Cambridge
O Level

**Cambridge International Examinations**
Cambridge Ordinary Level

CANDIDATE NAME

CENTRE NUMBER

CANDIDATE NUMBER



**STATISTICS** **4040/12**

Paper 1 **October/November 2014**

**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 an HB pencil for any diagrams or graphs.
Do not use staples, paper clips, 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.




CAMBRIDGE International Examinations

**Section A** [36 marks]

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

**1** In an industrial process, readings, $x$, of a particular gauge are recorded regularly. For 6 such recorded readings it is found that $\Sigma x = 279$ and $\Sigma x^2 = 13\,093$.

**(i)** Find the mean and standard deviation of $x$.

Mean = .....

Standard deviation = .....[4]

It was discovered later that one of the readings had been incorrectly recorded as 43, when in fact the correct reading was 34.

**(ii)** State, for each of the mean and standard deviation, whether its correct value will be smaller than, larger than, or the same as the value found in part **(i)**.

Mean .....

Standard deviation .....[2]

**2** A student calculated, correctly, five statistical measures for a set of data. The fi[illegible] obtained were, in ascending order, 6, 36, 43, 48 and 53.

**(i)** Insert these values in their correct positions in the table below.

| *Statistical measure* | *Value* |
|---|---|
| Median | |
| Lower quartile | |
| Upper quartile | |
| Standard deviation | |
| Variance | |

[5]

**(ii)** State the value of the 75th percentile for the student's original set of data.

.....[1]

**3** A national government plans a survey to obtain the responses of its citizens to its prop[illegible] wind farms as sources of renewable energy.

**(i)** The following survey methods are considered.

*A* Questionnaires will be mailed to 1000 citizens.

*B* Face to face interviews will be conducted with a total of 1000 citizens in shopping centres in different parts of the country.

*C* A questionnaire will be placed on the Internet inviting responses from anyone.

*D* Telephone calls will be made to 1000 citizens chosen from the telephone directory.

**(a)** Explain why method *D* would produce a biased sample.

.....

.....[1]

**(b)** Give one advantage of method *B* over method *A*.

.....

.....[1]

**(c)** Give one advantage and one disadvantage of method *C*.

Advantage .....

.....

Disadvantage .....

.....[2]

**(ii)** A closed question which will be asked in the survey is as follows.

Are you in favour of wind farms being built in your area?

| | |
|---|---|
| Yes | |
| No | |

**(a)** Give one disadvantage of asking closed questions.

.....

.....[1]

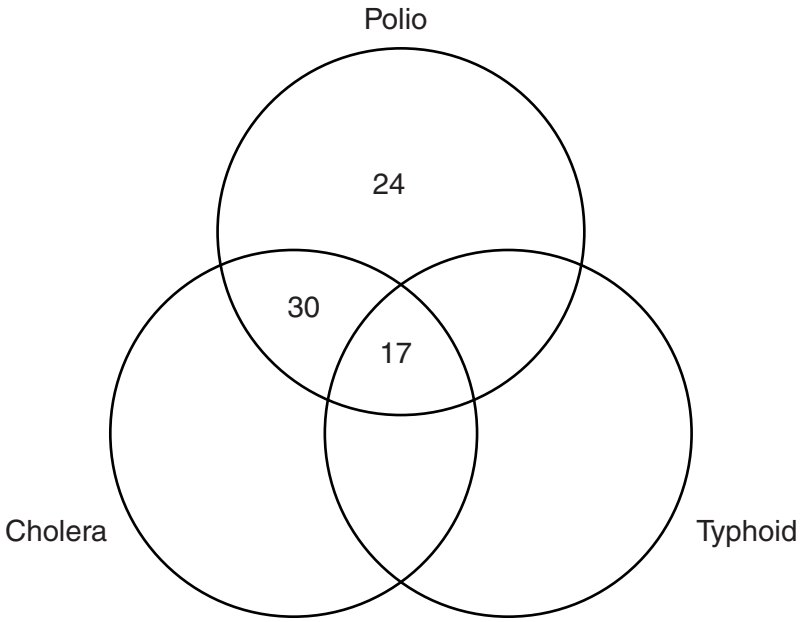
**(b)** Write down one open question which could be asked in the survey.

.....

.....

.....[1]

**4** The following diagram is to show the number of patients at a medical centre who ha[illegible] vaccine against one or more of the diseases polio, cholera and typhoid.




**(i)** **Complete the diagram** using the following information.

**(a)** The number of patients who have received only cholera vaccine is 5 fewer than the number of patients who have received only polio vaccine. [1]

**(b)** The number of patients who have received only typhoid vaccine is two thirds of the number of patients who have received polio and cholera vaccines but not typhoid vaccine. [1]

**(c)** The number of patients who have received polio and typhoid vaccines but not cholera vaccine is the same as the number of patients who have received all three vaccines. [1]

**(d)** Twice as many patients have received typhoid and cholera vaccines but not polio vaccine as have received all three vaccines. [1]

**(ii)** Find the mode of the number of these vaccines received by these patients.

.....[2]

**5** The table below gives information on the gender of, and number of books written by, authors attending a book fair.

| | *Number of books written* | | | | TOTAL |
|---|---|---|---|---|---|
| | 1 – 5 | 6 – 10 | 11 – 15 | More than 15 | |
| *Male* | 5 | 6 | 3 | 1 | 15 |
| *Female* | 12 | 8 | 4 | 1 | 25 |
| TOTAL | 17 | 14 | 7 | 2 | 40 |

One of these authors is chosen at random to speak at the opening ceremony.

Find the probability of choosing

**(i)** a male,

.....[1]

**(ii)** a female who has written 11 or more books,

.....[1]

**(iii)** an author who has written 6 – 10 books, given that the author is male.

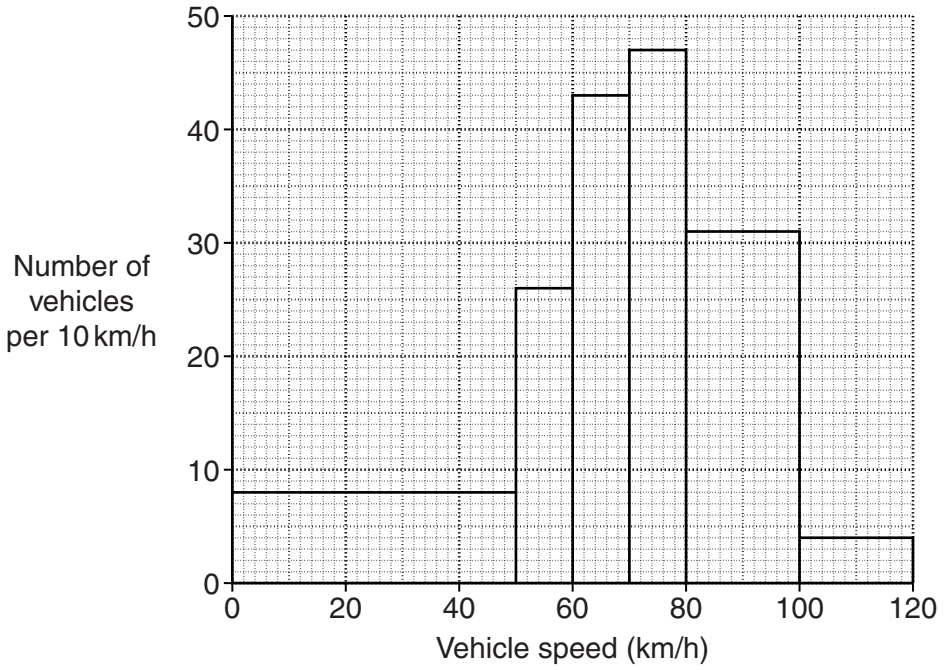
.....[1]

Two authors are chosen at random to lead discussion groups.

**(iv)** Find the probability that both have written 5 or fewer books.

.....[3]

**6** A police camera at the side of a road measures the speed, in km/h, of every vehicle [illegible] the road.
The following histogram represents the information it recorded over a certain period of time.




Use the histogram to find, for this period of time, the number of vehicles whose speeds were

**(i)** from 50 km/h to under 80 km/h,

.....[2]

**(ii)** from 80 km/h to under 100 km/h,

.....[2]

**(iii)** under 50 km/h.

.....[1]

The speed limit on this road is 100 km/h. Any driver of a vehicle travelling at a speed which is 5 km/h or more greater than the speed limit must pay a fine.

**(iv)** Estimate the number of drivers represented by this information who had to pay a fine.

.....[1]

**Section B** [64 marks]

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

Each question in this section carries 16 marks.

**7 In this question calculate all accident rates per thousand. Where values do not work out exactly give your answers to one decimal place.**

The table below gives information on the number of employees, and the number of accidents they suffered, at a building construction company, Kwikbuild, in the year 2012. It also shows the standard population for the building construction industry.

| *Job group* | *Number of accidents* | *Number of employees* | *Job group accident rate* | *Standard population* (%) |
|---|---|---|---|---|
| Management | 1 | 25 | | 8 |
| Office Administration | 8 | 167 | | 35 |
| Site Supervision | 3 | 40 | | 12 |
| Site Labour | 37 | 228 | | 45 |

**(i)** Calculate the crude accident rate for Kwikbuild.

.....[4]

**(ii)** Calculate the accident rate for each job group and insert the values in the table above.

[2]

**(iii)** Use your results from part **(ii)** to calculate the standardised accident rate for Kwi

.....[4]

Fastbuild is another building construction company. In 2012 its crude and standardised accident rates were 109.4 and 98.7 per thousand respectively.

**(iv)** State, with a reason, which of the two companies most likely operates in the safer environment.

.....

.....

.....[2]

For each company some of the accidents suffered by employees were classed as 'serious', and they all occurred in the Site Labour job group.
The table below gives information on the serious accidents suffered at the two companies.

| *Job group* | *Company* | *Number of serious accidents* | *Number of employees* |
|---|---|---|---|
| Site Labour | Kwikbuild | 7 | 228 |
| | Fastbuild | 5 | 154 |

**(v)** Calculate, for each company, for the Site Labour job group only, the serious accident rate, and hence state the company where an employee is less likely to suffer a serious accident.

.....[2]

**(vi)** State, with a reason, whether the values you have calculated in part **(v)** are crude or standardised rates.

.....

.....

.....[2]

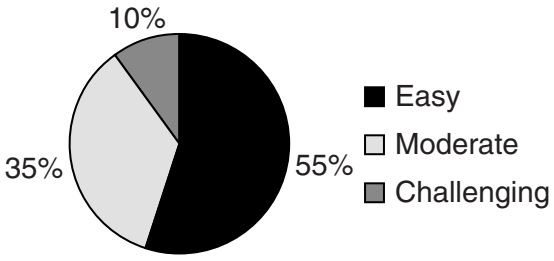
**8** A running club holds a cross-country race. Competitors enter in either the junior or [illegible] category. When they enter, they also choose to follow one of three routes: easy, moderate or challenging. Information about the number of competitors and the routes chosen is shown [illegible]

**Number of competitors entering the race by age category**

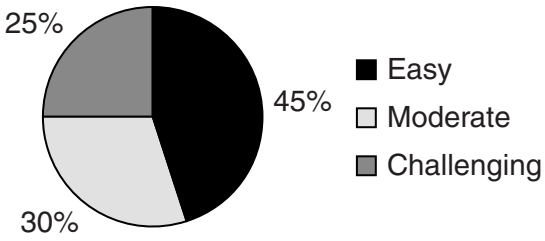
= 10 junior competitors

= 10 senior competitors

**Percentages of junior entrants by choice of route**




**Percentages of senior entrants by choice of route**




**(i)** Find the total number of competitors who entered the race.

.....[1]

**(ii)** Show that there were 42 junior competitors who chose the moderate route.

[1]

**(iii)** Find the number of senior competitors who chose the easy route.

.....[2]

Not all the entrants completed the race. The times taken by those who did complete [illegible] shown in the table below.

| *Completion time* (minutes) | *Number of competitors* | | | | | |
|---|---|---|---|---|---|---|
| | *Junior* | | | *Senior* | | |
| | Easy | Moderate | Challenging | Easy | Moderate | Challenging |
| 60 – under 90 | 16 | 4 | 0 | 19 | 3 | 2 |
| 90 – under 120 | 29 | 12 | 1 | 32 | 15 | 5 |
| 120 – under 150 | 18 | 15 | 3 | 20 | 17 | 14 |
| 150 – under 180 | 2 | 8 | 6 | 0 | 11 | 16 |
| TOTAL | 65 | 39 | 10 | 71 | 46 | 37 |

**(iv)** Find the number of competitors who entered the race but did not complete it.

.....[3]

**(v)** Estimate, to the nearest minute, the mean time taken by senior competitors who completed the challenging route.

.....[3]

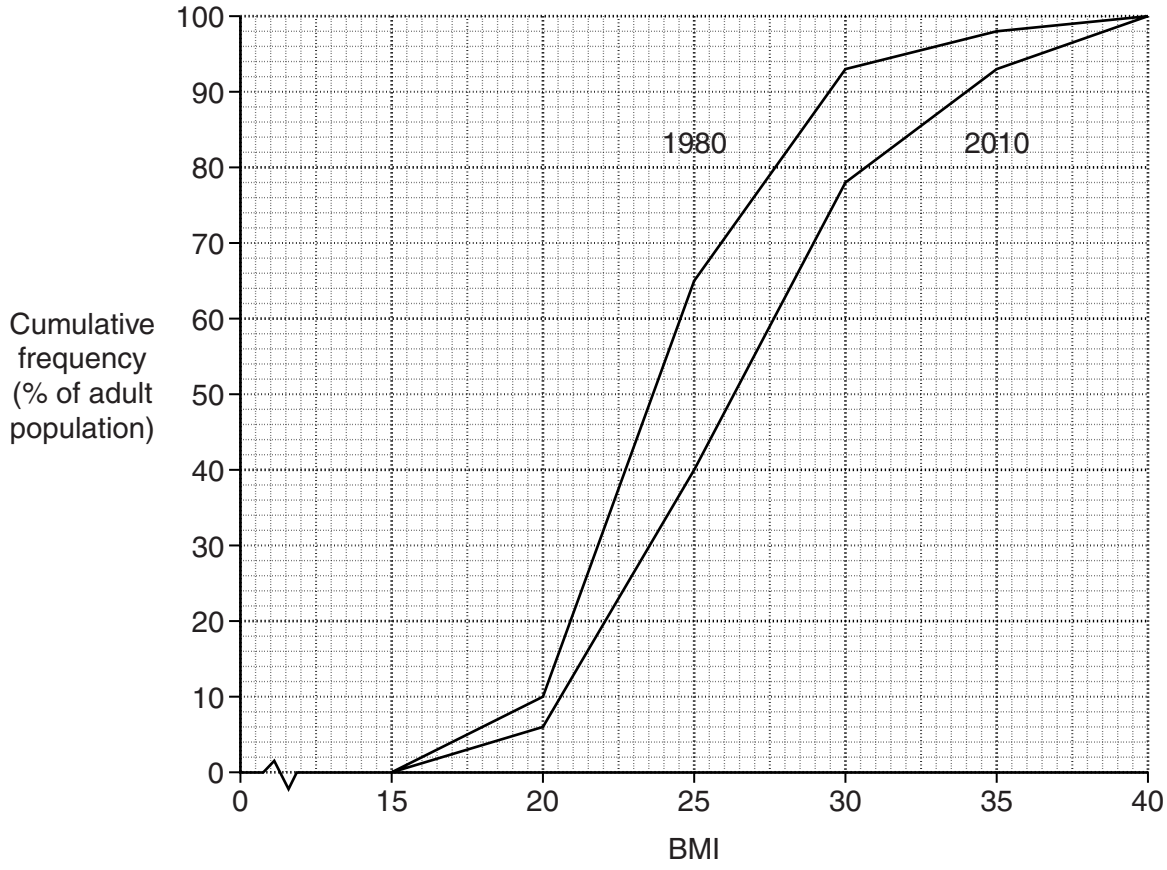
**(vi)** Of the junior competitors who completed the race in 2 hours or more, find the percentage who had chosen the challenging route.

.....[3]

**(vii)** Of all the senior competitors who had chosen the moderate route, find the percentage who completed the race in under 2 hours.

.....[3]

**9** One way to determine if an adult has a healthy weight, independent of age and g[illegible] measure their body mass index, BMI (a continuous variable).
The BMI values for the adult population of a particular country in the years 1980 and 20[illegible] summarised in the cumulative frequency polygons below.




Use these graphs to answer the following questions on the adult population of this country.

**(i)** Estimate

**(a)** the median BMI value in 1980,

.....[1]

**(b)** the median BMI value in 2010,

.....[1]

**(c)** the lower quartile BMI value in 1980,

.....[1]

**(d)** the upper quartile BMI value in 2010.

.....[1]

An adult's weight is classified as 'healthy' if their BMI value is between 18.5 and 25.

**(ii)** Estimate the percentage of the adult population whose weights were classified as 'hea

**(a)** in 1980,

.....[2]

**(b)** in 2010.

.....[1]

An adult is classified as 'overweight' if their BMI value is 25 or more.

**(iii)** Estimate the median BMI value of the 'overweight' adult population in 2010.

.....[3]

Adults with the highest BMI values are classified as 'obese'.
In 1980, 7% of the adult population were 'obese'.

**(iv)** Estimate the percentage of the adult population in 2010 who were 'obese'.

.....[4]

**(v)** By referring to any of the values you have estimated in parts **(i)**, **(ii)**, and **(iv)**, comment on how the health of the adult population of this country, assessed in terms of its weight, changed between 1980 and 2010.

.....

.....

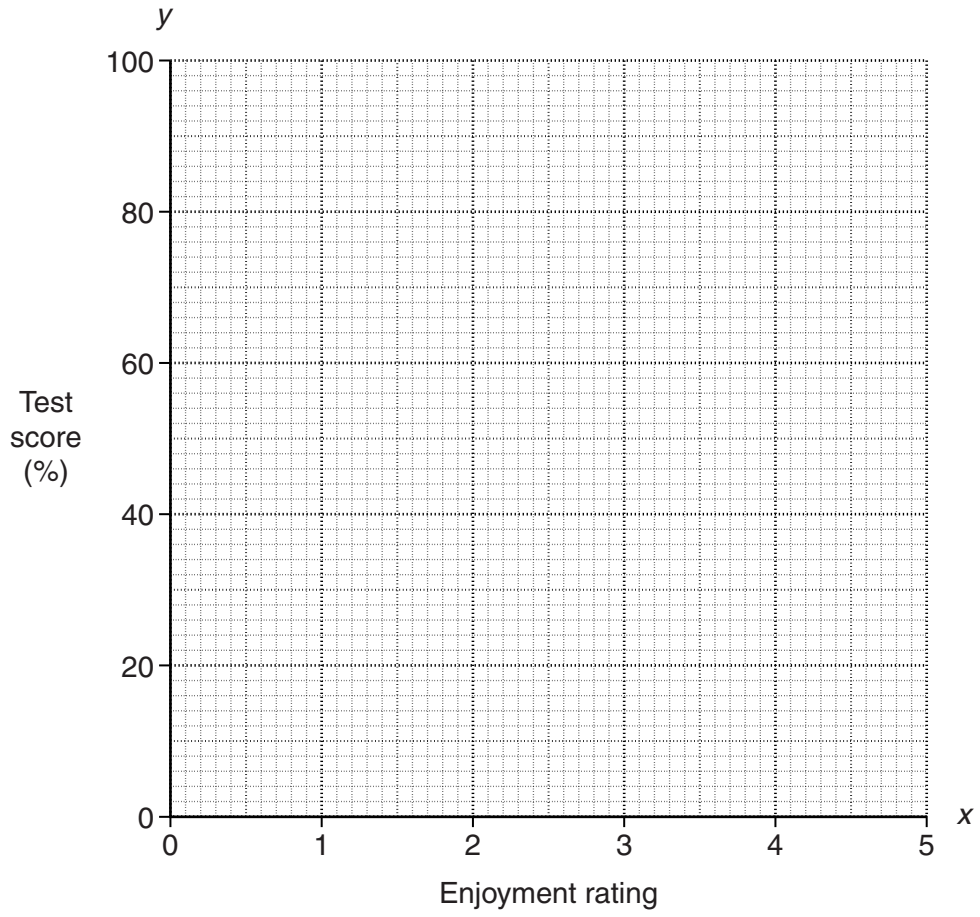
.....

.....[2]

**10** Baruti's teacher has suggested that pupils who enjoy studying a subject are likely to [illegible] in tests in the subject.
To investigate this, Baruti asked his friends to rate their enjoyment of Statistics on a linea[illegible] from 0 (dislike very much) to 5 (like very much), then recorded their scores on the next class[illegible] His results are shown in the following table.

| *Friend* | A | B | C | D | E | F | G | H |
|---|---|---|---|---|---|---|---|---|
| *Enjoyment rating*, *x* | 3 | 2 | 5 | 4 | 1 | 4 | 5 | 2 |
| *Test score* (%), *y* | 57 | 47 | 78 | 59 | 26 | 86 | 53 | 34 |

**(i)** Plot these data on the grid below.




[2]

**(ii)** Explain briefly why the points (5, 78) and (4, 59) are not used if the lower semi-average is calculated.

.....

.....[1]

**(iii)** Calculate the two semi-averages and the overall mean of the data, and plot t graph.

[5]

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

.....[4]

Another friend, who had rated his enjoyment of Statistics at 3, missed the test through illness.

**(v)** Use the line you have drawn in part **(iv)** to estimate the score this friend would have obtained if he had taken the test.

.....[1]

Baruti repeated his investigations for English and Science.
The equations he found for the lines of best fit were

$y = 1.24x + 53.8$ for English
and $y = 13.8x + 15.1$ for Science.

**(vi)** State, with a reason, in which of the subjects Statistics, English and Science a pupil's test score is most affected by their enjoyment rating.

.....

.....[2]

**(vii)** Explain briefly why Baruti may have experienced difficulty in deciding which of his two variables should be the independent and which the dependent.

.....

.....

.....[1]

**11** At a restaurant it is known from experience that 10% of the customers order an omele

Assume that customers make choices independently of each other and that nobody orde than one omelette.

**(i)** At table *A* there are 2 customers.

Find the probability that at this table an omelette is ordered by

**(a)** no customers,

.....[2]

**(b)** at least one customer.

.....[2]

The restaurant serves small omelettes and large omelettes. It is known from experience that 60% of those ordered are small and 40% are large.

**(ii)** At table *B* there are 4 customers.

Find the probability that at this table only one customer orders an omelette and it is a large omelette.

.....[4]

Small omelettes are made with 2 eggs and large omelettes with 3 eggs. The chef h[illegible] store of high quality eggs which are used only for making omelettes.

**(iii)** At table *C* there are 3 customers.

Find the probability that, in preparing the food for this table, from his special store the chef uses

**(a)** exactly 4 eggs,

.....[3]

**(b)** at most 4 eggs.

.....[5]

**BLANK PAGE**

**BLANK PAGE**

**BLANK PAGE**