O

UNIVERSITY OF CAMBRIDGE INTERNATIONAL EXAMINATIONS
General Certificate of Education Ordinary Level

CANDIDATE NAME	

CENTRE NUMBER		CANDIDATE NUMBER	

4555331710

STATISTICS **4040/01**

Paper 1 **October/November 2009**

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 **17** printed pages and **3** blank pages.

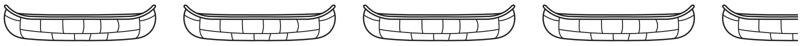
DC (SC/SLM) 11521/8

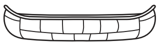
UNIVERSITY *of* CAMBRIDGE
International Examinations

Section A [36 marks]

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

1 A small boatyard makes canoes. The pictogram below shows the number of canoes made in the years 2002 and 2005.

Year	Number of canoes
2002	
2005	

 represents 10 canoes.

(i) State how many canoes were made in

(a) the year 2002,

....[1]

(b) the year 2005.

....[1]

(ii) Draw a pictogram to represent the 33 canoes which the boatyard made in the year 2008.

[2]

2 A large-scale survey is to be carried out, with the required information being collected by means of a questionnaire.

(i) Give **two** purposes of firstly conducting a small-scale survey using a pilot questionnaire.

Purpose 1 ....

....

Purpose 2 ....

....[2]

(ii) Give one advantage and one disadvantage of sending the questionnaire to respondents through the post, rather than using interviewers to ask the questions on it.

Advantage ....

....

Disadvantage ....

....[2]

3 In a large school a daily record was kept, from the start of each month, of the total number of absences **up to and including that day.**
The totals recorded for the first 11 school days of one month were as follows.

12	23	34	43	51	65	77	94	111	123	140

(i) Write the data in a form showing the number of absences on each individual day.

....[2]

(ii) Obtain the mean and the median of the data you have written in **(i)**.

Mean ....

Median ....[3]

(iii) Explain why there is a problem with stating a value for the mode of the data you have written in **(i)**.

....

....[1]

4 Five competitors in a quiz are isolated from each other. They are asked, in a random order, the same question, until one of them gives the correct answer.
At this point the questioning stops. Two of them know the correct answer, and three do not.

(i) State the maximum number of times the question would have to be asked.

....[1]

(ii) If N is the number of times the question is asked,

(a) state the possible values of N,

.... [1]

(b) calculate the probability of each value, presenting your results in a suitable table.

[4]

5 In a large factory the number of repair jobs reported during each of the three daily shifts (Early, Late and Night) was recorded over a three week period. The following table shows some of the results.

Week	Shift			Total
	Early	Late	Night	
1		34		
2	34		18	
3		32		84
Total			44	250

(i) The **total** number of repair jobs reported during the Early and Late shifts was the same. Use this information to insert two values into the table above. [1]

(ii) In week 1 and week 3 there were equal numbers of repair jobs reported during the Night shift. Use this information to insert a further two values into the table above. [1]

(iii) Showing all your working, calculate the remaining five values and insert them into the table above.

[5]

6 The table below gives, by category of employment in the year 2007, the number of employees, and the number of accidents they suffered, at a particular factory. It also shows the standard population for the whole industry of which the factory is part.

Category of employment	*Number of accidents*	*Number of employees*	*Standard population* (%)
Management	3	20	4
Technical	9	75	17
Skilled	18	230	41
Unskilled	25	115	38
Total	55	440	100

(i) Calculate the crude accident rate per thousand for the factory.

....[3]

(ii) Showing your full working for at least one of the categories, calculate the standardised accident rate per thousand for the factory.

....[5]

(iii) Explain briefly why the answers you have obtained to **(i)** and **(ii)** are not equal.

....

....[1]

BLANK PAGE

Section B starts on page 8

Section B [64 marks]

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

Each question in this section carries 16 marks.

7 The lengths of a certain engineering component coming off a production line are measured.

(i) State whether the variable 'length' is

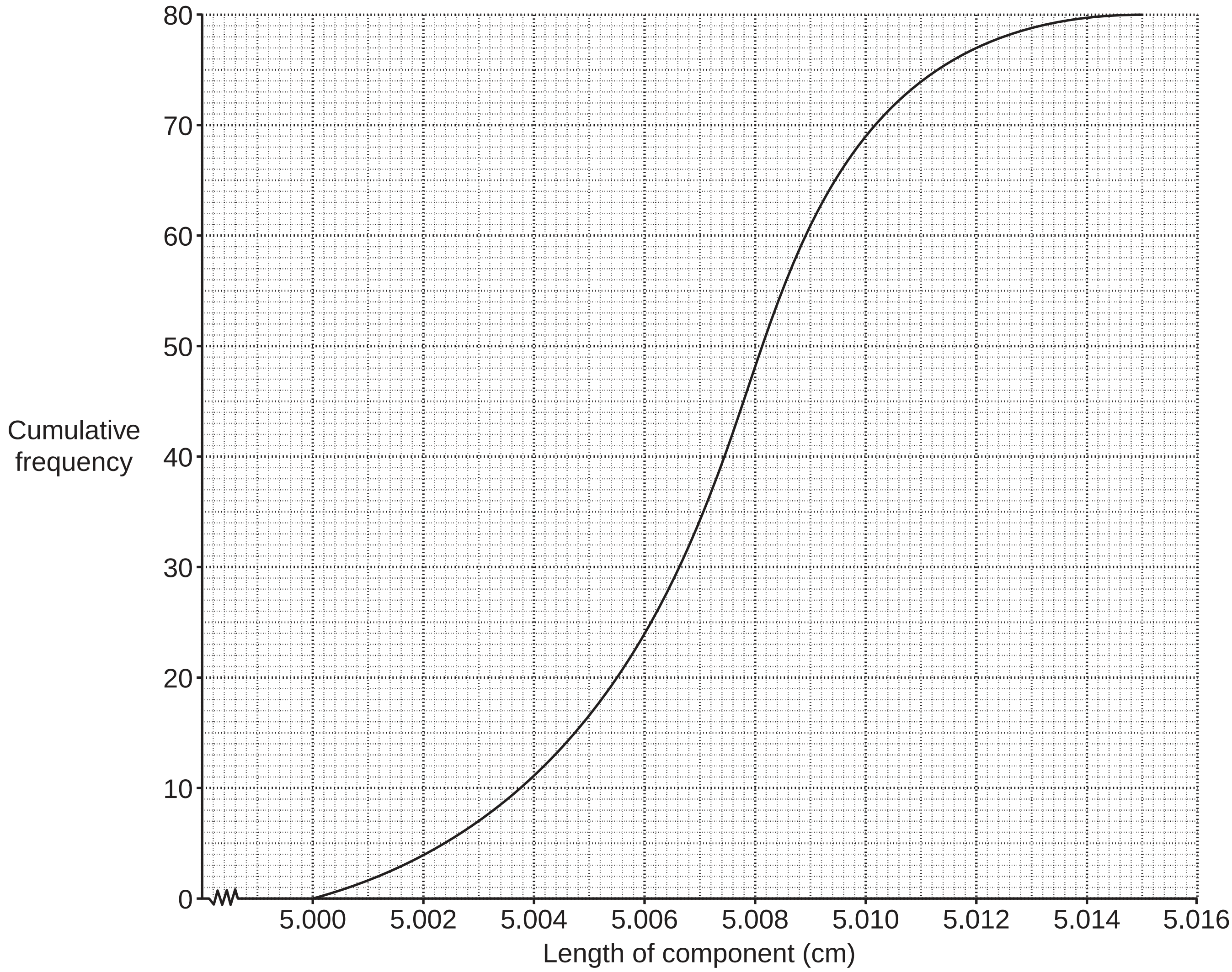
(a) qualitative or quantitative,

....[1]

(b) discrete or continuous.

....[1]

The length of each component is intended to be exactly 5.008 cm. Accurate measurements of a sample of 80 of these components from the production line produced the following cumulative frequency curve.



(ii) The graph is much steeper close to the intended length than it is at either end.

What does this tell you about the precision of the components produced?

....

....[2]

(iii) Use the graph to estimate, to 4 decimal places,

(a) the median of the lengths,

.... cm [1]

(b) the lower quartile of the lengths,

.... cm [1]

(c) the 90th percentile of the lengths.

.... cm [2]

(iv) Estimate the number of components which have a length of more than 5.0084 cm.

....[2]

Components shorter than 5.004 cm, or with length 5.012 cm or more, are rejected.

(v) Estimate from the graph how many components in this sample are **accepted**.

....[3]

(vi) Use the graph to estimate, to 4 decimal places, **for the accepted components** only,

(a) the median of the lengths,

.... cm [2]

(b) the lower quartile of the lengths.

.... cm [1]

8 **(a)** A turn in a game consists of rolling an unbiased six-sided dice, with faces numbered 1 to 6. If the face landing uppermost is 1, 2, 3, 4 or 5, then that number is the score for the turn. If the 6-face lands uppermost, the dice is rolled again, and the score for the turn is the sum of 6 and the number which lands uppermost on the second roll.

(i) List the scores which it is possible to obtain in a turn.

....[1]

(ii) Calculate, as a fraction, the probability of obtaining the lowest possible score in a turn.

....[1]

(iii) Calculate, as a fraction, the probability of obtaining the highest possible score in a turn.

....[1]

(iv) Calculate the probability that, in two consecutive turns, a player will obtain the lowest possible score and the highest possible score, **in either order**.

....[2]

(b) You are one of a class of 10 students. A bag contains one blue disc and 9 red discs.
The students take it in turn to draw a disc at random from the bag **without replacement**.
The student who draws the blue disc will win a prize.
You have been given the choice of drawing a disc first, fourth or last.

Calculate the probability of winning the prize for each of these three alternatives, and hence state which one you would choose.

....[5]

(c) Box *A* contains 7 green balls and 3 white balls. Box *B* contains 5 green balls and 9 white balls.

A ball is selected at random from Box *A* and placed in Box *B*. A ball is **then** selected at random from Box *B* and placed in Box *A*.

Calculate the probability that, after these two operations, the numbers of green and white balls in Box *A* and in Box *B* are the same as at the start.

....[6]

9 A new pressure gauge has been fitted onto a piece of equipment in a chemical plant.
Past experience has shown that such gauges may need to have their calibration (scale of measurement) adjusted.
To determine whether any adjustments are necessary on the new gauge, several pressure readings are taken using both it and a standard gauge which is known to be accurate. All readings are in units of megapascals (MPa).

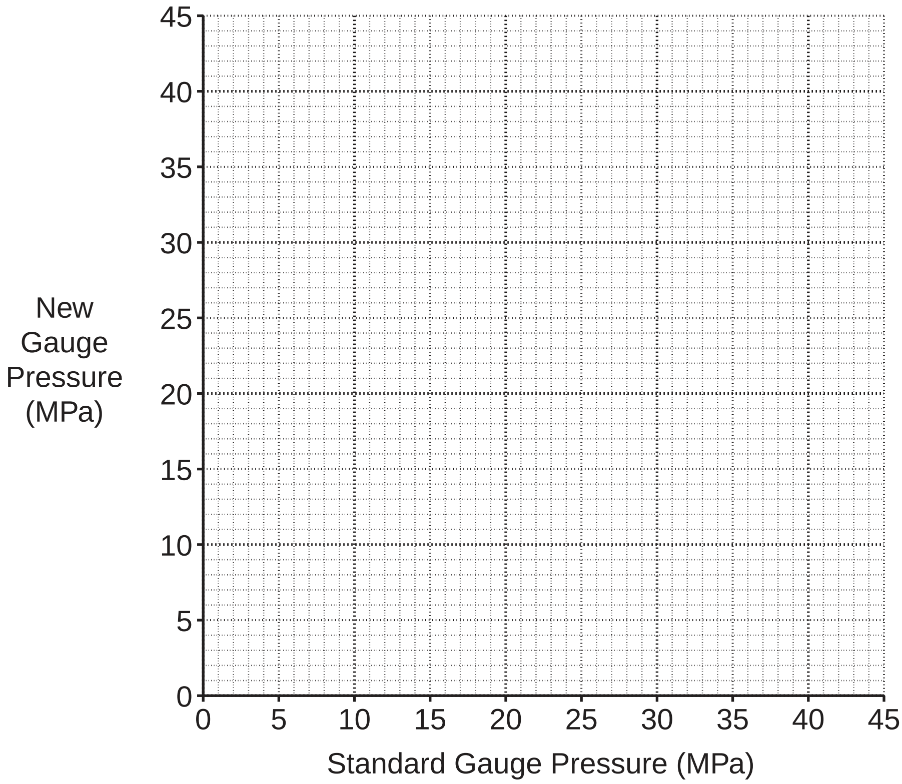
Reading	*Pressure* (MPa)	
	Standard gauge (x)	New gauge (y)
A	4	7
B	38	41
C	7	13
D	11	16
E	15	21
F	27	33
G	31	36
H	35	41

(i) Briefly explain why the standard gauge readings, rather than the new gauge readings, are plotted on the x-axis.

....

....[1]

(ii) Draw a scatter diagram of the data on the grid below.



[3]

(iii) Explain why you should use readings *A*, *C*, *D* and *E* to calculate one semi-average, and *B*, *F*, *G* and *H* to calculate the other semi-average.

....

....

....[1]

(iv) Calculate the overall mean and the two semi-averages, and plot them on your graph.

[5]

(v) Draw a line of best fit through your plotted averages. [1]

(vi) Either by calculation, or by using your graph, find the equation of the line of best fit, and write it in the form $y = \mathrm{m}x + \mathrm{c}$.

$y =$ ....[3]

(vii) Explain what the values of m and c in your equation tell you about any necessary adjustments to the new gauge.

....

....

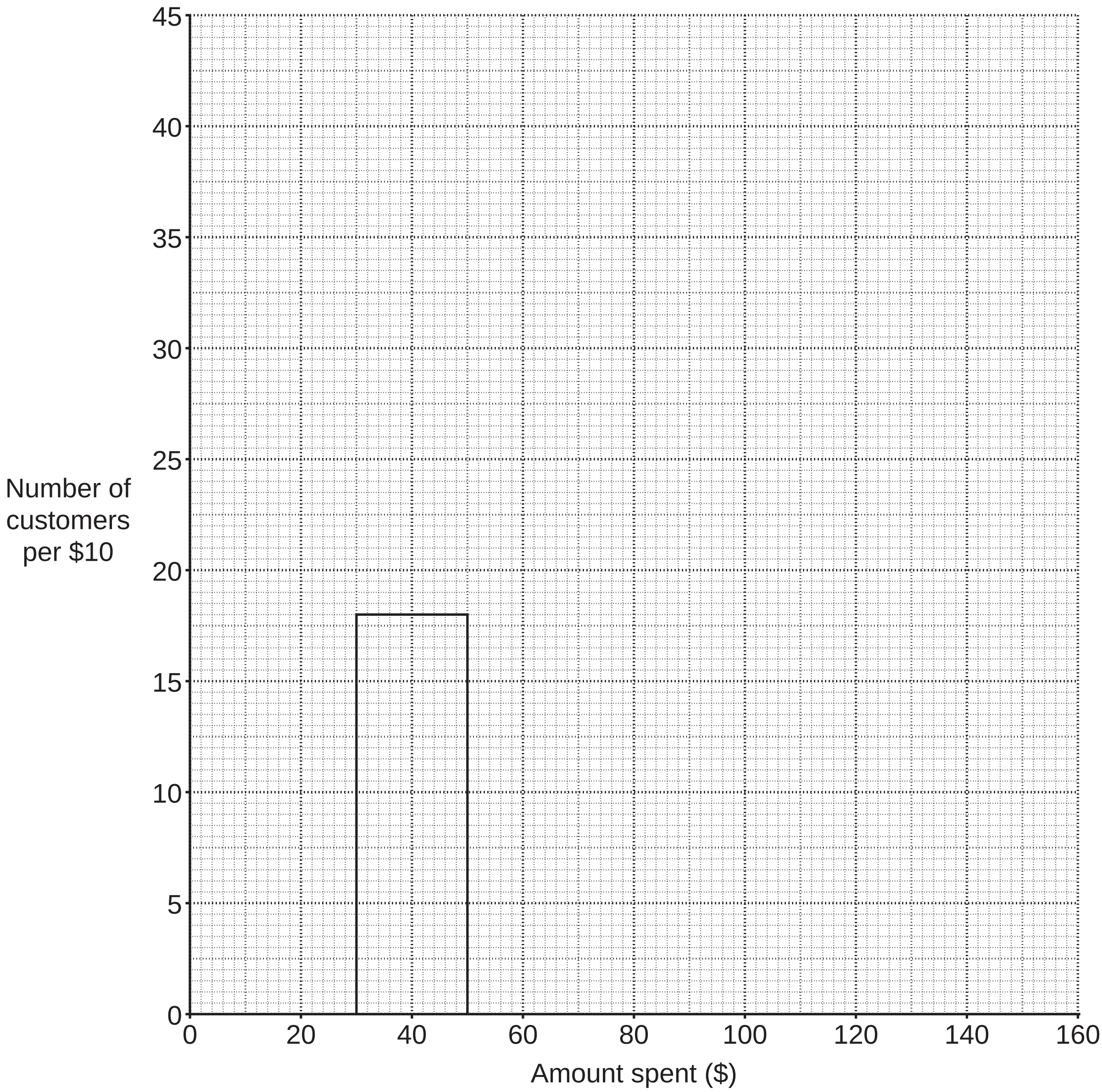
....[2]

10 A random sample of 200 customers leaving a supermarket during one day was asked how much money (in $) and how much time (in minutes) they had spent during their visit.

(i) The following table summarises the amounts spent.

Amount spent ($)	*Number of customers*
0 – under 30	18
30 – under 50	36
50 – under 60	32
60 – under 70	37
70 – under 80	21
80 – under 100	41
100 – under 150	15
Total	200

(a) Complete the histogram, which illustrates the amounts spent during each visit, on the grid below. The rectangle representing the 30 – under 50 class has been drawn.



[5]

(b) State the modal class.

....[1]

(c) If the 80 – under 100 and 100 – under 150 classes were merged to form a single class, 80 – under 150, calculate the value on the vertical axis of the height of the rectangle which would be drawn to represent the merged class.

....[2]

(ii) (a) The following table summarises the times spent by the customers in the supermarket.

Time (minutes)	*Number of customers*
0 – under 20	10
20 – under 40	49
40 – under 50	54
50 – under 60	45
60 – under 90	31
90 – under 120	11
Total	200

Estimate, to 3 significant figures, the mean and the standard deviation of the times spent in the supermarket by these customers.

Mean .... minutes

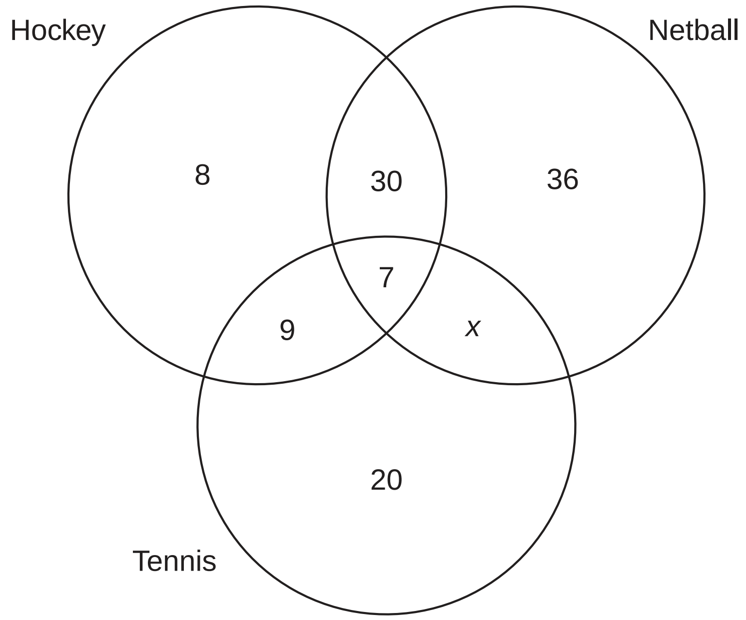
Standard deviation .... minutes [6]

(b) If the 60 – under 90 and 90 – under 120 classes are merged, the mean will increase. Explain why this is so.

....

....[2]

11 (a) At the start of the school year, all 125 girls at a school had to choose to play at least one of the sports hockey, netball and tennis. The following diagram illustrates the numbers who chose to play the different sports.



(i) Calculate the value of x.

....[2]

(ii) Explain what the value of x represents.

....

....[1]

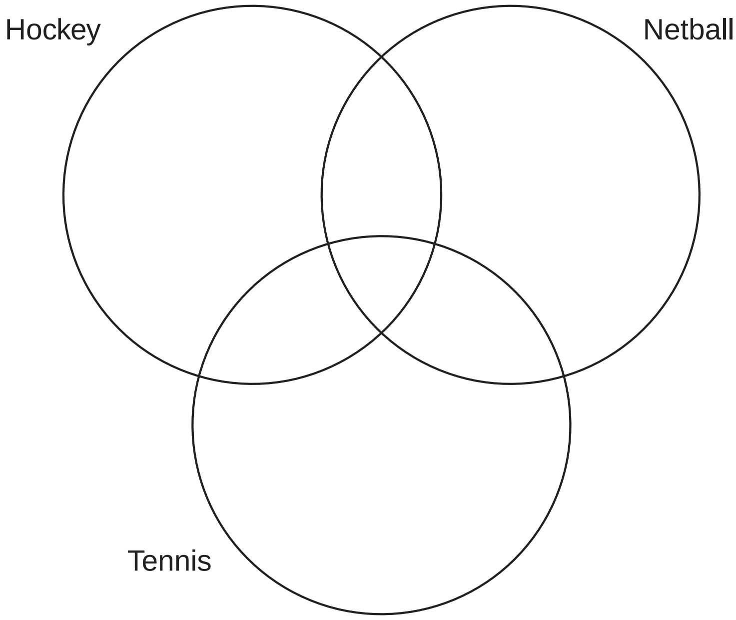
(iii) Showing all your working, determine which sport was chosen by the greatest number of girls.

....[2]

One month after the start of the year, girls were allowed to change their choice, as long as they still played at least one of the sports.

The following changes were made:
Two girls who had chosen to play only hockey decided to play **both** tennis **and** hockey.
Three girls who had chosen to play only tennis decided to play only netball **instead**.
Four girls who had chosen to play all three sports decided to stop playing tennis, but to continue playing both hockey and netball.

(iv) Insert, on the diagram below, the number of girls who have chosen to play the different sports **after these changes** have been made.



. [3]

(b) In this part of the question, you are NOT required to DRAW any pie charts.

The following table gives the number of different models of car sold by a large company in the year 2005.

Model	*Number sold*
Two-door saloon	491
Four-door saloon	646
Hatchback	542
Estate	521

(i) If these figures were to be illustrated by a pie chart, calculate, to the nearest degree, the angle of the sector representing sales of Estate cars.

.... ° [2]

(ii) In the year 2008, the company sold, in total, 26% more cars than it had sold in 2005. If the sales in 2005 were illustrated by a pie chart of radius 5 cm, calculate the radius of a pie chart illustrating the sales in 2008.

.... cm [3]

(iii) In a pie chart illustrating the sales in 2008, the angles of the sectors representing sales of two-door saloons, four-door saloons and hatchbacks are 72°, 103° and 85° respectively. Calculate the number of estate cars which were sold in 2008.

.... [3]

BLANK PAGE

BLANK PAGE