UNIVERSITY OF CAMBRIDGE INTERNATIONAL EXAMINATIONS
General Certificate of Education Ordinary Level

CANDIDATE NAME

CENTRE NUMBER

CANDIDATE NUMBER

STATISTICS **4040/13**

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 **21** printed pages and **3** blank pages.

Section A [36 marks]

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

For Examiner's Use

1 Four methods of sampling are:

simple random;

stratified random;

quota;

systematic.

State

(i) the method which does not require the use of any form of random numbers,

.....[1]

(ii) the method which does not require a sampling frame,

.....[1]

(iii) the method in which the choice of which individual items are selected is left to the interviewer/researcher,

.....[1]

(iv) the method which is most prone to bias if there is a pattern in the sampling frame which repeats at regular intervals,

.....[1]

(v) the two methods which require the use of random numbers to select every item in a sample,

.....

.....[1]

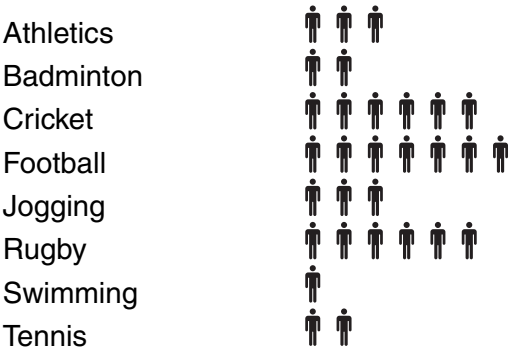
(vi) the two methods which require the population to be subdivided into appropriate categories.

.....

.....[1]

2 Twenty men were asked in which sports they participated regularly. The results are given in the diagram below.

For Examiner's Use



represents 1 man.

(i) State the name of this form of representation.

.....[1]

(ii) Give one disadvantage of this form of representing data.

.....

.....[1]

(iii) Name the most popular sport among these men.

.....[1]

(iv) Explain why the number of symbols in the diagram is greater than 20.

.....

.....

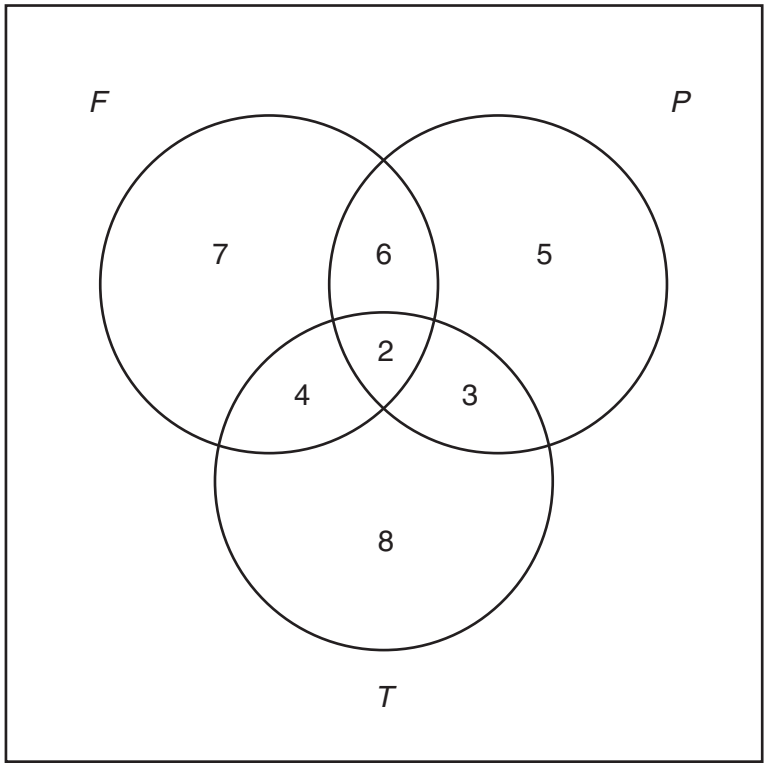
.....[1]

(v) Calculate the percentage of these men who played rugby regularly.

.....[2]

3 In a dancing competition, couples were judged on their performance in three dances. Each couple chose their dances from a list. Three of the dances on the list were the Foxtrot (*F*), the Paso Doble (*P*) and the Tango (*T*). The following diagram gives information on the choices of the couples entered in the competition.

For Examiner's Use



In total, there were 50 couples entered in the competition.

(i) Write, in the appropriate place in the diagram, the number of couples who did not choose any of these three dances.

[2]

(ii) State what the value 6 in the diagram represents.

.....

.....[1]

(iii) Find which of these three dances was chosen by the most couples, and state the number of couples who chose it.

Dance .....

Number of couples .....[3]

4 The date in the table below relates to the inhabitants of two streets, *S* and *T*.

For Examiner's Use

Street	*Median age (years)*	*Interquartile range of ages (years)*
S	27.5	18.0
T	58.3	4.2

(i) Possible descriptions of the ages of the inhabitants are:

C ages vary, generally old;
D ages vary, generally young;
E ages about the same, generally old;
F ages about the same, generally young.

Select which of these descriptions is most appropriate for the inhabitants of

(a) street *S*,

.....[1]

(b) street *T*.

.....[1]

25% of the population of street *S* is above 34.3 years old.

(ii) State the percentage of the population of street *S* which is between 27.5 and 34.3 years old.

.....[1]

(iii) Find the lower quartile age of the population of street *S*.

..... years [1]

(iv) For each of the following, state whether it is definitely true, possibly true, or definitely false.

(a) The lower and upper quartile ages of the population of street *T* are 56.2 and 60.4 years respectively.

.....[1]

(b) The oldest inhabitant of street *T* is older than the oldest inhabitant of street *S*.

.....[1]

5 A frequency distribution is given by the table below.

For Examiner's Use

Variable	*Frequency*
4	7
7	10
10	6
13	3

The arithmetic mean and the standard deviation of this distribution are 7.58 and 2.88 respectively, each correct to 3 significant figures.

(i) Using these results, and showing your method, find the mean and the **variance** of the frequency distribution in the following table.

Variable	*Frequency*
8	7
14	10
20	6
26	3

Mean = ...

Variance = ...[3]

(ii) The table below gives the grouped frequency distribution of a discrete variable.

Variable	*Frequency*
4, 5, 6	7
7, 8, 9	10
10, 11, 12	6
13, 14, 15	3

(a) State the class mid-points of this distribution.

..... , ... , ... , ...[1]

(b) Estimate, to 2 significant figures, the mean and the standard deviation of this distribution.

For Examiner's Use

Mean = ...

Standard deviation = ...[2]

6 The following table summarises the age and marital status of the female members of a sports club.

For Examiner's Use

Age	*Marital Status*				TOTAL
	Single	*Married*	*Widowed*	*Divorced*	
Under 20	18	2	0	0	20
20–24	9	13	4	7	33
25–29	7	89	7	11	114
30–39	5	4	4	2	15
40 and over	1	2	1	6	10
TOTAL	40	110	16	26	192

(i) A woman is chosen at random from the club.

(a) State the probability that she is single.

.....[1]

(b) Given that she is in the 30–39 age group, find the probability that she is single.

.....[2]

(ii) By considering divorced women aged 25–29, show that age and marital status are not independent for female members of this club.

[3]

BLANK PAGE

[Section B starts on page 10]

Section B [64 marks]

For Examiner's Use

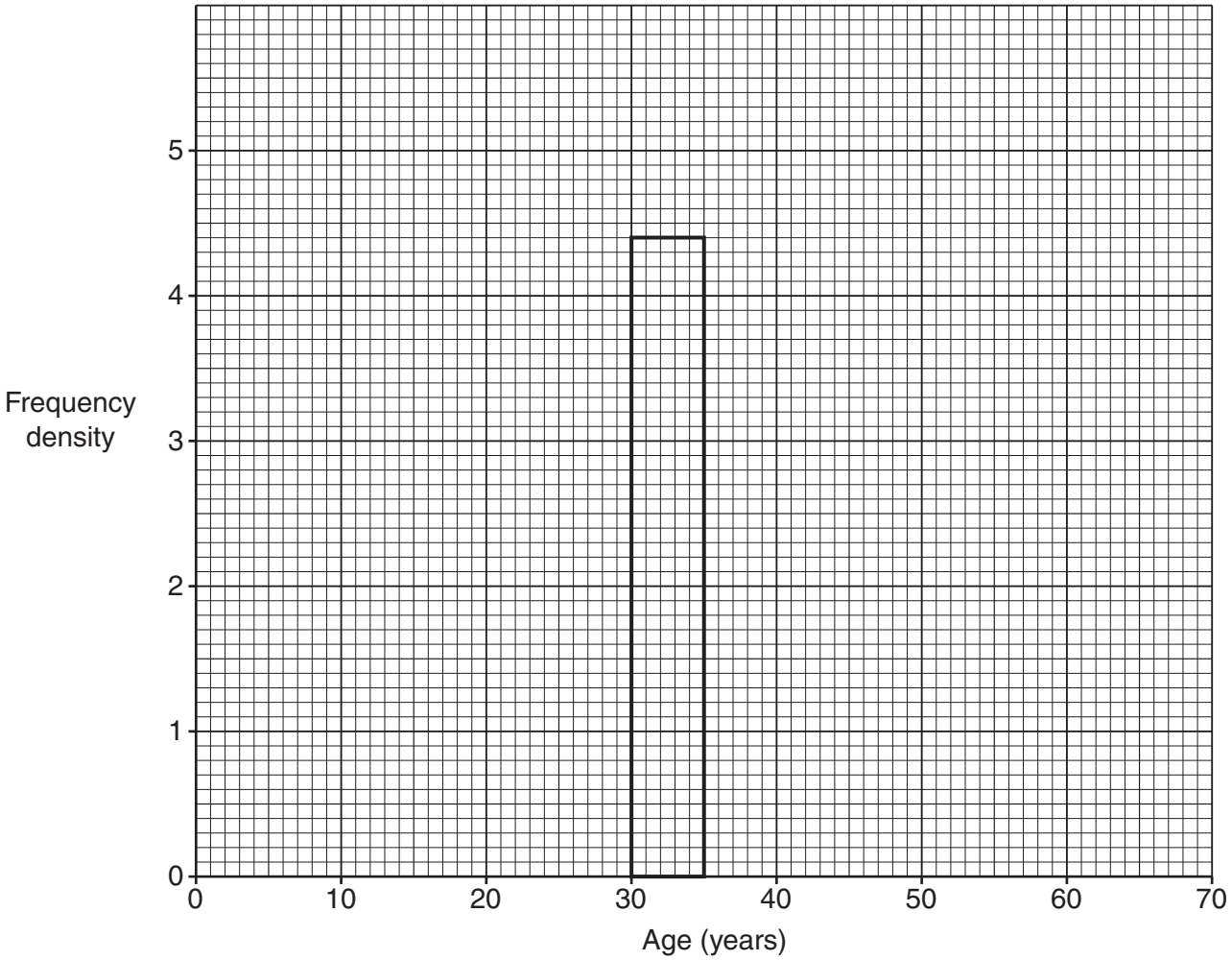
Answer not more than **four** of the questions 7 to 11.

Each question in this section carries 16 marks.

7 The table below gives the percentages in different age groups of anglers (people whose hobby is fishing) in the United Kingdom.

Age (years)	*Percentage of anglers*
15–under 25	28
25–under 30	20
30–under 35	22
35–under 45	16
45–under 65	14

On the grid below a histogram representing these ages is to be drawn. The rectangle for the 30–under 35 group is given.



(i) By considering the height of the given rectangle, state, on the answer line below, the units of the vertical axis.

.....[2]

(ii) Draw the rectangles representing the other four age groups.

[4]

(iii) Briefly explain why it is not possible to obtain an accurate value for the mode from this histogram.

.....

.....[1]

(iv) Write down the mid-points of the five age groups.

.....[1]

(v) Estimate, to 3 significant figures, the mean and the standard deviation of the anglers' ages.

Mean = ... years

Standard deviation = ... years [6]

(vi) State, giving a reason, which of the mean and median ages you would expect to be higher for this distribution.

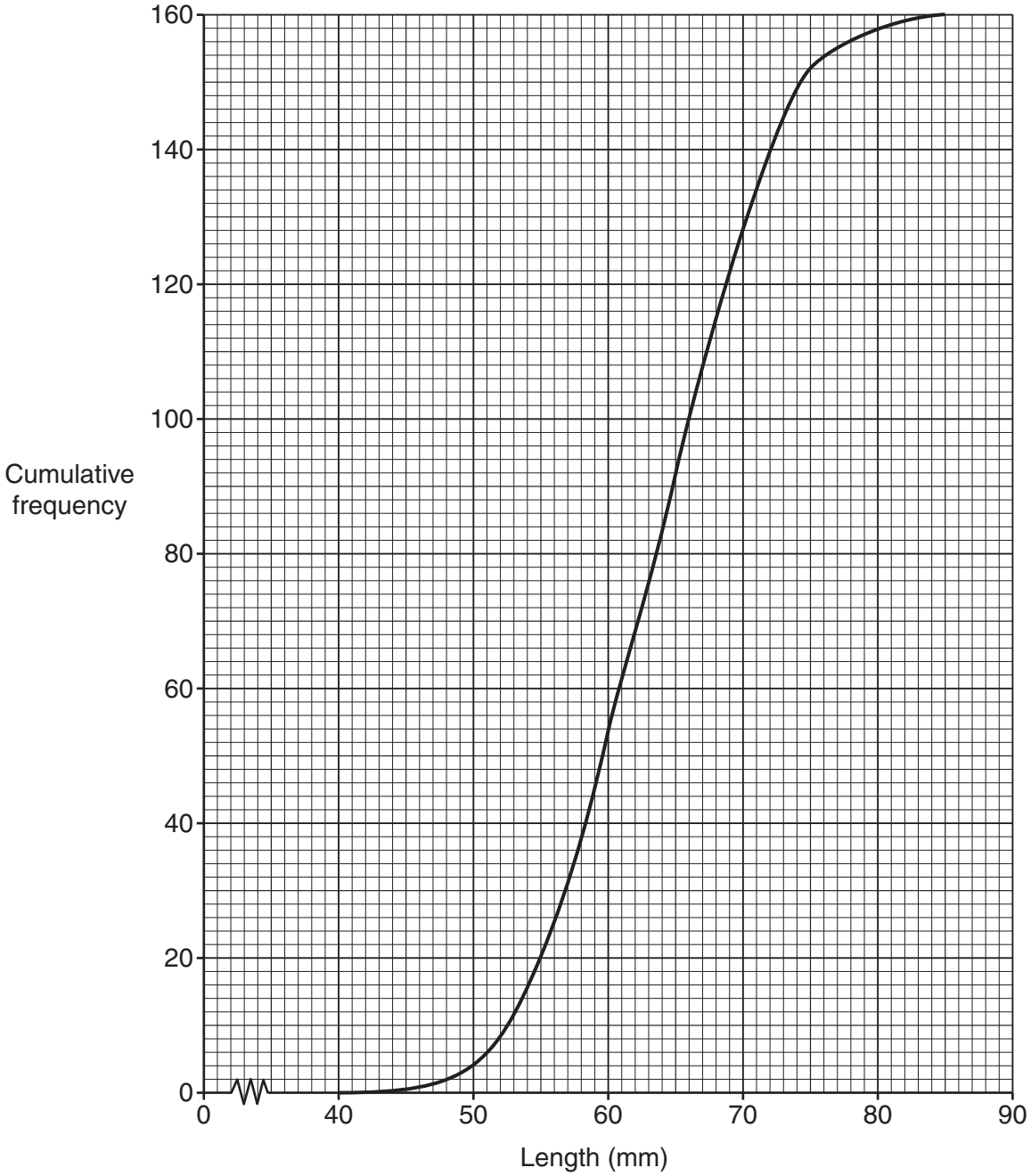
.....

.....

.....[2]

8 The lengths, in mm, of a sample of 160 rods produced on a machine are represented by the following cumulative frequency curve.

For Examiner's Use



(i) State which feature of the graph shows that the majority of rods had lengths in the middle of this range.

.....

.....[1]

For Examiner's Use

(ii) **Use the graph** to estimate

(a) the lower quartile length,

..... mm [2]

(b) the 80th percentile length.

..... mm [3]

The rods cost $5 each to produce.

Rods can only be used for their intended purpose if their lengths are in the range 55 mm to 75 mm inclusive.

Rods which are longer than 75 mm are shortened to be within the accepted limits, at a cost of $3 each.

Rods which are shorter than 55 mm cannot be used, and are sold for recycling at $2 each.

(iii) Estimate the number of rods in this sample which

(a) will have to be sold for recycling,

.....[1]

(b) can be used after having been shortened,

.....[2]

(c) can be used immediately after production.

.....[2]

[Turn over

(iv) By considering the cost of manufacturing all 160 rods in the sample, estimate the mean cost per rod of those rods which can be used.

For Examiner's Use

Mean = $ .....[5]

BLANK PAGE

[Question 9 is printed on the next page]

9 **(a)** The information below relates to a city in the year 2008.

The total population at the start of the year was 420 000.

The crude death rate for the city was 9.6 per 1000.

There were 5250 births in the city.

(i) Calculate the number of deaths in the city in the year 2008.

.....[2]

(ii) State, giving a reason, whether subtracting your answer to **(i)** from 5250 would give the increase in the city's population in 2008.

.....

.....

.....[2]

(b) The data in the table below relates to a town in the year 2009.

Age	*Population*	*Number of deaths*	*Death rate per 1000*	*Standard population*
0–24	10 000	70	7	3500
25–49		80	10	3000
50 and over				1500
TOTAL	25 000	262		8000

(i) Fill in the four values which are missing from the table.

[5]

(ii) Calculate, to 1 decimal place, and stating the units, the crude death rate and the standardised death rate of the town.

For Examiner's Use

Crude death rate = ...

Standardised death rate = ...[5]

It was subsequently discovered that the overall population of the region in which the town was situated was younger than had been thought, and that a more appropriate standard population for the three age groups would be 4000, 3000 and 1000 respectively.

(iii) Without carrying out any further calculations, state what effect using this new standard population would have on the crude death rate and on the standardised death rate of the town.

.....

.....

.....

.....

.....[2]

10 Zaheer only buys petrol when the tank of his car is almost empty, except that if he passes a garage selling cheap petrol he will buy some, however much is still in the tank.

For Examiner's Use

Whenever he buys petrol he **always** fills the tank and records the details in his log book, and **usually** re-sets the trip recorder to 0.

The following table is an extract from Zaheer's log book. The odometer records the total distance in km travelled by the car since its manufacture. The trip recorder records the distance in km travelled by the car since the trip recorder was last re-set to 0.

Date	*Odometer reading (km)*	*Trip recorder reading (km)*	*Petrol bought (litres)*
July 5	16319	510	38.5
July 11	16824	505	39.5
July 15	17099	275	20.5
July 18	17584	485	37
July 22	17789	690	15
July 24	18084	295	23.5
July 25	18579	495	38.5
July 26	18799	220	18

(i) Estimate the capacity of the car's petrol tank to the nearest 5 litres.

..... litres [1]

(ii) State the four dates on which Zaheer bought cheap petrol.

.....[1]

(iii) On one occasion when Zaheer bought petrol he forgot to re-set the trip recorder.

(a) State the date on which this happened.

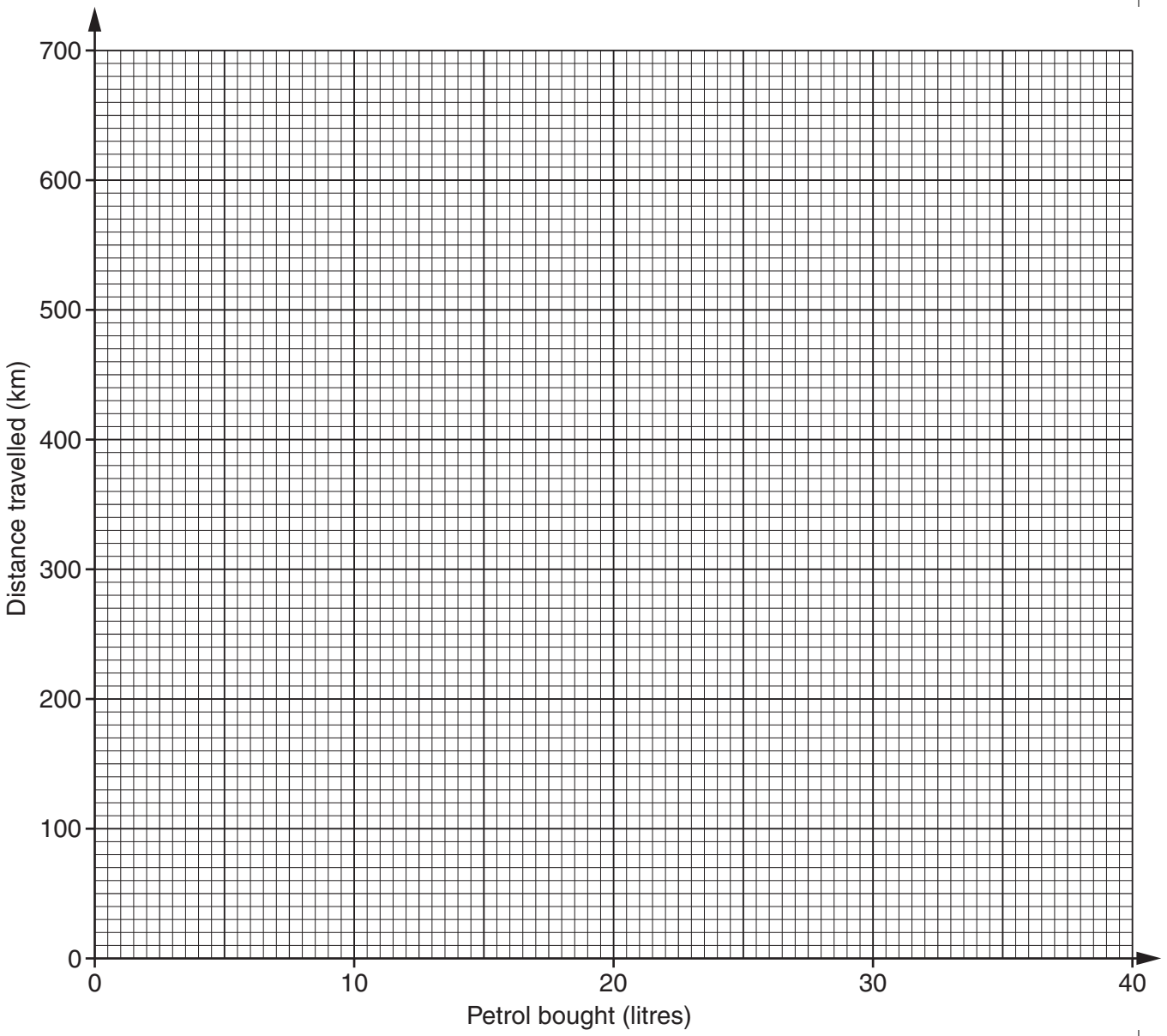
.....[1]

(b) Find the actual distance travelled between this purchase and the next one.

..... km [1]

The grid below gives petrol bought, in litres, on the *x*-axis, and distance travelled since previous purchase of petrol, in km, on the *y*-axis.

For Examiner's Use



(iv) Plot on the grid the eight points corresponding to the occasions on which Zaheer purchased petrol. [2]

(v) Calculate the overall mean and the two semi-averages, and plot them on the grid.

For Examiner's Use

[6]

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

(vii) Use your line of best fit to estimate

(a) the distance, to the nearest 10 km, which Zaheer could travel using 30 litres of petrol,

..... km [1]

(b) the petrol consumption of Zaheer's car, stating the units.

.....[2]

BLANK PAGE

[Question 11 is printed on the next page]

11 (a) Research organisations in the United Kingdom often divide the population into six categories, called social classes, denoted by *A*, *B*, *C1*, *C2*, *D* and *E*. The percentages of the population in the different social classes in the year 1987 are given in the following table.

Class	*A* and *B*	*C1*	*C2*	*D* and *E*
Percentage	13	21	38	28

(i) Draw and label a pie-chart of radius 4 cm to represent these percentages.

[4]

(ii) By the year 2001, the population of the U.K. was 7% larger than it had been in 1987. Calculate, to 2 decimal places, the radius of a corresponding pie chart representing the 2001 population.

(You are **not** required to draw this chart.)

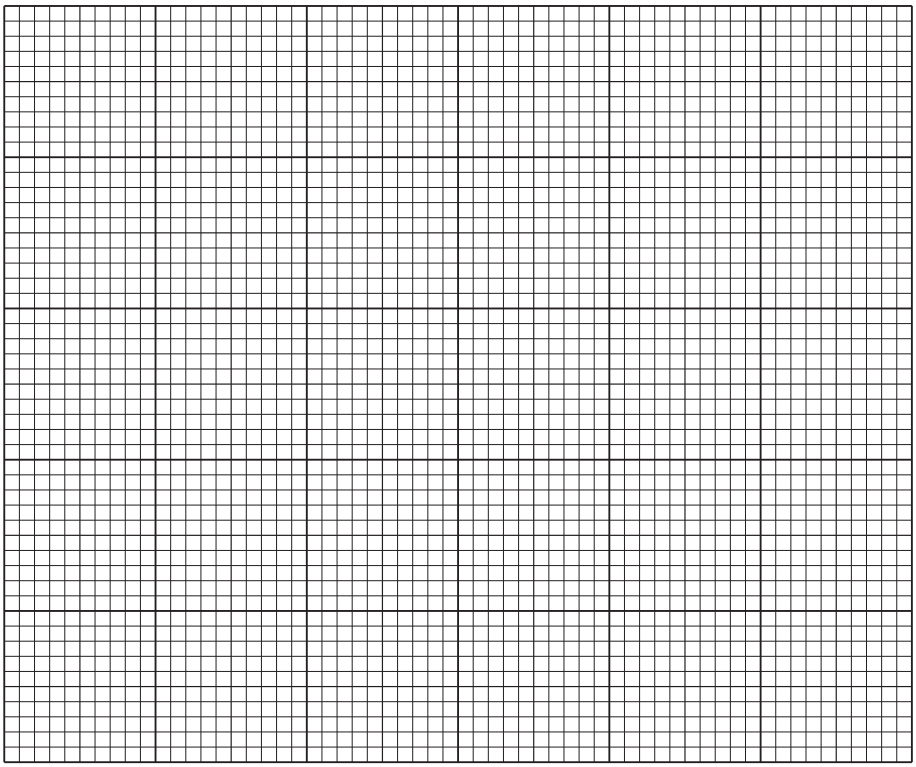
..... cm [4]

(b) The following table gives, for each gender separately, the percentage of the U.K. population of working age in different occupational groups in the year 2002.

	Males (%)	*Females (%)*
Professional	8	3
Managerial and Technical	27	24
Skilled (non-manual)	11	31
Skilled (manual)	27	7
Partly skilled and Unskilled	17	21
Others	10	14

'Others' included those in the Armed Forces, those who did not state their current or last occupation, and those who had not worked in the previous eight years.

(i) Draw, on the grid below, fully-labelled percentage component bar charts for males and females to illustrate the data in the table.



[4]

[Question 11 continues on the next page]

For Examiner's Use

(ii) Describe two differences between the occupations of males and females illustrated by the charts.

.....

.....

.....

.....

.....

.....[2]

(iii) State, with a reason, whether you would regard pie charts or bar charts as more appropriate to **compare** these two sets of percentages.

.....

.....

.....

.....[2]