UNIVERSITY OF CAMBRIDGE INTERNATIONAL EXAMINATIONS
General Certificate of Education Ordinary Level

CANDIDATE NAME

CENTRE NUMBER

CANDIDATE NUMBER

STATISTICS **4040/11**

Paper 1 **October/November 2010**

2 hours 15 minutes

Candidates answer on the question paper.

Additional Materials: Mathematical tables
Pair of compasses
Protractor

READ THESE INSTRUCTIONS FIRST

Write your Centre number, candidate number and name on all the work you hand in.
Write in dark blue or black pen.
You may use a soft pencil for any diagrams or graphs.
Do not use staples, paper clips, highlighters, glue or correction fluid.
DO **NOT** WRITE IN ANY BARCODES.

Answer **all** questions in Section A and not more than **four** questions from Section B.
If working is needed for any question it must be shown below that question.
The use of an electronic calculator is expected in this paper.

At the end of the examination, fasten all your work securely together.
The number of marks is given in brackets [ ] at the end of each question or part question.

This document consists of **19** printed pages and **1** blank page.

Section A [36 marks]

Answer **all** of the questions 1 to 6.

1 In a class test, the marks of eleven pupils were as follows.

4 25 7 8 4 4 8 9 6 5 19

(i) For these marks, find

(a) the mode,

.....[1]

(b) the median,

.....[1]

(c) the mean.

.....[1]

(ii) State which one of these three quantities you would choose as the most appropriate measure of central tendency (average) to represent these figures, and give a reason for rejecting **one** of the other two.

.....

.....[2]

(iii) State (but do **not** calculate) which measure of dispersion (spread) you would use as the most appropriate to represent these marks.

.....[1]

2 There are 34 pupils in a school class. Their method of travel to school is given in the following table.

Method of travel	*Number of pupils*
Car	7
Bus	10
Walk	7
Train	10

The data is to be illustrated by a pie chart of radius 4 cm.

(i) Calculate, **to the nearest degree**, the angles of the sectors representing Car and Bus.

Car ..... °

Bus ..... ° [2]

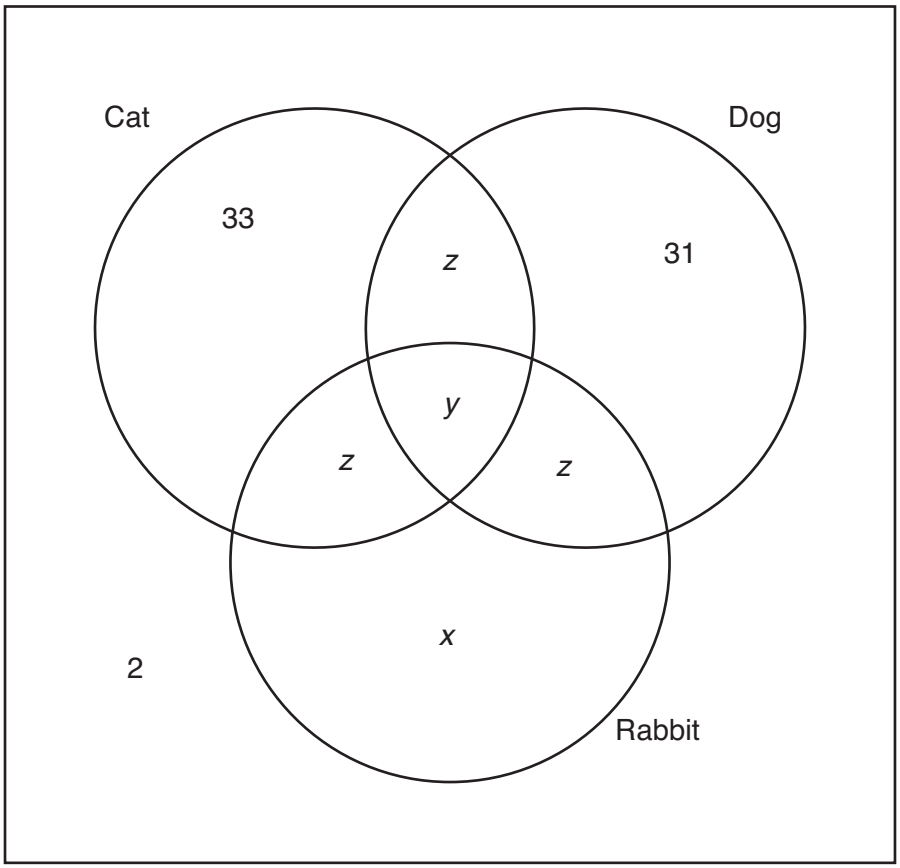
(ii) Draw and label the pie chart.

[2]

(iii) In another class there are 27 pupils. Calculate, correct to 1 decimal place, the radius of a comparable pie chart to illustrate this class. **You are NOT required to draw this pie chart.**

..... cm [2]

3 The diagram below shows how many pets (cat, dog, rabbit) are owned by each of 100 households. No household owns more than one of each different animal.



(i) Interpret the value 31 in the diagram.

.....[1]

(ii) 72 households own only one pet. Find the value of x.

$x =$ .....[1]

(iii) Three times as many households own only a cat as own all three types of pet. Find the value of y.

$y =$ .....[1]

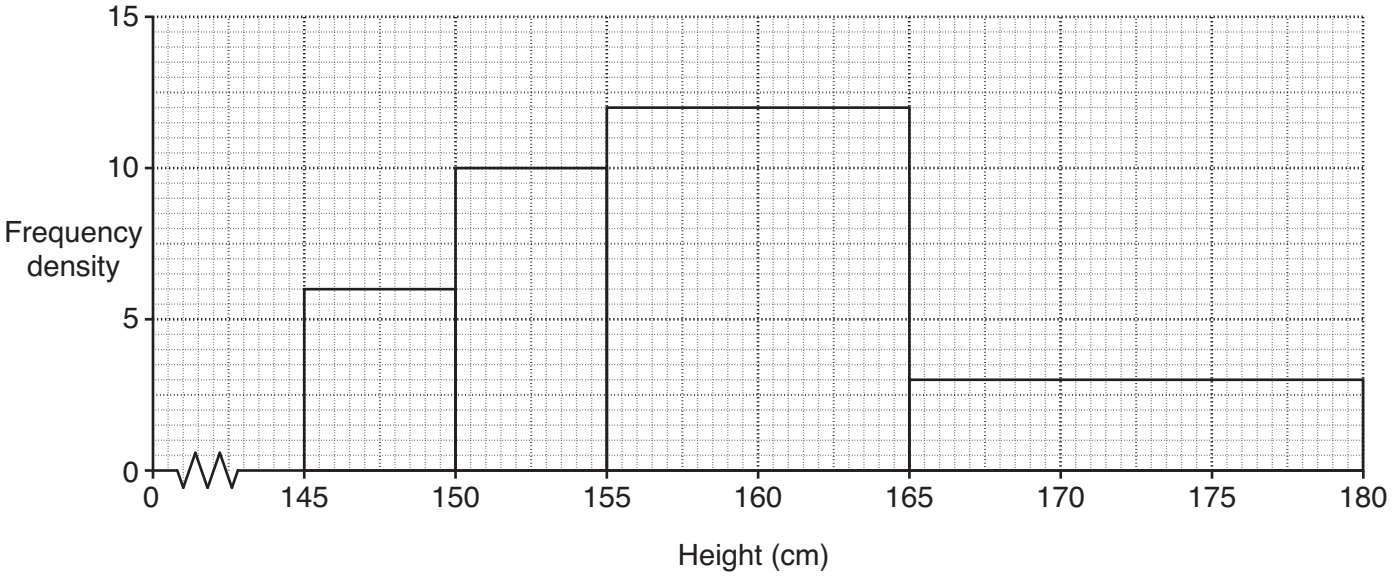
(iv) Of the 100 households, 2 own no pet. Find the value of z.

$z =$ .....[2]

(v) Find the number of households owning a dog.

.....[1]

4 The heights of the girls in a village were measured and are represented in the histogram below.



There are 24 girls in the 155 – under 165 cm class.

(i) Complete the table below.

Height (cm)	*Number of girls*
145 – under 150	
150 – under 155	
155 – under 165	24
165 – under 180	

[4]

(ii) Use the histogram to estimate the modal height of the girls.

..... cm [2]

5 Before an examination, teachers were required to predict the grades which their pupils would obtain, from *A* (highest) to *D* (lowest). The following table shows the numbers of pupils for whom the predicted grade was accurate.

Teacher's predicted grade	*Grade obtained*			
	A	*B*	*C*	*D*
A	132			
B		284		
C			203	
D				68

(i) Complete the table using the following information:

(a) Among pupils for whom grade *C* was predicted, 33 obtained grade *B* and 4 obtained grade *D*.

[1]

(b) There were no pupils for whom the predicted grade and the grade obtained were more than one grade apart.

[1]

(c) Of the candidates for whom grade *D* was predicted, one-third obtained grade *C*.

[1]

(d) Of the candidates for whom grade *A* was predicted, eleven times as many obtained grade *A* as obtained grade *B*.

[1]

(e) There were 50 candidates for whom grade *B* was predicted but who obtained another grade. Nine times as many obtained grade *A* as obtained grade *C*.

[1]

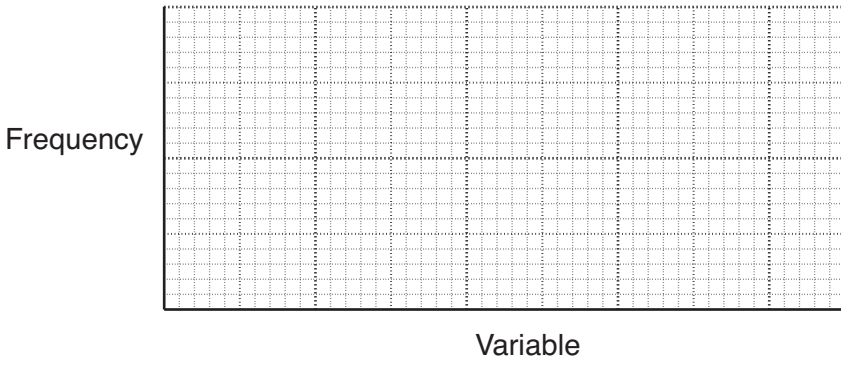
(ii) For all those pupils whose predicted grade was not accurate, state, with a reason, whether the grade obtained was likely to be higher or lower than the predicted grade.

.....

.....[1]

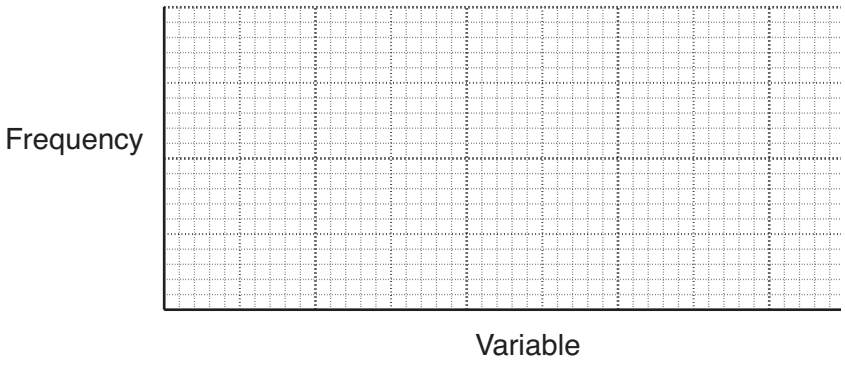
6 **(a)** On the grids below sketch the frequency curve of

(i) a distribution in which the mode is greater than the mean,



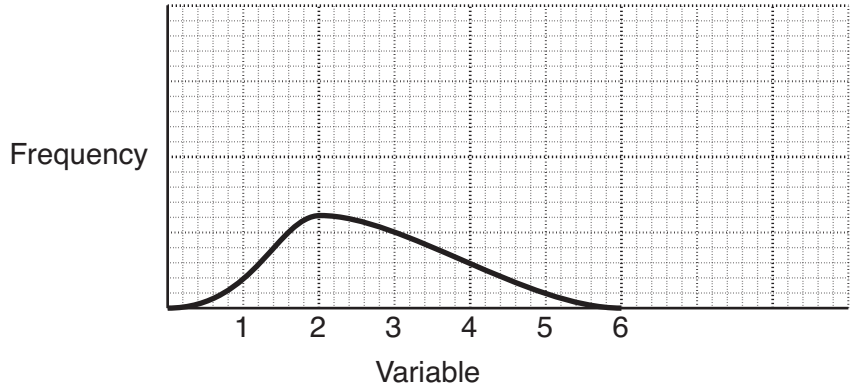
[2]

(ii) a distribution in which the mean, median and mode are all equal.

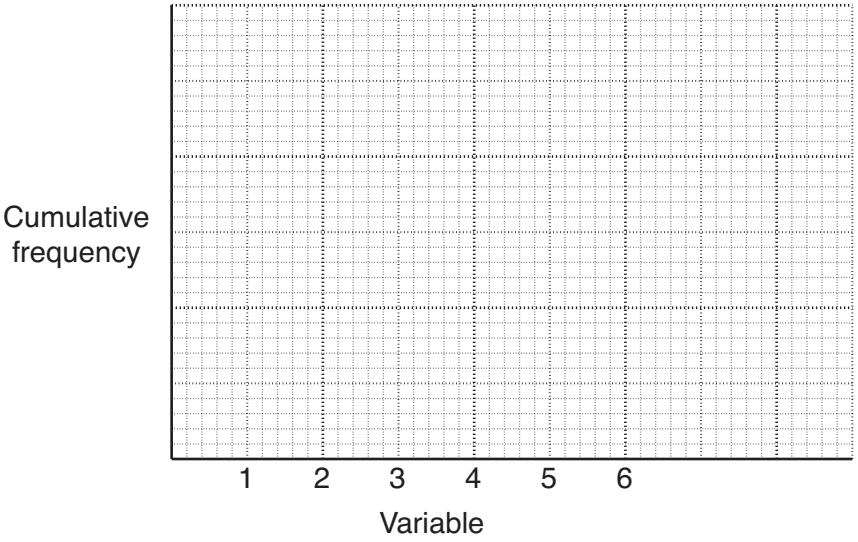


[2]

(b) A variable has the following frequency curve.



On the grid below sketch the corresponding cumulative frequency curve.



[2]

Section B [64 marks]

Answer not more than **four** of the questions 7 to 11.

Each question in this section carries 16 marks.

7 **(a)** A man has three unbiased coins, one gold, one silver and one bronze.
In an experiment, the gold coin is tossed. If a head lands uppermost the experiment stops, otherwise the silver coin is tossed. If a head on the silver coin lands uppermost the experiment stops, otherwise the bronze coin is tossed.
On each toss of a coin, if a head lands uppermost, H is recorded, and if a tail lands uppermost, T is recorded.

(i) State each of the possible sequences of outcome in the experiment together with its probability.

[4]

(ii) The experiment is carried out twice. Calculate the probability that both experiments stop with H being recorded.

.....[2]

(b) John and Zaheer play a game in which the player who starts the game has an advantage. When John starts, the probability that he will win is 0.7. When Zaheer starts, the probability that he will win is 0.8. All games end in a win for one player or the other.
They agree to play a number of games until one of them has won 3 games. Except for the first game, **each game is started by the loser of the previous game**.

(i) State the maximum number of games which have to be played for one player to win three games.

.....[1]

(ii) Calculate the probability that John will win by 3 games to 0 when

(a) John starts the first game,

.....[2]

(b) Zaheer starts the first game.

.....[2]

(iii) Denoting the event 'John wins' by J, and 'Zaheer wins' by Z

(a) write down the three different sequences of four results which lead to Zaheer winning by 3 games to 1,

[2]

(b) calculate the probability that Zaheer wins by 3 games to 1, **if John starts the first game.**

[3]

8 The table below gives certain information about the population and deaths in town *X* for the year 2008, together with the standard population of the area in which town *X* is situated.

Age group	*Deaths*	*Population in age group*	*Standard population (%)*
0 – under 20	11	2600	25
20 – under 40	*d*	4500	30
40 – under 60	11	4400	25
60 and over	65	1300	20

(i) The death rate for the 20 – under 40 age group was 2 per thousand. Show that $d = 9$.

[1]

(ii) Calculate the death rates per thousand for the other three age groups.

.....

.....

.....[2]

(iii) Calculate the crude death rate per thousand for town *X*.

.....[4]

(iv) Using the given rate for the 20 – under 40 age group, and the rates you have calculated in **(ii)**, calculate, correct to 2 decimal places, the standardised death rate per thousand for town *X*.

.....[4]

The table below gives information about town *Y*, another town in the same area as town *X*, for the year 2008. The crude death rate for town *Y* in 2008 was 7.97 per thousand.

Age group	*Death rate per thousand*	*Population in age group (%)*
0 – under 20	4	35
20 – under 40	3	24
40 – under 60	5	22
60 and over	25	19

(v) Calculate the standardised death rate per thousand for town *Y* in the year 2008, using the same standard population as for town *X*.

.....[2]

(vi) State, with a reason, which of the two towns would appear to have a healthier environment.

.....

.....[2]

For both town *X* and town *Y*, the standardised death rate is higher than the crude death rate.

(vii) State what this tells you about how their populations compare with the standard population.

.....

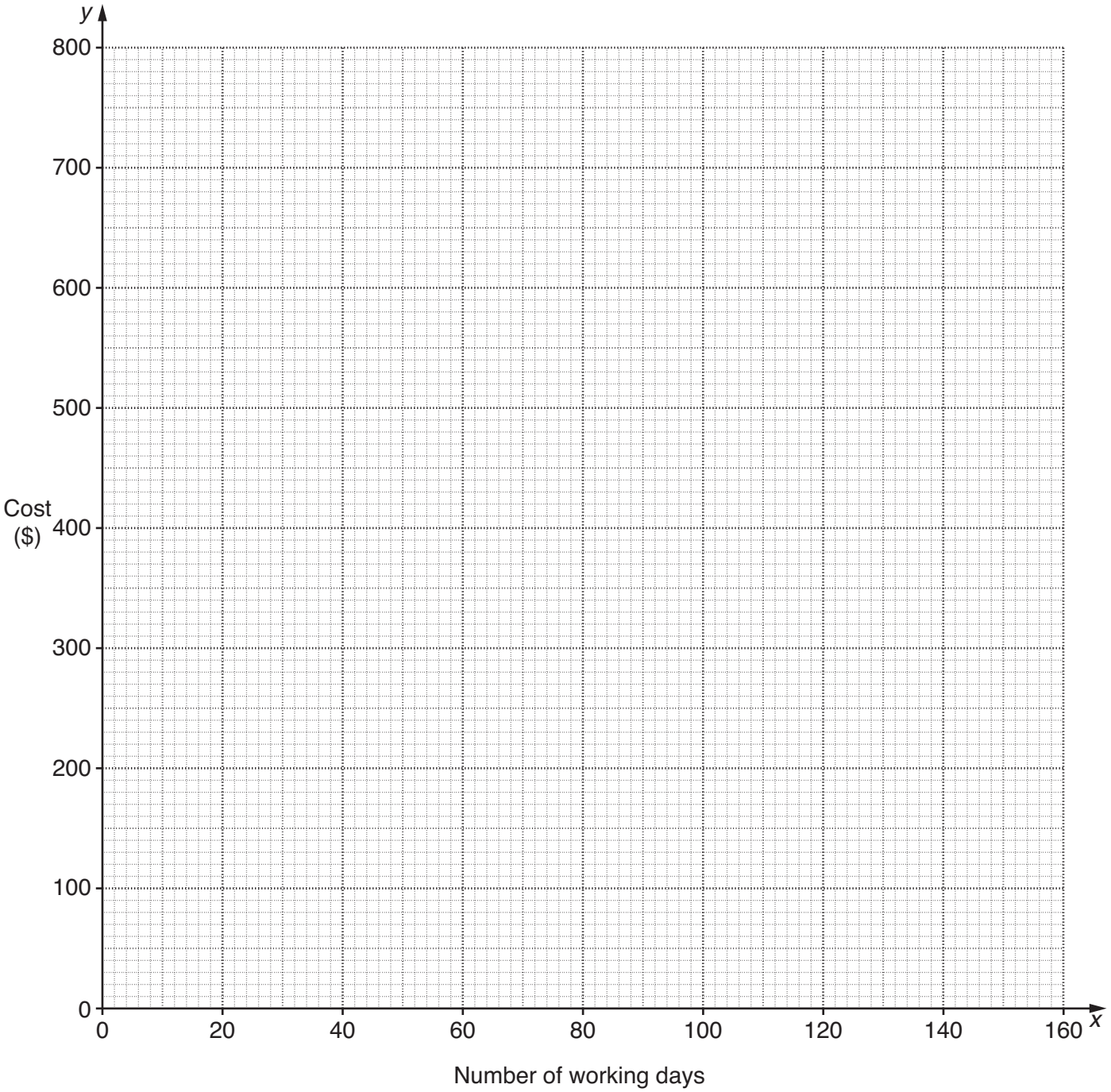
.....[1]

[Turn over

9 A man travels to and from work by car. The following table shows the cost (in $) over a period of 160 working days.

Number of working days (x)	20	40	60	80	100	120	140	160
Cost ($) (y)	230	290	340	390	445	500	560	610

(i) Plot a scatter diagram of cost against number of working days on the grid below.



[2]

(ii) Calculate three appropriate averages using the data, plot them on the grid, and use them to draw the line of best fit.

.....

.....

.....[6]

(iii) Obtain the equation of your line of best fit, giving the equation in the form $y = mx + c$.

.....[4]

(iv) Estimate the cost of travelling to and from work for 110 days.

$ .....[1]

(v) If the man used public transport, the journey to and from work would cost him $5 per day. On the grid, draw a line representing public transport cost against the number of working days.

[1]

(vi) State the number of days after which the man would have spent the same amount whether travelling by car or public transport.

.....[1]

(vii) Explain briefly why one of your lines passes through the origin, but the other does not.

.....

.....[1]

10 The owner of a travelling shop noted how many kilometres he travelled on each of the 300 days on which he worked during one particular year. He grouped his results as given in the following table.

Kilometres travelled	*Number of days*	*Class mid-points (m)*	$x = \frac{(m-25)}{5}$
under 10	6		
10 – under 20	13		
20 – under 30	62		
30 – under 40	96		
40 – under 60	65		
60 and over	58		

The lowest and highest classes in the above table are 'open-ended'.

(i) State, in each case with a reason, what you would regard as a suitable value to use for

(a) the lower limit of the lowest class,

.....

.....[2]

(b) the upper limit of the highest class.

.....

.....[2]

(ii) Using your answers to **(i)**, write **in the above table** the class mid-points (*m*) of the six classes. [2]

(iii) The values of m are scaled, by subtracting 25 and dividing by 5, to give values of x.

(a) Write the values of x **in the table**. [2]

(b) Estimate, to 3 significant figures, the mean and the standard deviation of the values of x.

Mean = ...

Standard deviation = ...[5]

(c) Use your results in **(iii)(b)** to estimate, to 3 significant figures, the mean and the standard deviation of the number of kilometres travelled per day by the owner of the shop.

Mean = ... km

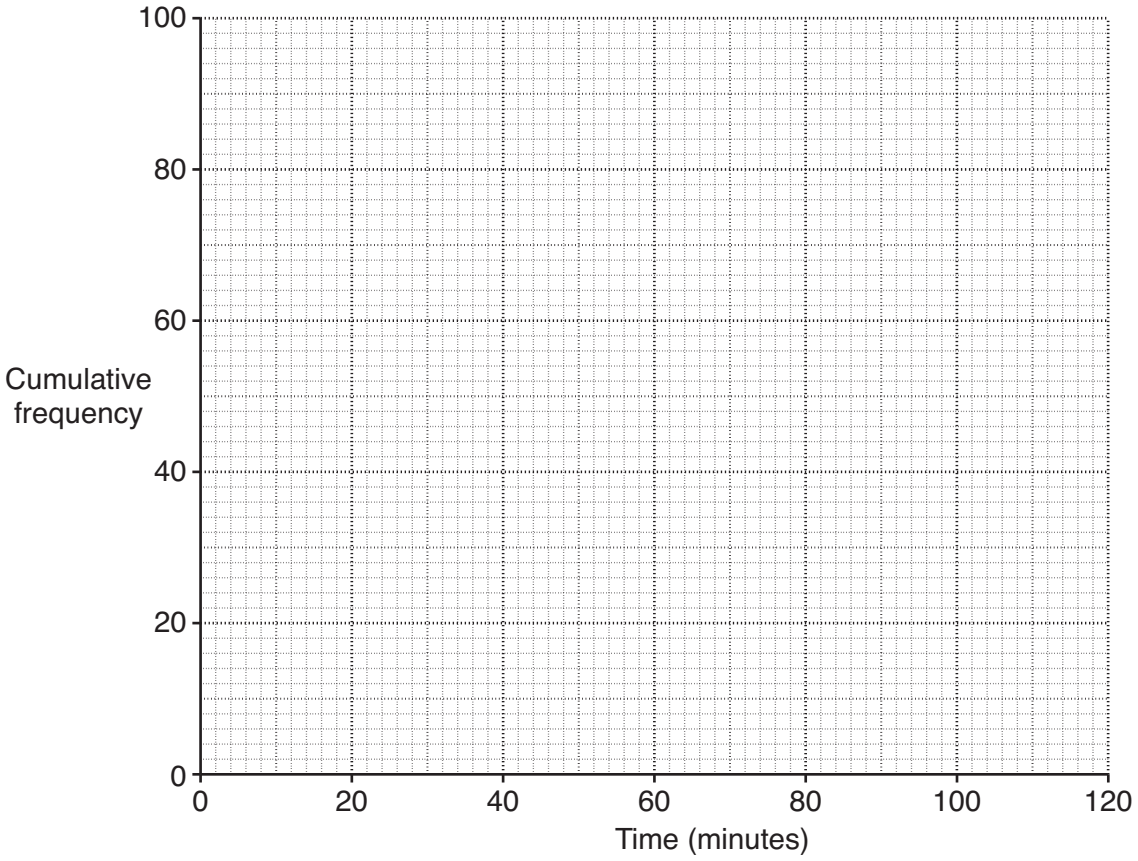
Standard deviation = ... km [3]

11 A company designed an aptitude test to gauge the suitability of people applying to join its workforce. The following table summarises the time taken (in minutes) to complete the test by 88 applicants.

Time to complete test (minutes)	*Number of applicants*	*Cumulative frequency*
10 – under 30	8	
30 – under 40	12	
40 – under 45	14	
45 – under 50	13	
50 – under 60	16	
60 – under 80	16	
80 – under 110	9	

(i) Complete the cumulative frequency column in the above table. [2]

(ii) Plot the cumulative frequencies on the grid below, joining the points by a smooth curve.



[3]

(iii) Use your graph to estimate the median of the times taken to complete the test.

..... minutes [2]

Applicants who complete the test in under 37 minutes are invited to attend an interview.

(iv) Use your graph to estimate the percentage of these 88 applicants who were **not** invited for an interview.

..... % [4]

In addition to the 88 applicants referred to in the table, 11 others had failed to complete the test after 110 minutes and were stopped from continuing with it.

(v) (a) Explain why, if these 11 applicants are included in the data, it is still possible to estimate the interquartile range of the times taken to complete the test by all 99 applicants.

.....

.....

.....

.....[2]

(b) Use your graph to estimate this interquartile range.

..... minutes [3]

BLANK PAGE