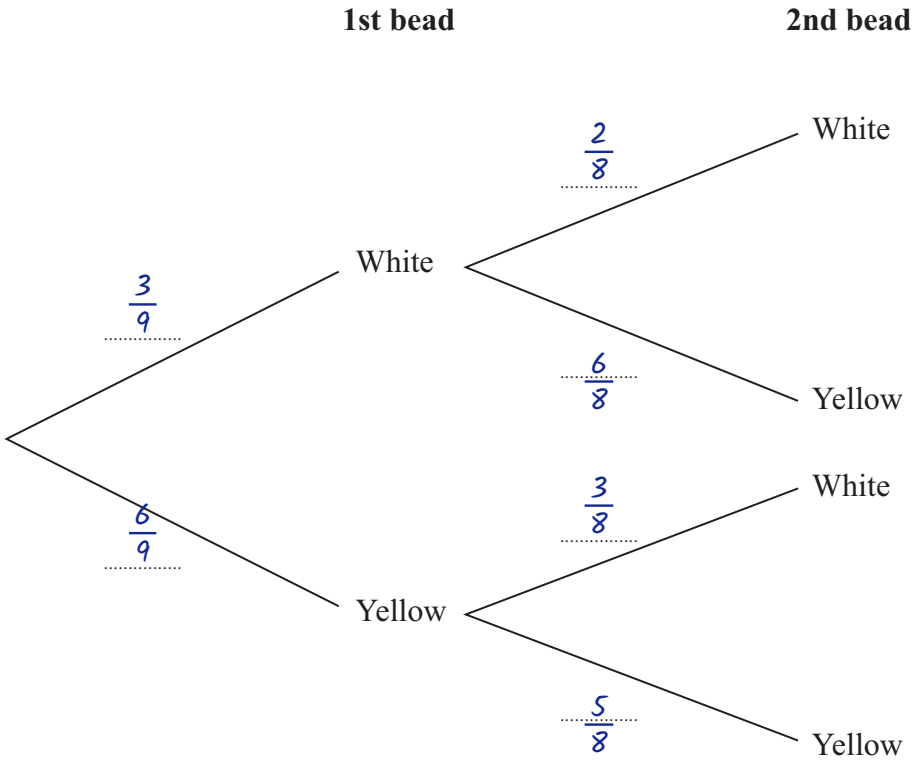
DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

stemup.co.uk

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

Question 7 continued



(d) $P(WW) = \frac{3}{9} \times \frac{2}{8} = \frac{6}{72} = \frac{1}{12}$

$P(WW \mid \text{at least one W}) = \frac{P(WW)}{P(\text{at least one W})} = \frac{1/12}{7/12} = \frac{1}{7}$

(e) Red $= \frac{4}{8}\, n = \frac{n}{2}$, yellow $= \frac{3n}{8}$, white $= \frac{n}{8}$

$P(R, R, Y, W) = \frac{n/2}{n} \times \frac{n/2 - 1}{n - 1} \times \frac{3n/8}{n - 2} \times \frac{n/8}{n - 3} = \frac{3}{235}$

$\frac{1}{2} \times \frac{n - 2}{2(n - 1)} \times \frac{3n}{8(n - 2)} \times \frac{n}{8(n - 3)} = \frac{3}{235}$

$\frac{3n^2}{256(n - 1)(n - 3)} = \frac{3}{235}$

$235n^2 = 256(n - 1)(n - 3) = 256n^2 - 1024n + 768$

$21n^2 - 1024n + 768 = 0$

$b = -1024, \quad c = 768$

Question 7 continued

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

stemup.co.uk

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 13 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

P 7 6 1 7 6 A 0 2 4 2 4

stemup.co.uk