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

Friday 17 October 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 8 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 ▶

P84842A

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

P 8 4 8 4 2 A 0 1 2 8

Pearson

1. A biologist is studying horses.

The biologist records the height, h cm, and the weight, w kg, of 8 horses.

The table below shows the biologist's results.

Height	140	145	150	157	138	140	148	142
Weight	500	521	539	554	492	497	532	525

[You may use $\sum h = 1160$ $\sum w = 4160$ $\sum hw = 604135$ $\sum w^2 = 2166600$ $S_{hh} = 286$]

(a) Calculate the value of S_{hw} and the value of S_{ww}

(3)

(b) Calculate the value of the product moment correlation coefficient between h and w

(2)

(c) Show that the equation of the regression line of w on h can be written as

$$w = 46.0 + 3.27h$$

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

(3)

(d) Give an interpretation, in context, of the gradient of the regression line.

(1)

Using the equation of the regression line in part (c), estimate the weight of a horse with a height of

(e) (i) 153 cm

(ii) 170 cm

(2)

(f) State, giving a reason, which of the two estimates found in part (e) would give the more reliable estimate.

(1)

(a) $S_{hw} = 604135 - \frac{1160 \times 4160}{8} = 604135 - 603200 = 935$

$S_{ww} = 2166600 - \frac{4160^2}{8} = 2166600 - 2163200 = 3400$

(b) $r = \frac{S_{hw}}{\sqrt{(S_{hh}\, S_{ww})}} = \frac{935}{\sqrt{(286 \times 3400)}} = 0.948181... = 0.948$

(c) $b = \frac{S_{hw}}{S_{hh}} = \frac{935}{286} = 3.269230...$

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

$\bar{h} = \frac{1160}{8} = 145, \quad \bar{w} = \frac{4160}{8} = 520$

$a = \bar{w} - b\,\bar{h} = 520 - 3.269230... \times 145 = 45.961538...$

$\therefore \; w = 46.0 + 3.27h \quad (3 \text{ s.f.})$

(d) On average, for each 1 cm increase in height, the weight of a horse increases by 3.27 kg.

(e)(i) $w = 46.0 + 3.27 \times 153 = 546.31$, weight $\approx$ 546 kg

(ii) $w = 46.0 + 3.27 \times 170 = 601.9$, weight $\approx$ 602 kg

(f) The estimate for 153 cm is more reliable, since 153 cm lies within the range of heights in the data (138 to 157), so it is interpolation, whereas 170 cm is outside the range (extrapolation).

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 12 marks)

2. The weights, to the nearest kilogram, of 35 rugby players from club A are summarised in the stem and leaf diagram below.

Weight (kg)		Totals
7	2	(1)
8	4 9	(2)
9	0 0 1 1 5 5 5 6 7	(9)
10	0 0 4 4 5 5 7 8 9 9	(10)
11	1 1 2 2 2 2 5	(7)
12	0 2 4 6 7	(5)
13	9	(1)

Key: 7 | 2 represents 72 kg

(a) Find the lower quartile and the upper quartile of the weights of these rugby players.

(2)

An outlier is defined as an observation that is

greater than $Q_3 + 1.5 \times (Q_3 - Q_1)$ or less than $Q_1 - 1.5 \times (Q_3 - Q_1)$

(b) Show that there is only one outlier for these data.

(3)

Figure 1 on page 7 shows a box plot for the weights of rugby players from club B

(c) On the same grid, draw a box plot to represent the weights of rugby players from club A

(4)

(d) Describe and interpret the skew of the distribution for **club B**, shown in Figure 1

(2)

(a) $n = 35$: Q_1 is the 9th value, Q_3 is the 27th value

$Q_1 = 95$ kg, $Q_3 = 112$ kg

(b) $Q_3 - Q_1 = 112 - 95 = 17$

Lower limit: $95 - 1.5 \times 17 = 69.5$

Upper limit: $112 + 1.5 \times 17 = 137.5$

Smallest value 72 > 69.5, so no outliers at the lower end.

Values above 137.5: only 139 (next largest is 127 < 137.5)

∴ there is only one outlier, 139.

(c) Median = 18th value = 105 kg

Whiskers to 72 and 127, with 139 marked as an outlier.

Box plot drawn on the grid below.

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

(d) For club B: $Q_2 - Q_1 = 101 - 97 = 4$, $Q_3 - Q_2 = 113 - 101 = 12$

$Q_3 - Q_2 > Q_2 - Q_1$, so club B has positive skew.

The weights of club B players are more spread out (more varied) above the median than below the median.

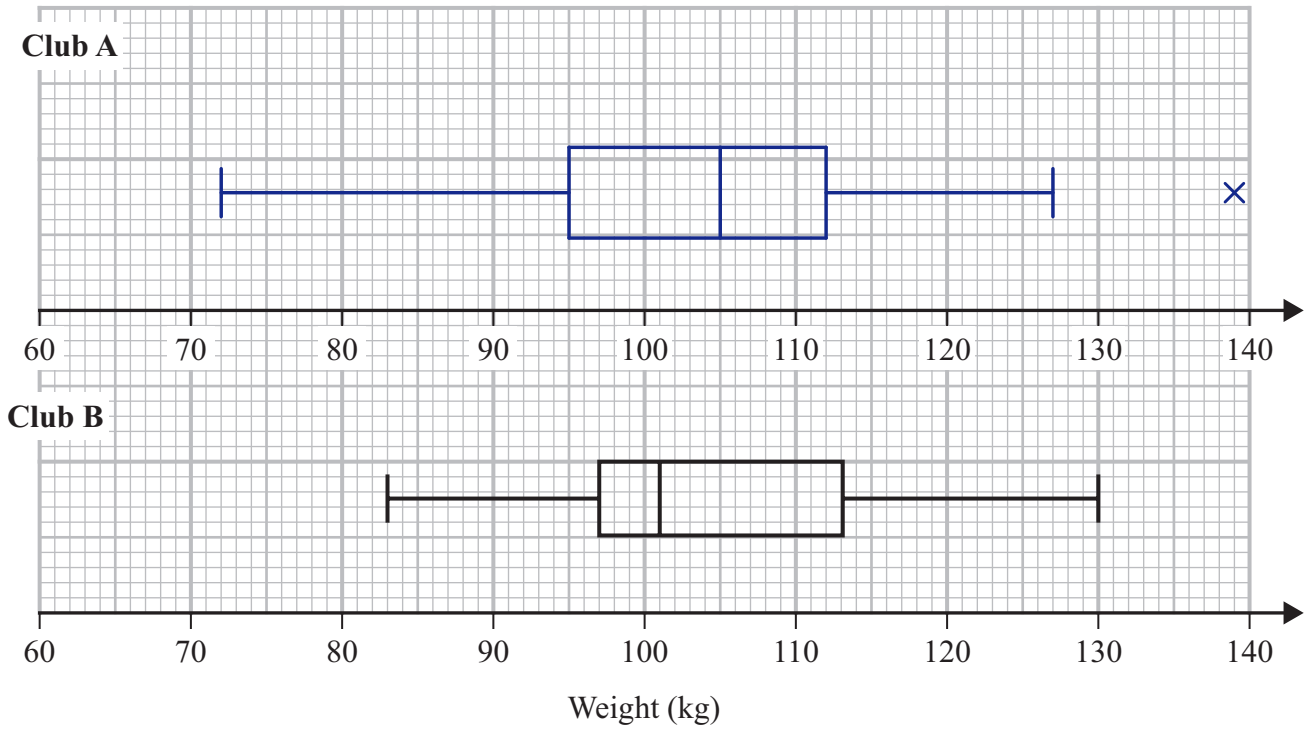


Figure 1

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

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

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

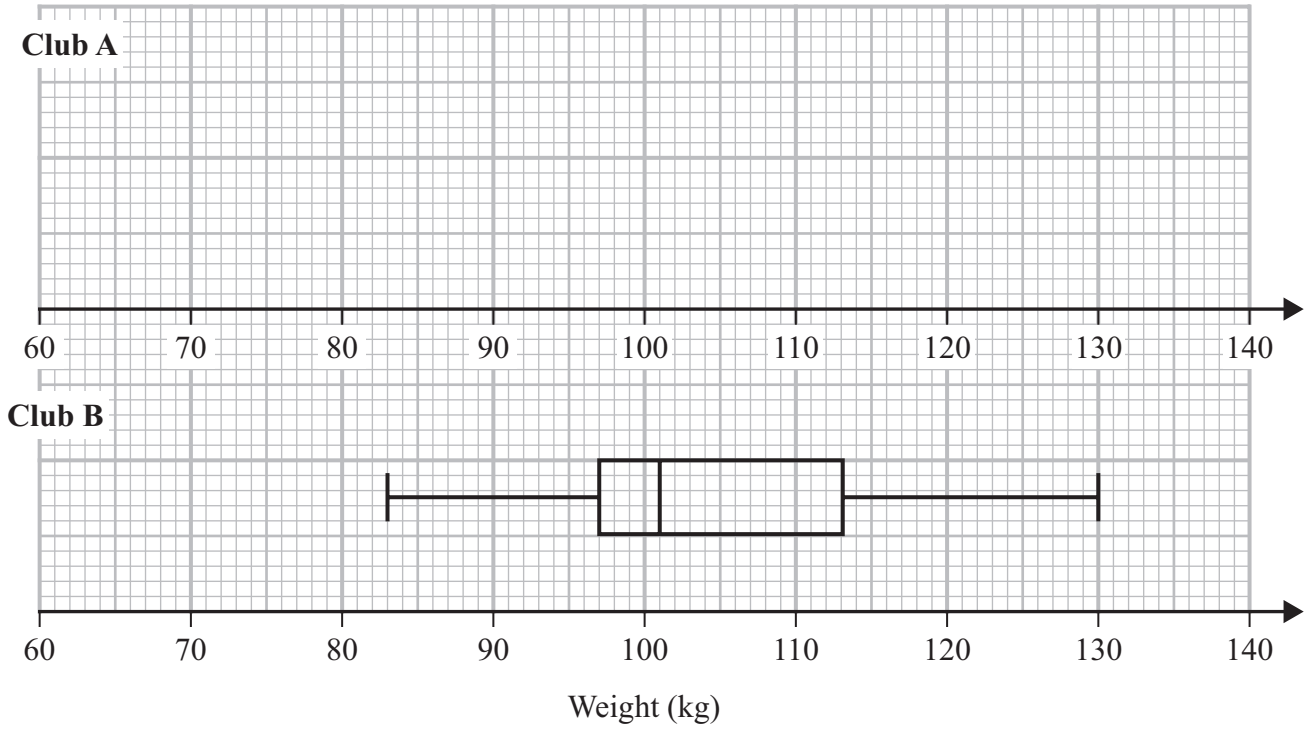


Figure 1

(Total for Question 2 is 11 marks)

3. A fair 6-sided die is rolled.

The random variable Y represents the score on the uppermost face.

(a) Write down the probability function of Y

(2)

(b) State the name of the distribution of Y

(1)

(c) Find $E(4Y + 3)$

(3)

(a) $P(Y = y) = \frac{1}{6}$, $y = 1, 2, 3, 4, 5, 6$

(b) Discrete uniform distribution

(c) $E(Y) = \frac{1 + 2 + 3 + 4 + 5 + 6}{6} = 3.5$

$E(4Y + 3) = 4E(Y) + 3 = 4 \times 3.5 + 3 = 17$

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 discrete random variable X has the probability distribution given in the table below, where a and b are constants.

x	–1	0	1	2	3
$P(X = x)$	0.05	0.1	a	b	0.35

Given that $E(X) = 1.8$

(a) find the value of a and the value of b

(3)

(b) Specify the cumulative distribution function $F(x)$ for $x = –1, 0, 1, 2$ and 3

(2)

(c) Find $P(2X < 1)$

(1)

(d) Find $\text{Var}(5 – 3X)$
You must show your working.

(4)

(a) $\Sigma P(X = x) = 1$: $0.05 + 0.1 + a + b + 0.35 = 1 \Rightarrow a + b = 0.5$

$E(X) = 1.8$: $-0.05 + 0 + a + 2b + 1.05 = 1.8 \Rightarrow a + 2b = 0.8$

Subtracting: $b = 0.3$, $a = 0.5 - 0.3 = 0.2$

$a = 0.2$, $b = 0.3$

(b)

x	-1	0	1	2	3
$F(x)$	0.05	0.15	0.35	0.65	1

(c) $P(2X < 1) = P(X < 0.5) = P(X \leq 0) = F(0) = 0.15$

(d) $E(X^2) = 1 \times 0.05 + 0 + 1 \times 0.2 + 4 \times 0.3 + 9 \times 0.35 = 4.6$

$\text{Var}(X) = E(X^2) - [E(X)]^2 = 4.6 - 1.8^2 = 1.36$

$\text{Var}(5 - 3X) = (-3)^2 \text{Var}(X) = 9 \times 1.36 = 12.24$

$\text{Var}(5 - 3X) = 12.24$ (= 12.2 to 3 s.f.)

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

Question 4 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 4 continued

(Total for Question 4 is 10 marks)

5. The weights, x grams, of 300 medium-sized apples are summarised in the table below.

Weight	Frequency
$105 \leqslant x < 115$	28
$115 \leqslant x < 125$	42
$125 \leqslant x < 130$	75
$130 \leqslant x < 140$	63
$140 \leqslant x < 150$	52
$150 \leqslant x < 170$	40

A histogram is drawn (not shown) to represent these data.

The bar that represents the class $125 \leqslant x < 130$ is 1.8 cm wide and 12 cm high.

(a) Calculate the width and the height of the bar that represents the class $150 \leqslant x < 170$

(3)

(b) Use linear interpolation to estimate the median of x
Give your answer to one decimal place.

(2)

The weights, w grams, of 300 large apples are coded using $y = \dfrac{w - 200}{2}$

These coded data are summarised by

$$\sum y = 6000 \qquad \sum y^2 = 150\,000$$

(c) Calculate

(i) the mean of w

(2)

(ii) the variance of w

(3)

(a) Width: class width 5 g ↔ 1.8 cm, so 20 g ↔ $\frac{1.8}{5} \times 20 = 7.2$ cm

Area of 125–130 bar = $1.8 \times 12 = 21.6$ cm² for a frequency of 75

Area for frequency 40 = $\frac{21.6}{75} \times 40 = 11.52$ cm²

Height = $\frac{11.52}{7.2} = 1.6$ cm. Width = 7.2 cm, height = 1.6 cm

(b) Median is the 150th value; cumulative frequencies 28, 70, 145, 208

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

$$Q_2 = 130 + \frac{150 - 145}{63} \times 10 = 130.7936... = 130.8 \text{ g}$$

(c)(i) $\bar{y} = \frac{6000}{300} = 20$

$w = 2y + 200 \Rightarrow \bar{w} = 2 \times 20 + 200 = 240$ g

(ii) $\text{Var}(y) = \frac{150000}{300} - 20^2 = 500 - 400 = 100$

$\text{Var}(w) = 2^2 \times \text{Var}(y) = 4 \times 100 = 400$

Question 5 continued

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

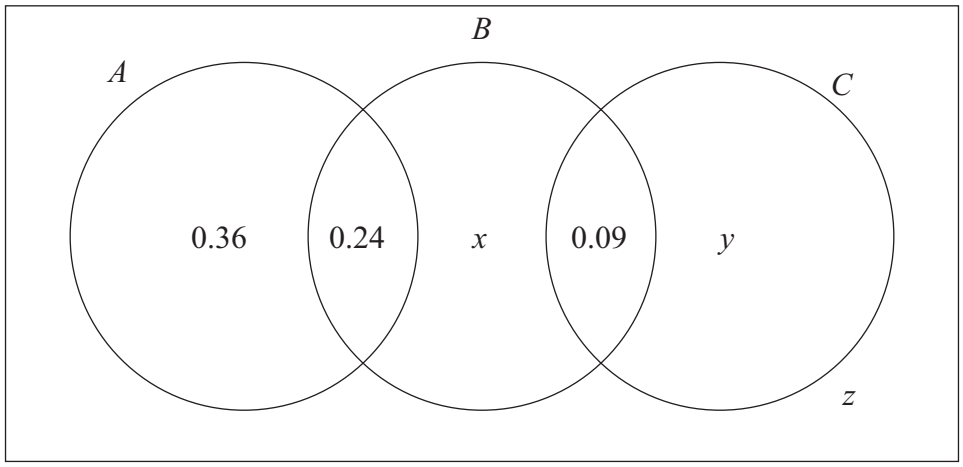
DO NOT WRITE IN THIS AREA

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 10 marks)

6. The Venn diagram, where x, y and z are probabilities, represents the events A, B and C and their associated probabilities.



Given that the events A and B are independent,

(a) find the value of x

(3)

Given also that $\mathrm{P}(B \mid C') = \dfrac{31}{75}$

(b) find the value of z

(2)

(c) Hence find the value of y

(2)

(d) Find

(i) $\mathrm{P}\left((B \cup C)' \cap A\right)$

(ii) $\mathrm{P}\left((A \cup B) \cap C'\right)$

(2)

(a) $P(A) = 0.36 + 0.24 = 0.6$, $P(B) = 0.24 + x + 0.09 = 0.33 + x$

$P(A \cap B) = 0.24$

Independent: $P(A \cap B) = P(A) \times P(B) \Rightarrow 0.6(0.33 + x) = 0.24$

$0.33 + x = 0.4 \Rightarrow x = 0.07$

(b) $P(B \mid C') = \dfrac{P(B \cap C')}{P(C')} = \dfrac{0.24 + 0.07}{0.36 + 0.24 + 0.07 + z} = \dfrac{0.31}{0.67 + z} = \dfrac{31}{75}$

$0.67 + z = 0.75 \Rightarrow z = 0.08$

(c) $y = 1 - 0.36 - 0.24 - 0.07 - 0.09 - 0.08$

$y = 0.16$

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

Question 6 continued

(d)(i) $P((B \cup C)' \cap A) = 0.36$

(ii) $P((A \cup B) \cap C') = 0.36 + 0.24 + 0.07 = 0.67$

(Total for Question 6 is 9 marks)

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

7. The wingspans, W mm, of Speckled Wood butterflies are normally distributed, with $W \sim N(51, 1.7^2)$

A Speckled Wood butterfly is selected at random.

(a) Use standardisation to find the probability that this butterfly has a wingspan between 48.45 mm and 54.40 mm.

(3)

Given that this butterfly has a wingspan between 48.45 mm and 54.40 mm,

(b) find the probability that the wingspan of this butterfly is greater than 53.55 mm. You must show your working.

(3)

The wingspans, X mm, of Large White butterflies are also normally distributed, with $X \sim N(\mu, \sigma^2)$

Given that $P(X < 64.58) = 0.30$ and $P(X > 69.46) = 0.10$

(c) calculate the value of μ and the value of σ
You must show your working.

(4)

(a) $P(48.45 < W < 54.40) = P\left(\frac{48.45 - 51}{1.7} < Z < \frac{54.40 - 51}{1.7}\right)$

$= P(-1.5 < Z < 2) = P(Z < 2) - P(Z < -1.5)$

$= 0.9772 - (1 - 0.9332) = 0.9772 - 0.0668 = 0.9104$

(b) $P(53.55 < W < 54.40) = P\left(\frac{53.55 - 51}{1.7} < Z < 2\right) = P(1.5 < Z < 2)$

$= 0.9772 - 0.9332 = 0.0440$

$P(W > 53.55 \mid 48.45 < W < 54.40) = \frac{P(53.55 < W < 54.40)}{P(48.45 < W < 54.40)} = \frac{0.0440}{0.9104}$

$= 0.048330... = 0.0483$

(c) $P(X < 64.58) = 0.30 \Rightarrow \frac{64.58 - \mu}{\sigma} = -0.5244$

$P(X > 69.46) = 0.10 \Rightarrow \frac{69.46 - \mu}{\sigma} = 1.2816$

$64.58 - \mu = -0.5244\sigma$ and $69.46 - \mu = 1.2816\sigma$

Subtracting: $4.88 = 1.8060\sigma \Rightarrow \sigma = 2.702104... = 2.70$

$\mu = 64.58 + 0.5244 \times 2.702104... = 65.996983... = 66.0$

$\mu = 66.0, \quad \sigma = 2.70$

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

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

(Total for Question 7 is 10 marks)

8. When Paul goes to work, he travels by car **or** by bicycle **or** he walks.

- the probability that he travels by car is $\frac{1}{2}$
- the probability that he travels by bicycle is $\frac{1}{6}$
- the probability that he walks is $\frac{1}{3}$

Given that Paul

- travels by car, the probability that he is late is $\frac{1}{5}$
- travels by bicycle, the probability that he is late is $\frac{1}{10}$
- walks, the probability that he is late is $\frac{1}{20}$

(a) Show that on any day that Paul goes to work, the probability that he is late is $\frac{2}{15}$

(1)

Given that Paul is late on a day that he goes to work,

(b) show that the probability that he travels to work by car is $\frac{3}{4}$

(2)

(c) Hence complete the tree diagram on page 28, to represent the given information.

(4)

(a) $P(\text{late}) = \frac{1}{2} \times \frac{1}{5} + \frac{1}{6} \times \frac{1}{10} + \frac{1}{3} \times \frac{1}{20} = \frac{1}{10} + \frac{1}{60} + \frac{1}{60} = \frac{8}{60} = \frac{2}{15}$

(b) $P(\text{car} \mid \text{late}) = \frac{P(\text{car} \cap \text{late})}{P(\text{late})} = \frac{1/2 \times 1/5}{2/15} = \frac{1/10}{2/15} = \frac{15}{20} = \frac{3}{4}$

(c) $P(\text{bicycle} \mid \text{late}) = \frac{1/6 \times 1/10}{2/15} = \frac{1/60}{2/15} = \frac{1}{8}$, $P(\text{walk} \mid \text{late}) = \frac{1/3 \times 1/20}{2/15} = \frac{1}{8}$

$P(\text{not late}) = 1 - \frac{2}{15} = \frac{13}{15}$

$P(\text{car} \mid \text{not late}) = \frac{1/2 \times 4/5}{13/15} = \frac{2/5}{13/15} = \frac{6}{13}$

$P(\text{bicycle} \mid \text{not late}) = \frac{1/6 \times 9/10}{13/15} = \frac{3/20}{13/15} = \frac{9}{52}$

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

Question 8 continued

$P(\text{walk} \mid \text{not late}) = \frac{1/3 \times 19/20}{13/15} = \frac{19/60}{13/15} = \frac{19}{52}$

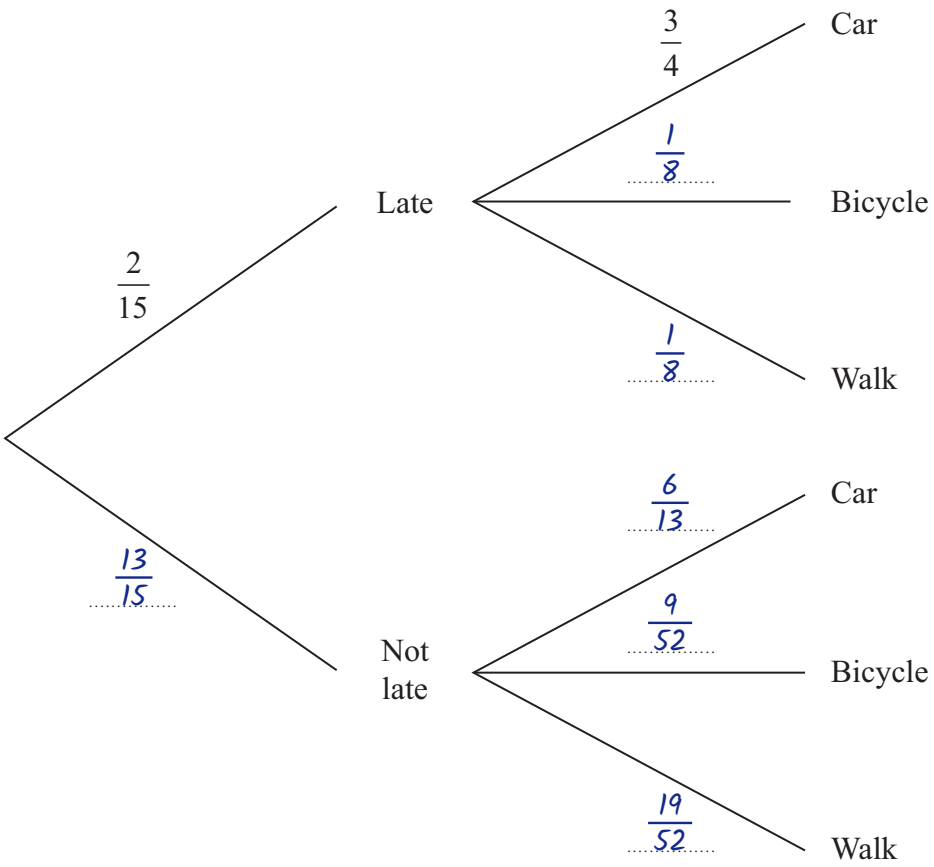
Tree diagram completed on page 28.

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

Question 8 continued



(Total for Question 8 is 7 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