

## Topic 4 Part 4 [94 marks]

1a. [2 marks]

### Markscheme

$$\text{Total} = 2 + 3 + 5 + 7 + 11 + 5 + 6 + 9 + 2 + 1 \quad (\text{M1})$$

(M1) is for a sum of frequencies.

$$= 51 \quad (\text{A1})(\text{G2})$$

[2 marks]

### Examiners report

a) b), c) There was much confusion about how to present the intervals. Often the mid-point only was seen. (eg. 65 instead of 60-70). Understanding of mode, median and mean was usually good but too many candidates wasted time calculating standard deviation by hand and got it wrong. In c(ii) 'greater than three' caused no problems but 'above the mean' was often ignored.

1b. [5 marks]

### Markscheme

Unit penalty (UP) is applicable where indicated in the left hand column.

(i) modal interval is 60 – 70

Award (A0) for 65 (A1)

(ii) median is length of fish no. 26, (M1)(A1)

also 60 – 70 (G2)

Can award (A1)(ft) or (G2)(ft) for 65 if (A0) was awarded for 65 in part (i).

(iii) mean is

$$\frac{2 \times 25 + 3 \times 35 + 5 \times 45 + 7 \times 55 + \dots}{51} \quad (\text{M1})$$

$$(\text{UP}) = 69.5 \text{ cm (3sf)} \quad (\text{A1})(\text{ft})(\text{G1})$$

Note: (M1) is for a sum of (frequencies multiplied by midpoint values) divided by candidate's answer from part (a). Accept mid-points 25.5, 35.5 etc or 24.5, 34.5 etc, leading to answers 70.0 or 69.0 (3sf) respectively. Answers of 69.0, 69.5 or 70.0 (3sf) with no working can be awarded (G1).

[5 marks]

### Examiners report

a) b), c) There was much confusion about how to present the intervals. Often the mid-point only was seen. (eg. 65 instead of 60-70). Understanding of mode, median and mean was usually good but too many candidates wasted time calculating standard deviation by hand and got it wrong. In c(ii) 'greater than three' caused no problems but 'above the mean' was often ignored.

1c. [3 marks]

## Markscheme

Unit penalty (UP) is applicable where indicated in the left hand column.

(UP) (i) standard deviation is 21.8 cm (GI)

For any other answer without working, award (G0). If working is present then (G0)(AP) is possible.

(ii)

$$69.5 + 3 \times 21.8 = 134.9 > 120 \quad (MI)$$

no fish (AI)(ft)(GI)

For 'no fish' without working, award (GI) regardless of answer to (c)(i). Follow through from (c)(i) only if method is shown.

[3 marks]

## Examiners report

a) b), c) There was much confusion about how to present the intervals. Often the mid-point only was seen. (eg. 65 instead of 60-70). Understanding of mode, median and mean was usually good but too many candidates wasted time calculating standard deviation by hand and got it wrong. In c(ii) 'greater than three' caused no problems but 'above the mean' was often ignored.

1d. [2 marks]

## Markscheme

5 fish are less than 40 cm in length, (MI)

Award (MI) for any of

$$\frac{5}{51}, \frac{46}{51}, 0.098 \text{ or } 9.8\%, 0.902, 90.2\% \text{ or } 5.1 \text{ seen.}$$

hence no fine. (AI)(ft)

Note: There is no G mark here and (M0)(AI) is never allowed. The follow-through is from answer in part (a).

[2 marks]

## Examiners report

d) This was often well done, even if earlier parts were poorly done.

1e. [2 marks]

## Markscheme

(i) and (iii) are correct. (AI)(AI)

[2 marks]

## Examiners report

e) Rather mixed performance here. It was hard to identify any consistency in the errors made.

Too much time was spent on this question. It was only worth two marks and candidates ought to have realised that it relied on a general pictorial understanding of the concepts, possibly supplemented by a little elementary arithmetic only, to compare (iii) and (vi). With good understanding, many of the options could be ruled out in a few seconds.

2a. [1 mark]

## Markscheme

$$45 \leq t < 60 \quad (A1)$$

[1 mark]

## Examiners report

Many candidates who had survived the previous two unit penalties, fell here with omission of units for the mean and standard deviation. The modal group was answered well. Part (b), finding the mean and standard deviation by GDC, was answered very poorly. Most did put the midpoints in one list and the frequencies in a second list but then either used the 2-Var stats button or 1-var stats button but only named L1 instead of L1, L2. Candidates who showed midpoints in their working did at least score a method mark.

2b. [3 marks]

## Markscheme

Unit penalty (UP) is applicable in question part (i)(b) only.

$$(UP) 42.4 \text{ minutes} \quad (G2)$$

$$21.6 \text{ minutes} \quad (G1)$$

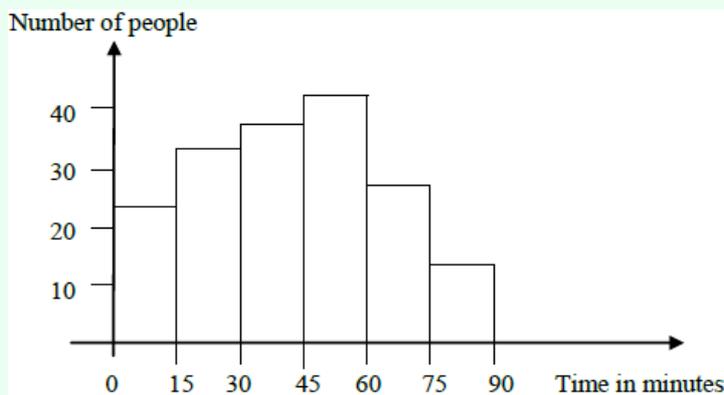
[3 marks]

## Examiners report

Many candidates who had survived the previous two unit penalties, fell here with omission of units for the mean and standard deviation. The modal group was answered well. Part (b), finding the mean and standard deviation by GDC, was answered very poorly. Most did put the midpoints in one list and the frequencies in a second list but then either used the 2-Var stats button or 1-var stats button but only named L1 instead of L1, L2. Candidates who showed midpoints in their working did at least score a method mark.

2c. [4 marks]

## Markscheme



(A4)

[4 marks]

## Examiners report

Many candidates who had survived the previous two unit penalties, fell here with omission of units for the mean and standard deviation. The modal group was answered well. Part (b), finding the mean and standard deviation by GDC, was answered very poorly. Most did put the midpoints in one list and the frequencies in a second list but then either used the 2-Var stats button or 1-var stats button but only named L1 instead of L1, L2. Candidates who showed midpoints in their working did at least score a method mark.

2d. [3 marks]

## Markscheme

	Drama	Comedy	Film	News
Males	58	119	157	52
Females	86	98	120	61

(M1)(M1)(A1)

[3 marks]

## Examiners report

The chi-squared question was answered well by the majority of candidates and almost all found the chi-squared statistic correctly by GDC, though many could not look up the correct critical value.

2e. [1 mark]

## Markscheme

$H_0$ : favourite TV programme is independent of gender or no association between favourite TV programme and gender

$H_1$ : favourite TV programme is dependent on gender (*must have both*) (A1)

[1 mark]

## Examiners report

The chi-squared question was answered well by the majority of candidates and almost all found the chi-squared statistic correctly by GDC, though many could not look up the correct critical value.

2f. [2 marks]

## Markscheme

$$\frac{365 \times 217}{751} \quad (M1)$$

$$= 105 \quad (A1)(ft)(G2)$$

[2 marks]

## Examiners report

The chi-squared question was answered well by the majority of candidates and almost all found the chi-squared statistic correctly by GDC, though many could not look up the correct critical value.

2g. [3 marks]

## Markscheme

$$12.6 \text{ (accept 12.558)} \quad (G3)$$

[3 marks]

## Examiners report

The chi-squared question was answered well by the majority of candidates and almost all found the chi-squared statistic correctly by GDC, though many could not look up the correct critical value.

2h. [3 marks]

### Markscheme

(i) 3 (AI)

(ii) reject  $H_0$  or equivalent statement (e.g. accept  $H_1$ ) (AI)(ft)

[3 marks]

## Examiners report

The chi-squared question was answered well by the majority of candidates and almost all found the chi-squared statistic correctly by GDC, though many could not look up the correct critical value.

3a. [1 mark]

### Markscheme

$H_0$ : The size of dog is independent of the time of day, (or equivalent) (AI)

Award (A0) for 'no correlation' (C1)

[1 mark]

## Examiners report

- a) Most of the students got the null hypothesis right but quite a few used the word *correlation* instead of *independent*.
- b) Candidates who used a GDC got it correct, while those who tried valiantly to calculate it by hand generally got an M1 but A0.
- c) Most of the students knew how to calculate the degrees of freedom.
- d) Many students did not have a clear idea which values to compare in order to arrive at a conclusion for the chi-squared test. Many compared the significance level with either the chi-squared value or the critical value. Some did not reject the hypothesis but either gave no reason or the wrong one.

As mentioned above, quite a number of candidates did not appear to have been taught this part of the course. There were many non-attempts. It was not a difficult question as indicated by the large number of candidates who scored full marks.

3b. [2 marks]

### Markscheme

$\chi^2 = 4.33$ . (accept 4.328) (MI)(AI)

Note: GDC use is anticipated but candidates might calculate this by hand. (MI) can be awarded for a reasonable attempt to use the formula. (C2)

[2 marks]

## Examiners report

- a) Most of the students got the null hypothesis right but quite a few used the word *correlation* instead of *independent*.
- b) Candidates who used a GDC got it correct, while those who tried valiantly to calculate it by hand generally got an M1 but A0.
- c) Most of the students knew how to calculate the degrees of freedom.
- d) Many students did not have a clear idea which values to compare in order to arrive at a conclusion for the chi-squared test. Many compared the significance level with either the chi-squared value or the critical value. Some did not reject the hypothesis but either gave no reason or the wrong one.

As mentioned above, quite a number of candidates did not appear to have been taught this part of the course. There were many non-attempts. It was not a difficult question as indicated by the large number of candidates who scored full marks.

3c. [1 mark]

## Markscheme

$$(3-1)(3-1) = 4 \quad (A1)$$

Award mark for left hand side seen. (C1)

[1 mark]

## Examiners report

- a) Most of the students got the null hypothesis right but quite a few used the word *correlation* instead of *independent*.
- b) Candidates who used a GDC got it correct, while those who tried valiantly to calculate it by hand generally got an M1 but A0.
- c) Most of the students knew how to calculate the degrees of freedom.
- d) Many students did not have a clear idea which values to compare in order to arrive at a conclusion for the chi-squared test. Many compared the significance level with either the chi-squared value or the critical value. Some did not reject the hypothesis but either gave no reason or the wrong one.

As mentioned above, quite a number of candidates did not appear to have been taught this part of the course. There were many non-attempts. It was not a difficult question as indicated by the large number of candidates who scored full marks.

3d. [2 marks]

## Markscheme

The hypothesis should not be rejected, (allow 'accept  $H_0$ ')  
**OR**

The size of dog is independent of the time of day (A1)(ft)

$$4.33 < 9.488 \text{ or}$$

$$0.363 > 0.05 \quad (R1)(ft)$$

Allow

$\chi_{calc}^2 < \chi_{crit}^2$  only if a value for  $\chi_{calc}^2$  is seen somewhere.

Note: Award (R1)(ft) for comparing the values and (A1)(ft) if the conclusion is valid according to the comparison given. If no reason is given, or if the reason is wrong **both** marks are lost. Note that (A0)(R1)(ft) can be awarded but (A1)(R0) cannot. (C2)

[2 marks]

## Examiners report

- a) Most of the students got the null hypothesis right but quite a few used the word *correlation* instead of *independent*.
- b) Candidates who used a GDC got it correct, while those who tried valiantly to calculate it by hand generally got an M1 but A0.
- c) Most of the students knew how to calculate the degrees of freedom.
- d) Many students did not have a clear idea which values to compare in order to arrive at a conclusion for the chi-squared test. Many compared the significance level with either the chi-squared value or the critical value. Some did not reject the hypothesis but either gave no reason or the wrong one.

As mentioned above, quite a number of candidates did not appear to have been taught this part of the course. There were many non-attempts. It was not a difficult question as indicated by the large number of candidates who scored full marks.

4a. [2 marks]

## Markscheme

(6, 13) (A1)(A1) (C2)

## Examiners report

[N/A]

4b. [2 marks]

## Markscheme

0.952 (  
0.95202... ) (A2) (C2)

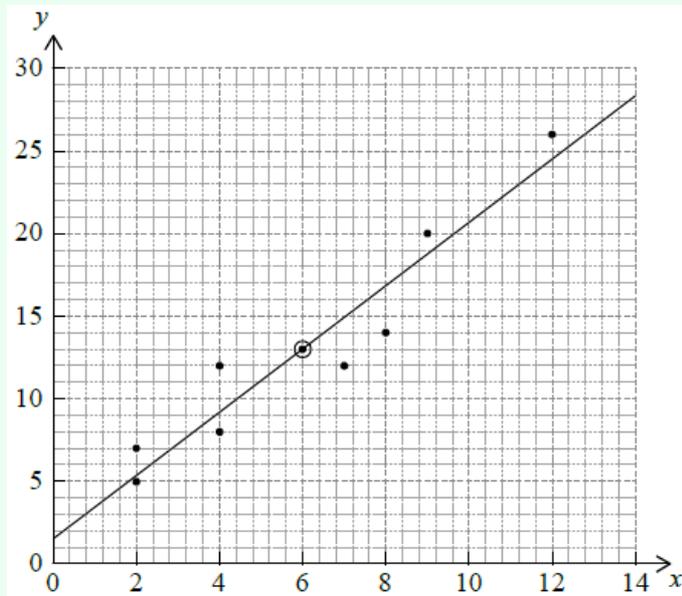
**Note:** Award (A0) for  
0.9.

## Examiners report

[N/A]

4c. [2 marks]

### Markscheme



(A1)  
 y intercept at  
 $y = 1.8$  (accept between 1 and 2)  
 (A1)(ft) line passes through their mean point (A1)(A1)(ft) (C2)

### Examiners report

[N/A]

5a. [2 marks]

### Markscheme

$\frac{50}{120} \times \frac{35}{120} \times 120$  OR  
 $\left(\frac{50 \times 35}{120}\right)$  (M1)  
 $= 14.6$  (  
 14.5833...) (A1) (C2)

### Examiners report

[N/A]

5b. [2 marks]

### Markscheme

0.0746 (A2) (C2)

### Examiners report

[N/A]

5c. [2 marks]

### Markscheme

Since

$p\text{-value} > 5\%$ , the choice of the sport is independent of gender. (RI)(AI)(ft) (C2)

**Note:** The (RI) is awarded for the explicit comparison, the (AI)(ft) is awarded for a consistent conclusion with their answer in part (c).

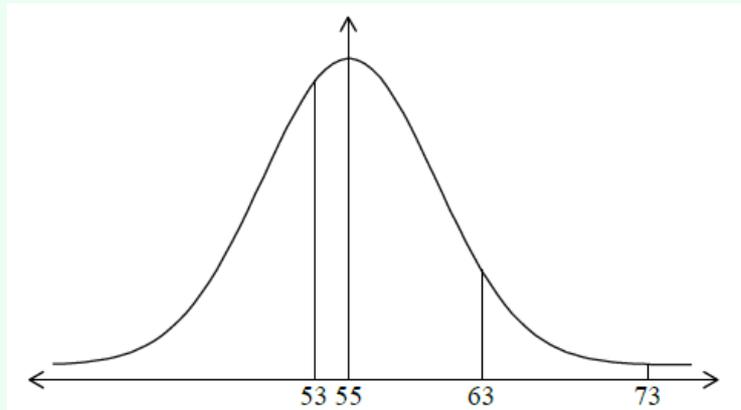
It is therefore possible that (RI)(A0) may be awarded, but (R0)(AI) can never be awarded.

### Examiners report

[N/A]

6a. [3 marks]

### Markscheme



(AI) for normal curve with mean of 55 indicated

(AI) for three lines in approximately the correct position

(AI) for labels on the three lines (AI)(AI)(AI)

### Examiners report

[N/A]

6b. [4 marks]

### Markscheme

(i)

$P(53 \leq \text{Weight} < 63) = 0.486$  (0.485902... ) (MI)(AI)(G2)

**Note:** Award (MI) for correct region indicated on labelled diagram.

(ii)

$P(\text{Weight} > 73) = 0.00506$  (0.00506402) (MI)(AI)(G2)

**Note:** Award (MI) for correct region indicated on labelled diagram.

### Examiners report

[N/A]

6c. [2 marks]

## Markscheme

$$P(\text{Weight} > w) = 0.3 \quad (M1)$$

$$w = 58.7 \quad ($$

$$58.6708\dots) \quad (A1)(G2)$$

**Note:** Award *(M1)* for correct region indicated on labelled diagram.

## Examiners report

[N/A]

6d. [2 marks]

## Markscheme

Expected number of large size eggs

$$= 2000(0.121) \quad (M1)$$

$$= 242 \quad (A1)(G2)$$

## Examiners report

[N/A]

6e. [3 marks]

## Markscheme

Expected income

$$= 2000 \times 0.30 \times 0.388 + 2000 \times 0.50 \times 0.486 + 2000 \times 0.65 \times 0.121 + 2000 \times 0.80 \times 0.00506 \quad (M1)(M1)$$

**Note:** Award *(M1)* for their correct products, *(M1)* for addition of 4 terms.

$$= 884.20 \text{ USD} \quad (A1)(ft)(G3)$$

**Note:** Follow through from part (b).

## Examiners report

[N/A]

7a. [2 marks]

## Markscheme

$$a = -0.134, b = 20.9 \quad (A1)$$

$$y = 20.9 - 0.134x \quad (A1) \quad (C2)$$

[2 marks]

## Examiners report

This question expected the candidates to use their GDC to find the equation of the regression line and to find the correlation coefficient and this was stated in the question parts. However, there were a number of candidates from specific schools who had not been taught to do this and they tried to find the equation from the formula in the information booklet. This would have wasted time and they did not manage to find the correct answer. This was also the case for the correlation coefficient. However, most candidates who found a correlation coefficient managed to comment on it correctly.

7b.

[1 mark]

## Markscheme

17 objects (AI)(ft)

accept only 17 (CI)

[1 mark]

## Examiners report

This question expected the candidates to use their GDC to find the equation of the regression line and to find the correlation coefficient and this was stated in the question parts. However, there were a number of candidates from specific schools who had not been taught to do this and they tried to find the equation from the formula in the information booklet. This would have wasted time and they did not manage to find the correct answer. This was also the case for the correlation coefficient. However, most candidates who found a correlation coefficient managed to comment on it correctly.

Many candidates did not give a whole number answer for the number of objects in part (b).

7c.

[1 mark]

## Markscheme

$r = -0.756$  (AI) (CI)

[1 mark]

## Examiners report

This question expected the candidates to use their GDC to find the equation of the regression line and to find the correlation coefficient and this was stated in the question parts. However, there were a number of candidates from specific schools who had not been taught to do this and they tried to find the equation from the formula in the information booklet. This would have wasted time and they did not manage to find the correct answer. This was also the case for the correlation coefficient. However, most candidates who found a correlation coefficient managed to comment on it correctly.

7d.

[2 marks]

## Markscheme

negative and moderately strong (AI)(ft)(AI)(ft) (C2)

[2 marks]

## Examiners report

This question expected the candidates to use their GDC to find the equation of the regression line and to find the correlation coefficient and this was stated in the question parts. However, there were a number of candidates from specific schools who had not been taught to do this and they tried to find the equation from the formula in the information booklet. This would have wasted time and they did not manage to find the correct answer. This was also the case for the correlation coefficient. However, most candidates who found a correlation coefficient managed to comment on it correctly.

8a. [4 marks]

## Markscheme

(i)  
 $\frac{280}{400}$  (0.7, 70% or equivalent) (AI)(AI)(G2)

**Note:** (AI) for correct numerator, (AI) for correct denominator.

(ii)  
 $\frac{57}{210}$  ( $\frac{19}{70}$ , 0.271, 27.1%) (AI)(AI)(G2)

**Note:** (AI) for correct numerator, (AI) for correct denominator.

[4 marks]

## Examiners report

Candidates answered part (a) correctly. Some lost one out of the 4 marks for making an error in the denominator of the conditional probability. In (b) many students failed to see that (b) was 'without replacement'. Parts (c), (d) and (e) seemed to be very well done by some centres and uniformly badly by others. In (e) many gave the table from the GDC and highlighted the value 63 for which no mark was gained. Expected value formula should have been used instead.

8b. [3 marks]

## Markscheme

$$\frac{180}{400} \times \frac{179}{399} \quad (AI)(MI)$$

**Note:** (AI) for correct values seen, (MI) for multiplying their two values, (AI) for correct answer.

$$= \frac{537}{2660} (= 0.202) \quad (AI)(G3)$$

[3 marks]

## Examiners report

Candidates answered part (a) correctly. Some lost one out of the 4 marks for making an error in the denominator of the conditional probability. In (b) many students failed to see that (b) was 'without replacement'. Parts (c), (d) and (e) seemed to be very well done by some centres and uniformly badly by others. In (e) many gave the table from the GDC and highlighted the value 63 for which no mark was gained. Expected value formula should have been used instead.

8c. [1 mark]

## Markscheme

$H_0$  : 'the preference of brand of cereal is independent of the city'. (AI)

OR

$H_0$  : 'there is no association between the brand of cereal and city'.

[1 mark]

## Examiners report

Candidates answered part (a) correctly. Some lost one out of the 4 marks for making an error in the denominator of the conditional probability. In (b) many students failed to see that (b) was 'without replacement'. Parts (c), (d) and (e) seemed to be very well done by some centres and uniformly badly by others. In (e) many gave the table from the GDC and highlighted the value 63 for which no mark was gained. Expected value formula should have been used instead.

8d. [1 mark]

## Markscheme

$$df = 2 \quad (AI)$$

[1 mark]

## Examiners report

Candidates answered part (a) correctly. Some lost one out of the 4 marks for making an error in the denominator of the conditional probability. In (b) many students failed to see that (b) was 'without replacement'. Parts (c), (d) and (e) seemed to be very well done by some centres and uniformly badly by others. In (e) many gave the table from the GDC and highlighted the value 63 for which no mark was gained. Expected value formula should have been used instead.

8e. [2 marks]

## Markscheme

$$\frac{210 \times 120}{400} \quad (MI)(AI)$$

**Note:** (MI) for substituting in correct formula, (AI) for correct values.

$$= 63 \quad (AG)$$

**Note:** Final line must be seen or previous (AI) mark is lost.

[2 marks]

## Examiners report

Candidates answered part (a) correctly. Some lost one out of the 4 marks for making an error in the denominator of the conditional probability. In (b) many students failed to see that (b) was 'without replacement'. Parts (c), (d) and (e) seemed to be very well done by some centres and uniformly badly by others. In (e) many gave the table from the GDC and highlighted the value 63 for which no mark was gained. Expected value formula should have been used instead.

8f. [2 marks]

## Markscheme

$$39.3 \quad (G2)$$

**Note:** Award (GI)(A0)(AP) if answers not to 3 significant figures.

[2 marks]

## Examiners report

Candidates answered part (a) correctly. Some lost one out of the 4 marks for making an error in the denominator of the conditional probability. In (b) many students failed to see that (b) was 'without replacement'. Parts (c), (d) and (e) seemed to be very well done by some centres and uniformly badly by others. In (e) many gave the table from the GDC and highlighted the value 63 for which no mark was gained. Expected value formula should have been used instead.

8g.

[2 marks]

## Markscheme

$p$  – value  $< 0.05$  (RI)(ft)

Do not accept

$H_0$  . (AI)(ft)

**Notes:** Allow 'Reject

$H_0$  or equivalent'. (ft) from their

$\chi^2$  statistic.

Award (RI)(ft) for comparing the appropriate values. (AI)(ft) can be awarded only if the conclusion is valid according to the comparison given. If no reason given or if reason is wrong both marks are lost. Note that (RI)(A0)(ft) can be awarded but (R0)(AI)(ft) cannot.

[2 marks]

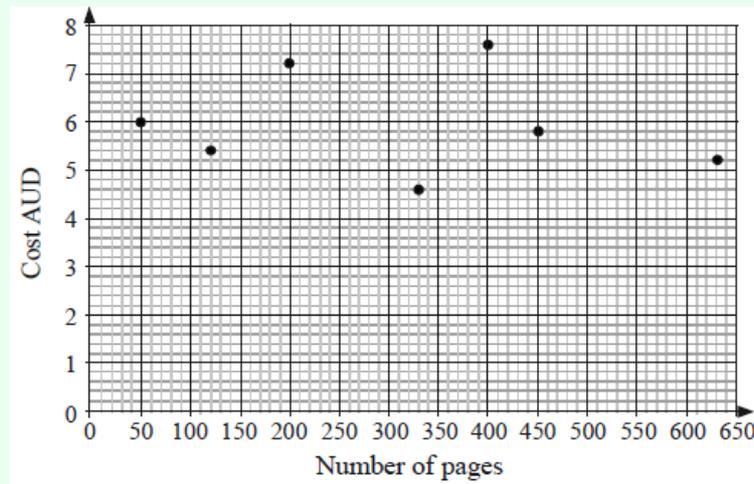
## Examiners report

Candidates answered part (a) correctly. Some lost one out of the 4 marks for making an error in the denominator of the conditional probability. In (b) many students failed to see that (b) was 'without replacement'. Parts (c), (d) and (e) seemed to be very well done by some centres and uniformly badly by others. In (e) many gave the table from the GDC and highlighted the value 63 for which no mark was gained. Expected value formula should have been used instead.

8h.

[3 marks]

## Markscheme



(AI)(AI)(AI)

**Notes:** (AI) for label and scales, (A2) for all points correct, (AI) for 5 or 6 correct. Award a maximum of (A2) if points are joined.

[3 marks]

## Examiners report

The graph was well done with almost all candidates labelling and scaling the axes correctly. A minority of students joined the points or drew the graph on lined paper which prevented them from gaining full marks in this part of the question.

In (b) some candidates were not able to calculate the linear correlation coefficient. A few G2 comments pointed out that the command term used may have been ambiguous to some candidates and they did not think that they could use their GDC to find  $r$ . Some attempted to use the formula even though the value of  $S_{xy}$  was not given. The guide says that 'A GDC can be used to calculate  $r$  when raw data is given'. This potential unfairness was taken into consideration during the setting of boundaries so that no candidate was disadvantaged by the possible ambiguous wording of the question. In future the command term 'Using your GDC' or 'Write down' will be used in similar questions.

Some students who did use the GDC gave

$r^2$  instead of

$r$ . This really caught the attention of many examiners as

$r^2$  is not in the syllabus.

8i.

[2 marks]

## Markscheme

$$r = -0.141 \quad (G2)$$

**Note:** If negative sign is missing award  $(G1)(G0)$ .

[2 marks]

## Examiners report

The graph was well done with almost all candidates labelling and scaling the axes correctly. A minority of students joined the points or drew the graph on lined paper which prevented them from gaining full marks in this part of the question.

In (b) some candidates were not able to calculate the linear correlation coefficient. A few G2 comments pointed out that the command term used may have been ambiguous to some candidates and they did not think that they could use their GDC to find  $r$ . Some attempted to use the formula even though the value of

$S_{xy}$  was not given. The guide says that 'A GDC can be used to calculate  $r$  when raw data is given'. This potential unfairness was taken into consideration during the setting of boundaries so that no candidate was disadvantaged by the possible ambiguous wording of the question. In future the command term 'Using your GDC' or 'Write down' will be used in similar questions.

Some students who did use the GDC gave

$r^2$  instead of

$r$ . This really caught the attention of many examiners as

$r^2$  is not in the syllabus.

8j.

[2 marks]

## Markscheme

'The coefficient of correlation is too low, (very) weak (linear) relationship'.  $(R1)$

Not a sensible thing to do, *accept 'no'*.  $(A1)$

**Note:** Do not award  $(R0)(A1)$ . The correlation coefficient has to be mentioned in their reasoning.

[2 marks]

## Examiners report

The graph was well done with almost all candidates labelling and scaling the axes correctly. A minority of students joined the points or drew the graph on lined paper which prevented them from gaining full marks in this part of the question.

In (b) some candidates were not able to calculate the linear correlation coefficient. A few G2 comments pointed out that the command term used may have been ambiguous to some candidates and they did not think that they could use their GDC to find  $r$ . Some attempted to use the formula even though the value of

$S_{xy}$  was not given. The guide says that 'A GDC can be used to calculate  $r$  when raw data is given'. This potential unfairness was taken into consideration during the setting of boundaries so that no candidate was disadvantaged by the possible ambiguous wording of the question. In future the command term 'Using your GDC' or 'Write down' will be used in similar questions.

Some students who did use the GDC gave

$r^2$  instead of

$r$ . This really caught the attention of many examiners as

$r^2$  is not in the syllabus.