Please check the examination details below before entering your candidate information

Candidate surname StemUp Model Solution | Other names

Pearson Edexcel
International
Advanced Level

Centre Number | Candidate Number

Wednesday 16 October 2019

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 (Blue), 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 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 ▶

P58493A

©2019 Pearson Education Ltd.

1/1/1/

P 5 8 4 9 3 A 0 1 2 8

Pearson

stemup.co.uk

Leave blank

1. Jeremiah is investigating the relationship between the annual heating bill, h dollars, and the total floor area, f square metres, of buildings.

A random sample of 8 buildings is taken and the data for each building are coded using

$$x = \frac{f - 3500}{80} \text{ and } y = \frac{h - 4500}{80}$$

The results for the coded data are summarised below

$$\sum x = 5 \qquad \sum y = 0 \qquad \sum xy = 1818 \qquad S_{xx} = 1754$$

(a) Calculate S_{xy}

(1)

(b) Find the equation of the regression line of h on f in the form $h = a + bf$

(6)

(c) Give an interpretation of (i) the value of b in your regression line,

(ii) the value of a in your regression line.

(2)

(d) Estimate the annual heating bill for a building with a total floor area of 4600 square metres.

(2)

(a) $S_{xy} = \Sigma xy - \frac{\Sigma x \Sigma y}{n} = 1818 - \frac{5 \times 0}{8} = 1818$

(b) $b = \frac{S_{xy}}{S_{xx}} = \frac{1818}{1754} = 1.036488...$

$\bar{x} = \frac{5}{8} = 0.625, \quad \bar{y} = \frac{0}{8} = 0$

$c = \bar{y} - b\bar{x} = 0 - 1.036488... \times 0.625 = -0.647805...$

$y = -0.6478... + 1.0365x$

$\frac{h - 4500}{80} = -0.6478... + 1.036488... \times \frac{f - 3500}{80}$

$h - 4500 = -51.8244... + 1.036488...(f - 3500)$

$h - 4500 = -51.8244... + 1.036488...f - 3627.708...$

$h = 820.467... + 1.036488...f$

$h = 820 + 1.04f$

(c)(i) On average, for every increase of 1 m^2 in the total floor area, the annual heating bill increases by about \$1.04

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

Leave blank

Question 1 continued

(ii) $a = 820$ is the annual heating bill ($820) for a building with zero floor area, i.e. a fixed "standing charge" of about $820

(d) $h = 820.467... + 1.036488... \times 4600$

$= 5588.31...$

Estimated annual heating bill = $5590 (3 s.f.)

Leave blank

Question 1 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

Leave blank

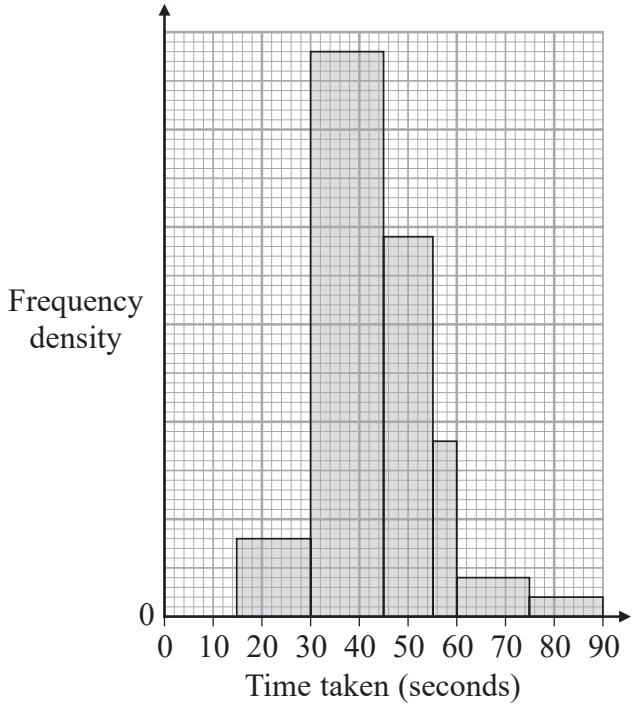
Question 1 continued

(Total 11 marks)

Q1

Leave blank

2. The histogram shows the times taken, in seconds, by each of 260 people to complete a puzzle.



(a) Use the histogram to complete the frequency table for the times taken to complete the puzzle.

Time taken (seconds)	15–30	30–45	45–55	55–60	60–75	75–90
Frequency (f)	20	145	65	15	10	5
Time midpoint (t seconds)	22.5	37.5	50	57.5	67.5	82.5

(3)

Given that $\sum ft = 11\,087.5$ and $\sum ft^2 = 505\,718.75$

(b) find an estimate for

(i) the mean time taken to complete the puzzle,

(1)

(ii) the standard deviation of the times taken to complete the puzzle.

(2)

(c) Use linear interpolation to estimate the median time taken to complete the puzzle.

(2)

(d) Describe the skewness of these data. Give a reason for your answer.

(1)

Three of the 260 people are chosen at random.

(e) Estimate the probability that all 3 of their times are less than 36 seconds.

(4)

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

Leave blank

Question 2 continued

(a) frequency = frequency density × class width

30–45: $fd = 145 \div 15 = 9.67$; 15–30: $fd = 20 \div 15 = 1.33$

45–55: $fd = 6.5$, frequency $= 6.5 \times 10 = 65$

55–60: $fd = 3$, frequency $= 3 \times 5 = 15$

Check: $20 + 145 + 65 + 15 + 10 + 5 = 260$

(b)(i) mean $\bar{t} = \frac{\Sigma ft}{\Sigma f} = \frac{11\,087.5}{260} = 42.644\ldots = 42.6$ seconds

(ii) $\sigma = \sqrt{\left(\frac{\Sigma ft^2}{\Sigma f} - \bar{t}^2\right)} = \sqrt{\left(\frac{505\,718.75}{260} - 42.644\ldots^2\right)}$

$= \sqrt{(1945.072\ldots - 1818.531\ldots)} = \sqrt{126.54\ldots}$

$= 11.249\ldots = 11.2$ seconds

(c) Median is the $260 \div 2 = 130$th value, in the 30–45 class

$Q_2 = 30 + 15 \times \frac{130 - 20}{145}$

$= 41.379\ldots = 41.4$ seconds

(d) Positive skew, since the median (41.4) < the mean (42.6)

(e) Number of people with times less than 36 s:

$20 + \frac{36 - 30}{15} \times 145 = 20 + 58 = 78$

$P(\text{time} < 36) = \frac{78}{260} = 0.3$

$P(\text{all 3} < 36) = \frac{78}{260} \times \frac{77}{259} \times \frac{76}{258}$

$= \frac{209}{7955} = 0.026272\ldots = 0.0263$

Leave blank

Question 2 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

Leave blank

Question 2 continued

(Total 13 marks)

Q2

Leave blank

3. The distance that Jenny throws a javelin is normally distributed with mean 42 m and standard deviation 5 m.

(a) Find the probability that the distance Jenny throws the javelin is less than 40 m.

(3)

Jenny enters a javelin competition. To qualify for the final she has 3 attempts to throw the javelin a distance of more than 40 m. Once she has qualified she does not use any of her remaining attempts.

Given that Jenny qualified for the final and that throws of the javelin are independent,

(b) find the probability that she qualified for the final on her third throw with a distance greater than 45 m.

(5)

(a) $X \sim N(42, 5^2)$

$P(X < 40) = P\left(Z < \frac{40 - 42}{5}\right) = P(Z < -0.4)$

$= 1 - 0.6554$

$= 0.3446$

(b) Let $p = P(X < 40) = 0.3446$

$P(\text{qualify}) = (1 - p) + p(1 - p) + p^2(1 - p) = 1 - p^3$

$= 1 - 0.3446^3 = 0.959079...$

$P(X > 45) = P\left(Z > \frac{45 - 42}{5}\right) = P(Z > 0.6)$

$= 1 - 0.7257 = 0.2743$

$P(\text{fail, fail, throw} > 45) = 0.3446^2 \times 0.2743 = 0.032572...$

$P(\text{3rd throw} > 45 \mid \text{qualified}) = \frac{0.3446^2 \times 0.2743}{0.959079...}$

$= \frac{0.032572...}{0.959079...}$

$= 0.033962... = 0.0340$

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

Leave blank

Question 3 continued

Leave blank

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

Leave blank

Question 3 continued

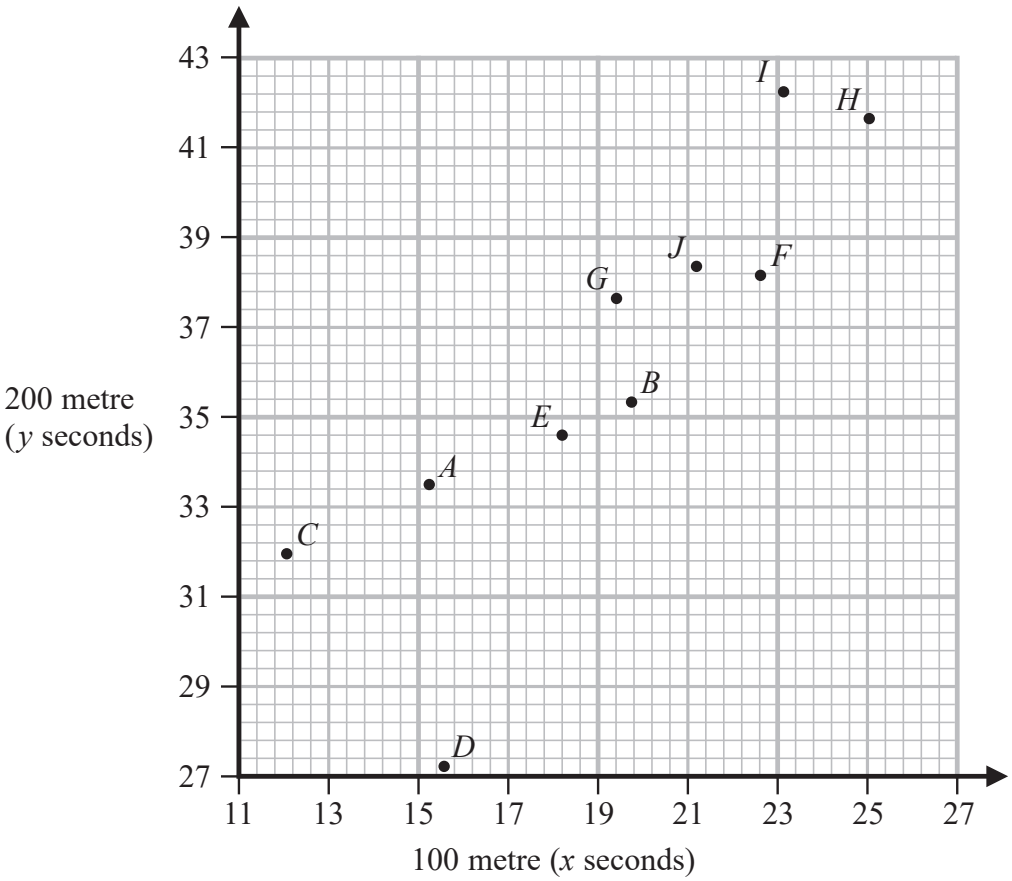
Q3

(Total 8 marks)

stemup.co.uk

Leave blank

4. A random sample of 10 boys A, B, C, D, E, F, G, H, I and J is taken from a junior athletics club. Each boy selected is asked to run a 100 metre race and a 200 metre race. The time taken, x seconds, by each boy to run the 100 metre race is recorded and the time taken, y seconds, by each boy to run the 200 metre race is recorded. The results are plotted on the scatter diagram below.



(a) State, without calculation, which of the 3 values below is most likely to be a value of the product moment correlation coefficient for the data in the scatter diagram.

0.72 0.05 0.95

(1)

In the sample of 10 boys, one is a junior champion 100 metre runner and one is a junior champion 200 metre runner.

(b) Write down the boy who is most likely to be the 100 metre junior champion.

(1)

The data for the two junior champions are removed and the remaining data are summarised below

$\sum x^2 = 3445.26$ $\sum x = 164.4$ $S_{yy} = 67.52$ $S_{xy} = 60.85$

(c) (i) Calculate the value of the product moment correlation coefficient for the remaining data.

(3)

(ii) Comment, in context, on the value of the product moment correlation coefficient that you obtained in part (i).

(1)

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

stemup.co.uk

Leave blank

Question 4 continued

(a) 0.72

(b) C

(c)(i) $S_{xx} = \Sigma x^2 - \frac{(\Sigma x)^2}{n} = 3445.26 - \frac{164.4^2}{8}$

$= 3445.26 - 3378.42 = 66.84$

$r = \frac{S_{xy}}{\sqrt{(S_{xx}\, S_{yy})}} = \frac{60.85}{\sqrt{(66.84 \times 67.52)}}$

$= 0.905787... = 0.906$

(ii) The faster a boy runs the 100 metres, the faster he tends to run the 200 metres (strong positive correlation between times).

Q4

(Total 6 marks)

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

Leave blank

5. A housing estate consists of 320 houses.

A house on the estate is selected at random.

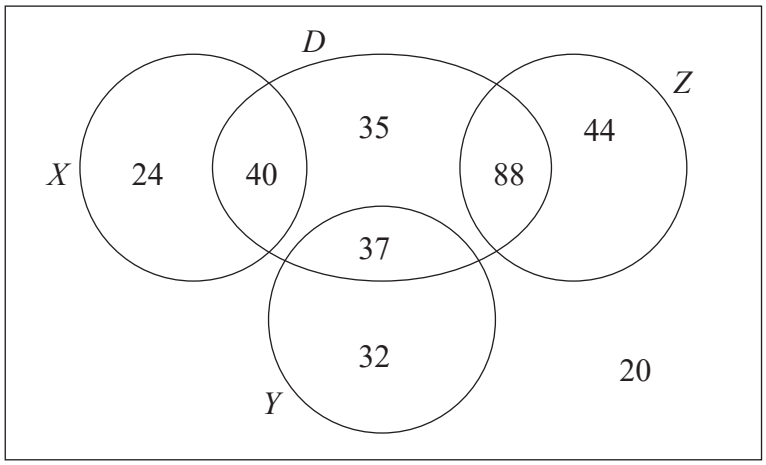
D denotes the event 'the house has a driveway'

X denotes the event 'no cars per household'

Y denotes the event 'exactly 1 car per household'

Z denotes the event 'exactly 2 cars per household'

The Venn diagram shows the number of households in each region for the events D, X, Y and Z.



(a) Find (i) P(D)

(1)

(ii) P($D \cap X'$)

(1)

(iii) P($D' \cup Z'$)

(2)

Given that the house has a driveway,

(b) find the probability that there are exactly 2 cars in the household.

(2)

(c) Write down 2 of the events D, X, Y and Z that are mutually exclusive.

(1)

(d) Determine whether the events D and X are independent. Justify your answer.

(2)

(e) Define in the context of this question, the event

(i) $D' \cap Z$

(1)

(ii) $D \cap (X \cup Y)$

(2)

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

Leave blank

Question 5 continued

(a)(i) $P(D) = \frac{35 + 40 + 88 + 37}{320} = \frac{200}{320} = \frac{5}{8}$

(ii) $P(D \cap X') = \frac{35 + 37 + 88}{320} = \frac{160}{320} = \frac{1}{2}$

(iii) $P(D' \cup Z') = 1 - P(D \cap Z) = \frac{320 - 88}{320} = \frac{232}{320} = \frac{29}{40}$

(b) $P(Z \mid D) = \frac{P(Z \cap D)}{P(D)} = \frac{88/320}{200/320} = \frac{88}{200} = \frac{11}{25}$ $(= 0.44)$

(c) X and Y (also X and Z, or Y and Z)

(d) $P(D) = \frac{5}{8} = 0.625$, $P(X) = \frac{24 + 40}{320} = \frac{64}{320} = 0.2$

$P(D) \times P(X) = 0.625 \times 0.2 = 0.125$

$P(D \cap X) = \frac{40}{320} = 0.125$

$P(D \cap X) = P(D) \times P(X)$, so D and X are independent.

(e)(i) A house that does not have a driveway and has exactly 2 cars.

(ii) A house that has a driveway and has fewer than 2 cars (i.e. no cars or exactly 1 car).

Leave blank

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

Leave blank

Question 5 continued

(Total 12 marks)

Q5

Leave blank

6. A machine cuts wood into pieces. The lengths, W metres, of the pieces produced by the machine are normally distributed with mean μ metres and standard deviation σ metres. It is known that

$$P(W < 3.968) = 0.1 \text{ and } P(3.968 < W < 4.026) = 0.75$$

(a) Calculate the value of μ and the value of σ

(5)

A second machine cuts wood into logs. The lengths, L cm, of the logs produced by this second machine are normally distributed with $L \sim N(30, 0.5^2)$

An outlier is a value that is greater than $Q_3 + 1.5 \times (Q_3 - Q_1)$ or smaller than $Q_1 - 1.5 \times (Q_3 - Q_1)$

A log is selected at random.

Given that $Q_1 = 29.7$ to 3 significant figures,

(b) find the probability that the length of this log is an outlier.

(5)

(a) $P(W < 3.968) = 0.1 \Rightarrow z = -1.2816$

$P(W < 4.026) = 0.1 + 0.75 = 0.85 \Rightarrow z = 1.0364$

$\frac{3.968 - \mu}{\sigma} = -1.2816 \Rightarrow \mu - 1.2816\sigma = 3.968$

$\frac{4.026 - \mu}{\sigma} = 1.0364 \Rightarrow \mu + 1.0364\sigma = 4.026$

Subtract: $2.3184\sigma = 0.058$

$\sigma = 0.025021... = 0.0250$

$\mu = 3.968 + 1.2816 \times 0.025021... = 4.00006...$

$\mu = 4.00$ metres, $\sigma = 0.0250$ metres

(b) $L \sim N(30, 0.5^2)$. By symmetry $Q_3 = 30 + (30 - 29.7) = 30.3$

$IQR = Q_3 - Q_1 = 30.3 - 29.7 = 0.6$

Upper limit $= 30.3 + 1.5 \times 0.6 = 31.2$

Lower limit $= 29.7 - 1.5 \times 0.6 = 28.8$

$P(L > 31.2) = P\left(Z > \frac{31.2 - 30}{0.5}\right) = P(Z > 2.4)$

$= 1 - 0.9918 = 0.0082$

By symmetry $P(L < 28.8) = 0.0082$

$P(\text{outlier}) = 2 \times 0.0082 = 0.0164$

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

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

stemup.co.uk

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

Leave blank

Question 6 continued

(Total 10 marks)

Q6

Leave blank

7. The number of cakes, X, bought by customers at a particular shop has probability distribution

x	1	2	3	4	5	6	>6
$P(X = x)$	0.35	a	a	0.15	b	b	0

where a and b are constants.

Given that $E(X) = 2.5$

(a) (i) show that $a = 0.2$

(ii) find the value of b

(5)

(b) Calculate $\text{Var}(4X + 3)$

(4)

The cost to produce each cake is 20 cents and the shopkeeper sells each cake for 80 cents.

(c) Find the expected profit made by the shopkeeper for a randomly selected customer buying cakes.

(2)

The shopkeeper decides to run a promotion where she gives 1 free cake to all customers who buy 4 or more cakes. During the promotion the number of cakes, Y, taken away by a customer buying cakes has the following probability distribution

y	1	2	3	4	5	6
$P(Y = y)$	$\frac{3}{40}$	$\frac{4}{40}$	$\frac{3}{40}$	0	$\frac{22}{40}$	$\frac{8}{40}$

(d) Find the expected profit made by the shopkeeper for a randomly selected customer buying cakes during the promotion.

(4)

(a)(i) $\Sigma P = 1$: $0.35 + 2a + 0.15 + 2b = 1$

$\Rightarrow \quad 2a + 2b = 0.5 \quad \Rightarrow \quad b = 0.25 - a$

$E(X) = 0.35 + 2a + 3a + 0.6 + 5b + 6b = 2.5$

$\Rightarrow \quad 5a + 11b = 1.55$

$b = 0.25 - a$: $\quad 5a + 11(0.25 - a) = 1.55$

$5a + 2.75 - 11a = 1.55 \quad \Rightarrow \quad 6a = 1.2$

$a = 0.2$

(ii) $b = 0.25 - 0.2 = 0.05$

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

P 5 8 4 9 3 A 0 2 4 2 8

stemup.co.uk

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

Leave blank

Question 7 continued

(b) $E(X^2) = 1^2(0.35) + 2^2(0.2) + 3^2(0.2) + 4^2(0.15)$
$+ 5^2(0.05) + 6^2(0.05)$
$= 0.35 + 0.8 + 1.8 + 2.4 + 1.25 + 1.8 = 8.4$
$\text{Var}(X) = E(X^2) - [E(X)]^2 = 8.4 - 2.5^2 = 2.15$
$\text{Var}(4X + 3) = 4^2\ \text{Var}(X) = 16 \times 2.15 = 34.4$

(c) Profit per cake = 80 − 20 = 60 cents
Expected profit = $60 \times E(X) = 60 \times 2.5 = 150$ cents ($1.50)

(d) Let W = profit in cents per customer.
$y = 1$: $w = 60$; $y = 2$: $w = 120$; $y = 3$: $w = 180$
$y = 5$ (4 bought): $w = 4 \times 80 - 5 \times 20 = 220$
$y = 6$ (5 bought): $w = 5 \times 80 - 6 \times 20 = 280$

$E(W) = \frac{1}{40}(60 \times 3 + 120 \times 4 + 180 \times 3 + 220 \times 22 + 280 \times 8)$

$= \frac{8280}{40} = 207$ cents ($2.07)

Leave blank

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

Leave blank

Question 7 continued

Leave blank

Question 7 continued

Q7

(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

stemup.co.uk