UNIVERSITY OF CAMBRIDGE INTERNATIONAL EXAMINATIONS
General Certificate of Education Ordinary Level

CANDIDATE NAME

CENTRE NUMBER

CANDIDATE NUMBER

3225555521

STATISTICS **4040/12**

Paper 1 **October/November 2012**

2 hours 15 minutes

Candidates answer on the question paper.

Additional Materials: 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.

For Examiner's Use

1 A supermarket carried out a survey to see if customers were satisfied with the services it provided. The results are shown in the pictogram below.

☺ = 60 satisfied customers

☹ = 6 dissatisfied customers

(i) State how many customers were

(a) satisfied,

.....[1]

(b) dissatisfied.

.....[1]

(ii) Explain why this is not a good pictogram.

.....

.....

.....[2]

The survey was carried out by a team of interviewers who spoke to customers in the supermarket on one particular weekday afternoon.

(iii) State, giving a reason for your answer, whether or not you would expect the results of the survey to be biased.

.....

.....

.....[2]

2 A car hire company has 5 small cars, 15 medium-sized cars and 10 large cars. An inspector selects a sample of cars from this population to test their mechanical condition.

For Examiner's Use

(i) For the different possible sampling methods, state whether each of the following is true or false.

(a) A (simple) random sample of size 10 might contain only medium-sized cars.

.....[1]

(b) A systematic sample will require the use of a sampling frame.

.....[1]

(c) If quota sampling is used, the inspector may select as many cars as he wishes, and as many of each type of car as he wishes.

.....[1]

(d) If systematic sampling is used, every car has an equal chance of being selected after the first car has been selected.

.....[1]

(ii) The inspector decides he has time to test only 9 cars. If he wishes to obtain a random sample stratified by car size, find how many cars of each size he must select.

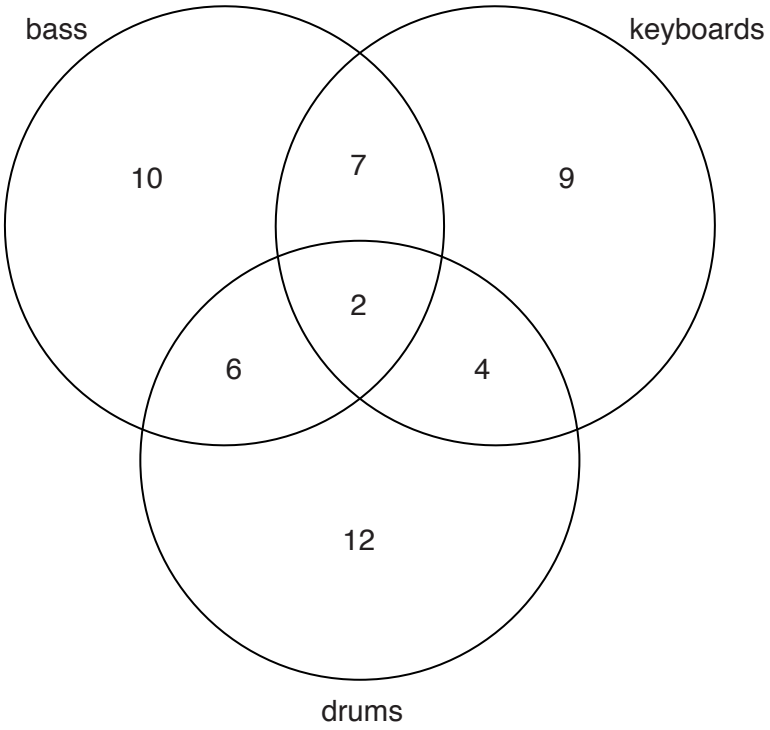
Small .....

Medium .....

Large .....[2]

3 The diagram below shows the number of musicians performing at a music festival who play one or more of the instruments bass, keyboards and drums.

For Examiner's Use



Use this information to find the number of musicians who play

(i) keyboards,

.....[1]

(ii) bass and drums,

.....[1]

(iii) keyboards or drums or both.

.....[2]

The guitar is also played by all the bass players, but not by any of the other musicians.

For Examiner's Use

For the four instruments, bass, keyboards, drums and guitar, find the number of musicians who play

(iv) exactly three instruments,

.....[1]

(v) exactly two instruments.

.....[1]

4 The products leaving a factory are checked every day to ensure that faulty items are not sent out to customers. The table below shows values of the variable X, the number of faulty items found in a day, for a period of 50 days.

Number of faulty items, x	0	1	2	3	4	5 or more
Number of days, f	19	10	7	6	5	3

For example, there were 7 days when 2 faulty items were found.

(i) State the mode of the distribution, and explain why it is a poor measure of central tendency in this case.

.....

.....

.....[2]

(ii) Find the median of the distribution.

.....[1]

The mean of the distribution has been calculated to be 1.75 faulty items per day.

(iii) Find the upper class limit of the 5 or more class which has been used in the calculation.

.....[3]

5 Each person at a committee meeting receives a plate of seven biscuits for refreshments, of which four are ginger, two are chocolate, and one is plain. One committee member chooses four biscuits at random from her plate.

For Examiner's Use

If X is the number of chocolate biscuits chosen,

(i) state the possible values of X,

.....[1]

(ii) find the probability of each value of X, and present your results in a suitable table.

[5]

6 At a medical centre the dosage for a particular vaccination is 0.500 ml, but small variations from this are acceptable.

The exact dosages (in millilitres) in a prepared batch of six were found to be as follows.

0.512 0.506 0.493 0.518 0.491 0.514

(i) Calculate, to 5 decimal places, the standard deviation of this batch.

.....[2]

It is very important that, for a batch of six, the standard deviation is no greater than 0.01250 ml, and the range is no more than three times the standard deviation.

(ii) Find whether or not the prepared batch satisfies both of these conditions.

.....[3]

A trainee technician at the medical centre suggests that, as well as checking that the standard deviation is not too large, they must also check that it is not too small.

(iii) State, giving a reason, whether or not you agree with the trainee.

.....

.....[1]

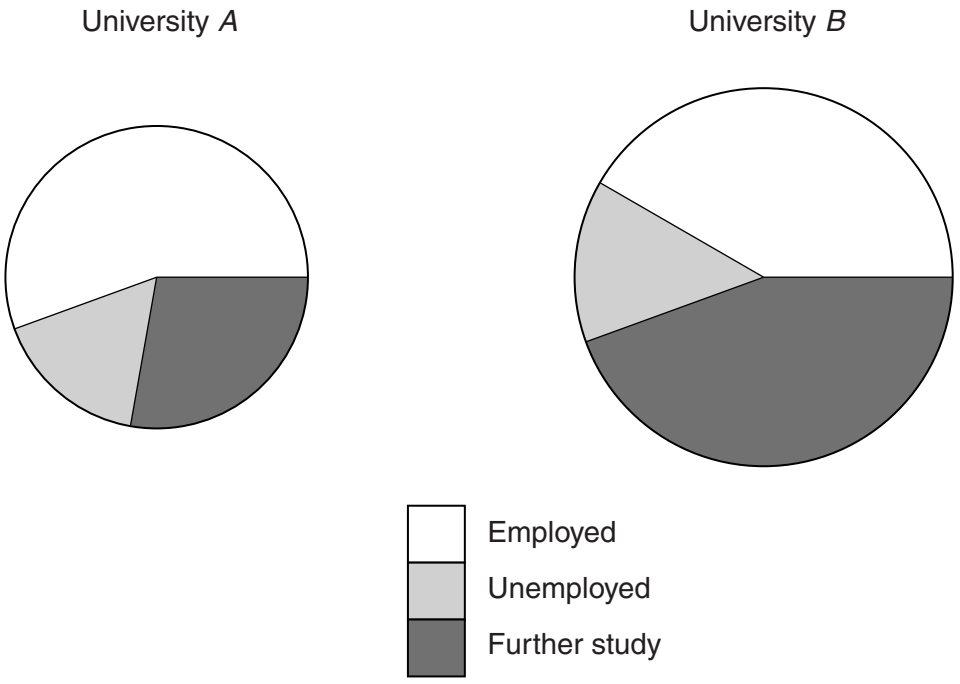
Section B [64 marks]

For Examiner's Use

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

Each question in this section carries 16 marks.

7 Two universities carried out a survey of their graduates one year after graduation. The graduates were classified as employed, unemployed, or engaged in further study, as shown in the following pie charts, which are drawn to scale.



The chart for University *A* represents a total of 1170 graduates.

Find the number of graduates who, one year after graduation, were

(i) from University *A* and engaged in further study,

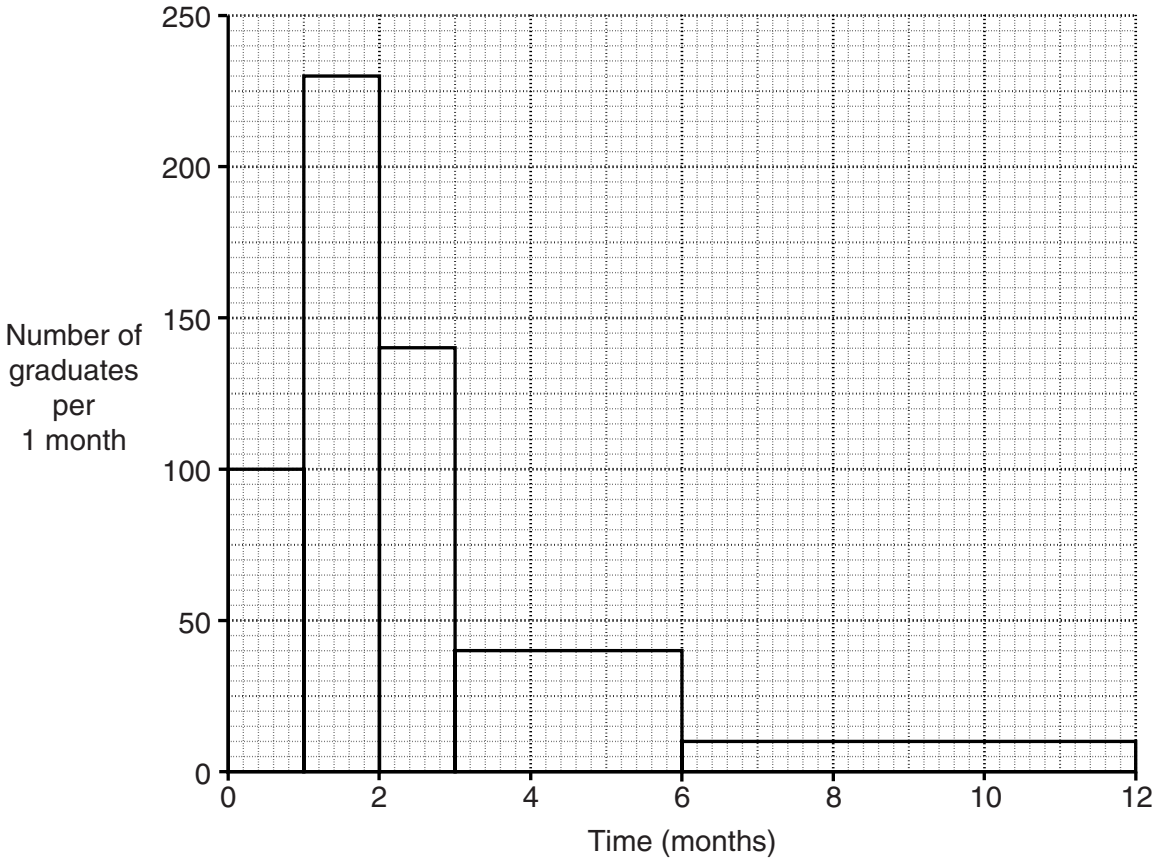
.....[2]

(ii) from University *B* and unemployed.

.....[4]

The graduates from University *A* who were employed were asked how long it took them to find employment. The times are represented in the histogram below.

For Examiner's Use



(iii) State the modal class.

.....[1]

Use the histogram to find the number of graduates from University *A* for whom the time taken to find employment was

(iv) from 0 months to 3 months,

.....[2]

(v) from 3 months to 6 months,

.....[2]

(vi) from 6 months to 12 months.

.....[1]

Of the employed graduates from University *A,* 163 were science graduates.

For Examiner's Use

(vii) Estimate the number of science graduates from University *A* who found employment within 6 months of graduation.

.....[3]

(viii) State the assumption which you have made in your calculation in part **(vii)**.

.....

.....[1]

8 In this question calculate all fertility rates per thousand, and to the nearest whole number.

For Examiner's Use

The fertility rate is defined as the number of births per 1000 females.

The table below gives information about the female population and births in the town of Bluedorf for the year 2010, together with the standard female population of the area in which Bluedorf is situated.

Age group of females	*Births*	*Population of females in age group*	*Age group fertility rate*	*Standard population of females* (%)
Under 20	112	3200		25
20 – 30	459	2250		15
31 – 40	488	3050		20
Over 40	76	4000		40

(i) Calculate the crude fertility rate for Bluedorf.

.....[4]

(ii) Calculate the fertility rate for each age group and insert the values in the table above.

[2]

(iii) Calculate the standardised fertility rate for Bluedorf.

.....[4]

The table below gives information about Redville, another town in the same area as Bluedorf, also for the year 2010.

For Examiner's Use

Age group of females	*Fertility rate* (per 1000 females)	*Population of females in age group*
Under 20	32	3000
20 – 30	225	1560
31 – 40	180	1700
Over 40	20	4950

(iv) Calculate the standardised fertility rate for Redville in the year 2010, using the same standard population as for Bluedorf.

.....[2]

(v) Find how many more births there were in Bluedorf than in Redville in the year 2010.

.....[2]

The local government of the area in which Bluedorf and Redville are situated wishes to limit population growth, but only has sufficient funds for a publicity campaign on birth control in one of these two towns.

(vi) State, with a reason, in which of these two towns the campaign should be conducted.

.....

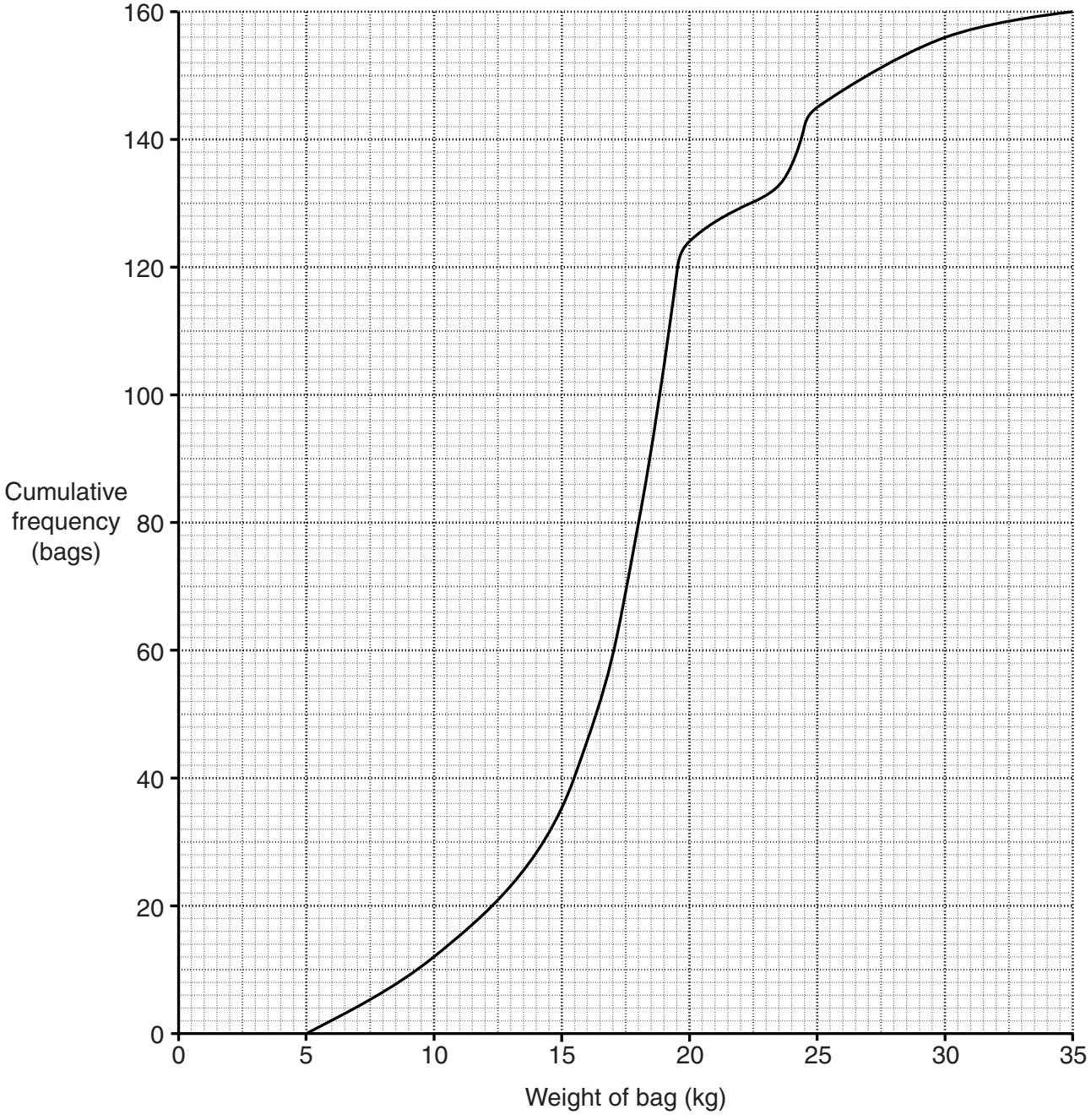
.....

.....[2]

9 In answering all parts of this question you are required to show your working, either as the calculations leading to your result, or by drawing appropriate lines on the graph.

For Examiner's Use

Passengers on an aeroplane are allowed to check-in one bag for carriage in the aeroplane's hold. The weights of checked-in bags for one particular flight are illustrated in the cumulative frequency curve below.



(i) Use the graph to estimate

(a) the median of the weights,

..... kg [1]

(b) the interquartile range of the weights,

For Examiner's Use

..... kg [4]

(c) the value of k, if the kth percentile of the weights is 24 kg.

k = ...[2]

The weight allowance for a checked-in bag is 20 kg. Bags heavier than 20 kg are classified as overweight.

(ii) Use the graph to estimate

(a) the number of overweight bags,

.....[2]

(b) the median weight of the overweight bags.

..... kg [2]

For overweight bags passengers are charged \$8 per kilogram for the amount by which the weight of their bag exceeds 20 kg. For example, for a bag weighing 21.5 kg, the passenger would be charged \$12.

(iii) Assuming that the median and mean of the overweight bags are approximately the same, use your answers to part **(ii)** to estimate the total money received in charges for the overweight bags on this particular flight.

\$ .....[4]

(iv) Interpret the point at which the cumulative frequency curve meets the horizontal axis.

.....

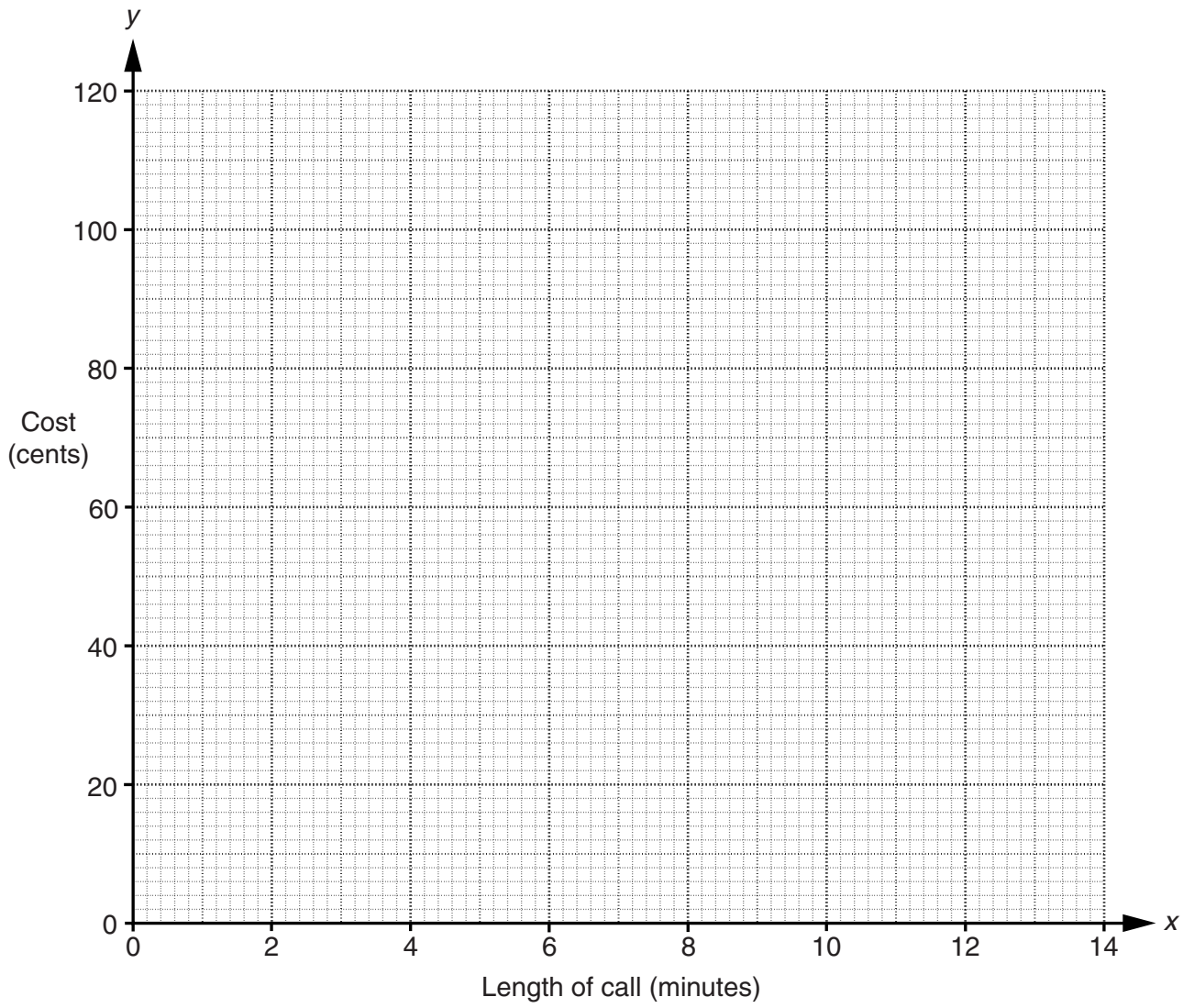
.....[1]

10 Soraya is checking her telephone bill. For all the national calls listed she writes down the length of the call, to the nearest minute, and the cost of the call, to the nearest cent. The results are shown in the following table.

For Examiner's Use

Length of call (minutes), x	1	8	14	3	7	1	2	11
Cost of call (cents), y	25	71	108	39	63	28	34	93

(i) Plot these data on the grid below.



[2]

The data have an overall mean of (5.875, 57.625) and a lower semi-average of (1.75, 31.5).

(ii) Show how the value 31.5 is calculated.

[2]

(iii) Find the upper semi-average, and plot this and the two given averages on your graph.

For Examiner's Use

.....[3]

(iv) Use your plotted averages to draw a line of best fit, and find its equation in the form $y = mx + c$.

.....[4]

For national calls, the cost is made up of a connection charge for the call, and a charge per minute for the length of the call.

(v) Write down the amounts of these charges.

Connection charge = ... cents

Charge per minute = ... cents
[1]

For local calls, there is a fixed charge of 10 cents for any call lasting up to 2 minutes. There is a further charge of 3 cents per minute for any length of time for which the call is longer than 2 minutes.

(vi) Draw, on the grid in part **(i)**, a graph showing the cost of local calls lasting between 0 and 14 minutes.

[2]

Later, Soraya makes a national call which costs 40 cents, and a local call which costs 40 cents.

(vii) Use your graphs to estimate how much longer the local call lasted than the national call. Give your answer to the nearest minute.

.....minutes [2]

11 In this question give your answers either as fractions, or as exact decimals, or as decimals correct to 3 significant figures.

For Examiner's Use

(a) In an office there are 15 workers, of whom 3 are supervisors and 12 are assistants. Two workers are chosen at random, without replacement.

Find the probability of choosing

(i) two assistants,

.....[2]

(ii) one supervisor and one assistant.

.....[2]

(b) Kwame likes quizzes, and estimates that he knows the correct answer to 60% of quiz questions. He enters a television quiz, in which, for each question asked, four alternative answers are offered to the contestant, only one of which is correct. If Kwame knows the correct answer he chooses the correct answer, otherwise he makes a random guess. Assume that his own estimate of his knowledge is accurate.

(i) Find the probability that he gets a question he is asked correct.

.....[2]

(ii) If he is asked three questions, find the probability that

(a) he answers them all correctly,

.....[2]

(b) he knows none of the correct answers, but answers all the questions correctly.

.....[3]

(c) In a game, a turn consists of rolling an unbiased six-sided dice with faces numbered 1, 2, 3, 4, 5 and 6.
If the number obtained is 3, 4, 5 or 6, then that is the score for the turn.
If the number obtained is 2, then the dice is rolled one more time, and the score for the turn is the sum of 2 and the number obtained on the second roll.
If the number obtained is 1, the dice is rolled two more times, and the score for the turn is the sum of 1 and the numbers obtained on the second and third rolls.

For Examiner's Use

If, in a turn, a player obtains a score of 5, find the probability that the dice has been rolled two times altogether.

.....[5]

BLANK PAGE