UNIVERSITY OF CAMBRIDGE INTERNATIONAL EXAMINATIONS
General Certificate of Education Ordinary Level

CANDIDATE NAME			
CENTRE NUMBER		CANDIDATE NUMBER	

0910845626

STATISTICS **4040/22**

Paper 2 **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 **17** printed pages and **3** blank pages.

DC (LEO/JG) 51083/2

UNIVERSITY *of* CAMBRIDGE International Examinations

[Turn over

For Examiner's Use

Section A [36 marks]

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

1 **(i)** A student collected data about her Statistics textbook. She classifies her data into the following categories.

A a qualitative variable

B a discrete quantitative variable

C a continuous quantitative variable

D not a variable

For each of the following, enter into the table which of *A*, *B*, *C* or *D* is the correct description.

The number of words in each sentence on a randomly selected page.	
The number of pages in the book.	
The types of diagram used in the book.	
The number of letters in each of a random selection of 100 words.	
The area of each page used for diagrams.	

[4]

(ii) Another student collected data about the trees in the college grounds.

(a) Give an example of a discrete variable he could study.

....[1]

(b) Give an example of a continuous variable he could study.

....[1]

2 The total electricity production, measured in terawatt-hours (TWh), in 2010, of two countries *A* and *B* is shown in the table below.
The electricity production has been categorised as coming from fossil fuels, nuclear power or renewable sources.

For Examiner's Use

	Electricity production (TWh)	
	Country A	*Country B*
Fossil Fuels	334	109
Nuclear Power	165	53
Renewable	43	41
TOTAL	542	203

(i) Display the data in a percentage sectional (component) bar chart to allow the proportions of each source of electricity in each country to be compared.

[5]

(ii) Use your percentage sectional bar chart to compare the electricity production that comes from renewable sources in the two countries.

....

....[1]

3 The 30 members of a running club competed in a marathon. The grouped frequency distribution below shows the times they took to complete the race.

For Examiner's Use

Time taken, t (hours and minutes)	*Number of runners*	
2 h 50 m $\leqslant t <$ 3 h 00 m	5	
3 h 00 m $\leqslant t <$ 3 h 10 m	14	
3 h 10 m $\leqslant t <$ 3 h 20 m	7	
3 h 20 m $\leqslant t <$ 3 h 30 m	4	

(i) **Using an assumed mean of 3 hours 05 minutes**, find an estimate of the mean time taken by the runners, in hours and minutes, to the nearest minute.

....[5]

(ii) Explain why the mean is an appropriate average (measure of central tendency) to use in this case.

....

....[1]

4 **(a)** The values of a variable are grouped into classes labelled 0 – 9, 10 – 19, 20 – 29, etc.

State the mid-point of the 10 – 19 class, if the values are

(i) numbers of people,

....[1]

(ii) distances, measured to the nearest km,

....[1]

(iii) ages, measured in whole numbers of years.

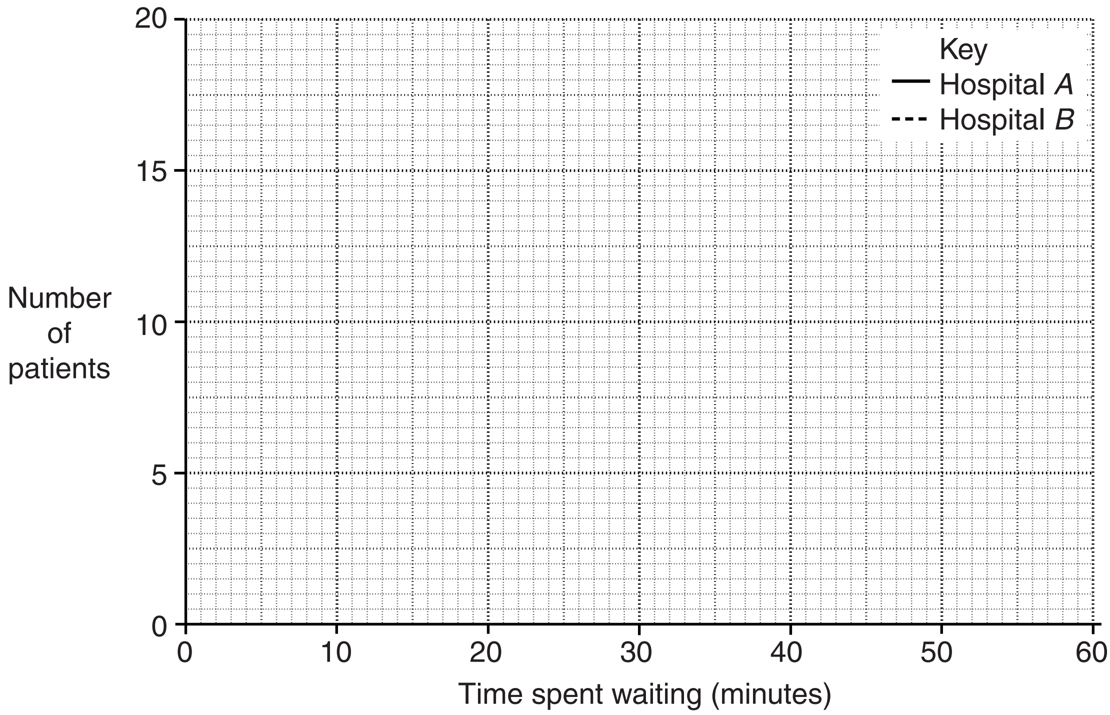
....[1]

(b) The length of time, to the nearest minute, that patients spent waiting to see doctors at two hospitals *A* and *B* were recorded on one particular morning.

For Examiner's Use

Length of time spent waiting (minutes)	*Hospital A*	*Hospital B*
0 – 9	0	0
10 – 19	5	3
20 – 29	17	7
30 – 39	6	13
40 – 49	2	7
50 – 59	0	0

(i) On the grid below, draw a frequency polygon for each hospital, using the given key.



[2]

(ii) Using your frequency polygons, compare the waiting times at the two hospitals.

....

....[1]

5 The pupils in a class sat a History examination and a Geography examination. The raw marks for these examinations are to be scaled so that they each have a mean of 60% and a standard deviation of 12%.

For Examiner's Use

In the History examination the mean and standard deviation of the raw marks of the pupils in the class were 48% and 8% respectively.

(i) Find the scaled History mark for a pupil whose raw mark in the History examination was 44%.

.... [2]

(ii) Find the History mark that will remain unchanged by the scaling process.

.... [2]

In the Geography examination the standard deviation of the raw marks of the pupils in the class was 10%. One particular pupil whose raw mark was 65% has a scaled mark of 78%.

(iii) Find the mean of the raw marks of the pupils in the class for the Geography examination.

.... [2]

6 The masses, in grams, of the 31 eggs produced on a small farm in one day are shown below.

For Examiner's Use

Mass, m (grams)	*Number of eggs*
$40 \leqslant m < 50$	1
$50 \leqslant m < 55$	5
$55 \leqslant m < 60$	13
$60 \leqslant m < 65$	8
$65 \leqslant m < 70$	3
$70 \leqslant m < 80$	1

(i) Without drawing a graph, calculate, to one decimal place, an estimate of the median mass of the eggs.

.... g [3]

Eggs with a mass over 63 grams can be sold as 'large'.

(ii) Without drawing a graph, calculate an estimate of the number of eggs produced on this day that could be sold as 'large'.

.... [3]

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 local government wishes to gather data from the pupils attending schools in its region on their attitudes to school.

(i) Give two reasons why the local government would choose to take a **sample** of the pupils for its survey.

Reason 1 ....

....

Reason 2 ....

.... [2]

A small school, *A*, in the region consists of 60 pupils, numbered 00 to 59, and a sample of six of its pupils is to be chosen for the survey.

(ii) Give a reason why each of the following sampling methods may lead to a biased sample:

(a) choosing the first six pupils to arrive at school on a particular day;

....

.... [1]

(b) choosing the pupils numbered 10 to 15.

....

.... [1]

EXTRACT FROM A RANDOM NUMBER TABLE

57	04	23	95	16	23	38	17	67	05	84	30	74	18	49	35	24
08	39	62	75	42	61	83	95	29	82	11	54	96	07	30	42	18

(iii) Using the random number table above, starting at the beginning of the first row and working along the row, find a **simple random sample** of six different pupils from school *A*.

.... [2]

For Examiner's Use

(iv) Another school, *B*, has 600 pupils, numbered 000 to 599. A **systematic sample** of size six is to be taken from this school.

(a) State the smallest and largest possible three-digit numbers of the first pupil selected.

.... [2]

The systematic sample is selected by starting at the beginning of the second row of the random number table and working along the row.

(b) Write down the number of the first pupil selected.

.... [1]

(c) Write down the numbers of the other five pupils selected for the systematic sample.

.... [2]

(v) Comment on the suitability of taking samples of the same size from both school *A* and school *B*.

....

.... [1]

(vi) It is decided to take a random sample of size 6 from school *A* and a random sample of size 60 from school *B*.

(a) State the name of this type of sampling.

.... [1]

(b) Using this method, state the number of pupils that would be selected from another school, *C*, which has 150 pupils.

.... [1]

(vii) Explain how a **quota sample** of pupils from the region might be selected.

....

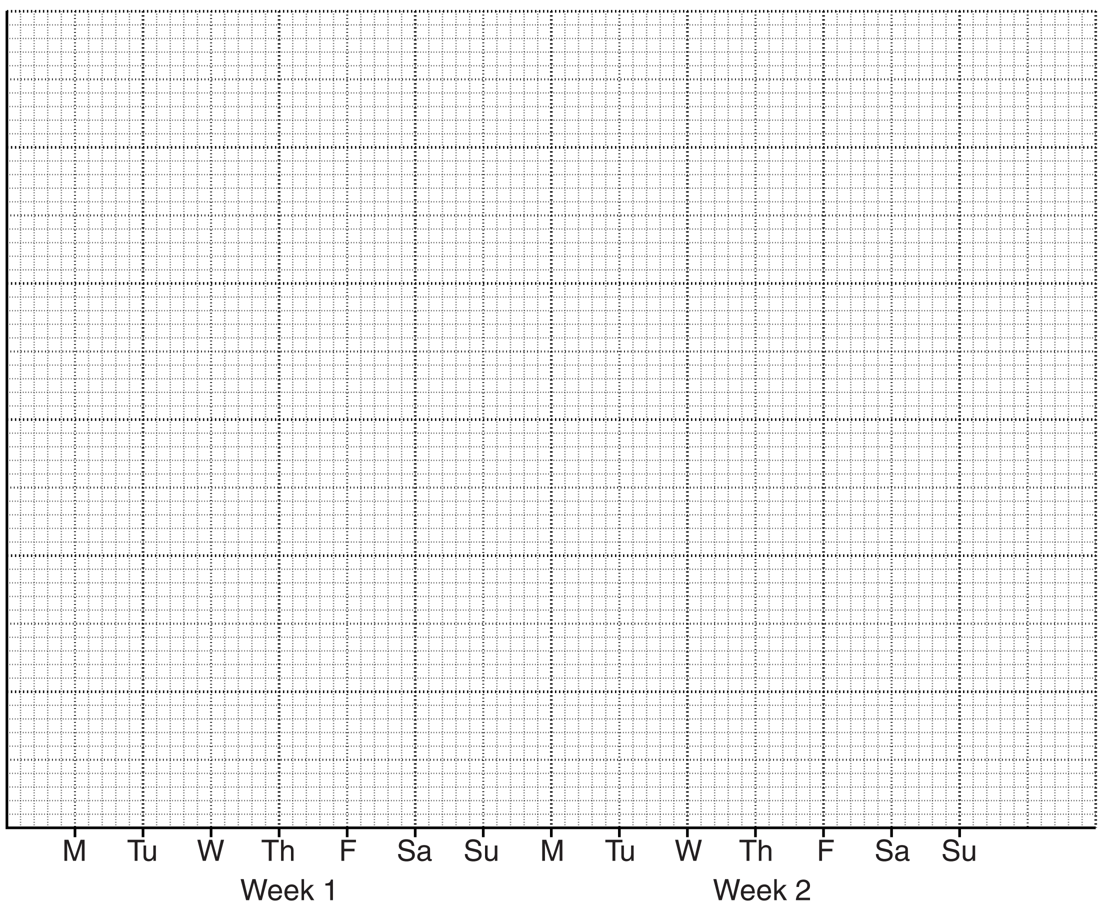
....

.... [2]

8 The money spent in a shop was recorded over a two-week period.

Week	*Day*	*Money spent* ($)	*7-day totals*	*7-day moving average values*
	Monday	162		
	Tuesday	143		
	Wednesday	138		
1	Thursday	138		
	Friday	149		
	Saturday	204		
	Sunday	90		
	Monday	155		
	Tuesday	130		
	Wednesday	123		
2	Thursday	132		
	Friday	142		
	Saturday	200		
	Sunday	88		

(i) Using the above data, draw, on the grid below, a time series graph of the money spent in the shop. Use an appropriate scale on the vertical axis.



[5]

For Examiner's Use

(ii) Explain why it is appropriate to calculate 7-day moving average values in this case.

....

....[1]

(iii) The 7-day totals, in order, are 1024, 1017, 1004, 989, 983, 976, 972 and 970. Insert these values in the correct places in the table. [1]

(iv) **(a)** Calculate the 7-day moving average values, correct to one decimal place, and insert them in the correct places in the table.

[1]

(b) Plot the moving average values on your time series graph, and draw a trend line through the points. [2]

(v) Explain what the trend line you have drawn tells you.

....

....[1]

(vi) The seasonal component for a Monday is 14.2. Estimate the money spent in the shop on the Monday of week 3.

$..[2]

(vii) Esme calculates the seasonal component for a Wednesday to be 12.4. State, with a reason, whether or not she is correct.

....

....[1]

(viii) Similar results were collected for another shop which does not open on Sundays. Explain whether it would be necessary to centre the moving average values calculated for this shop, giving a reason for your answer.

....

....

....[2]

9 Three boxes each contain 1 red, 2 blue and 3 green counters.
In a game, a turn consists of randomly selecting one counter from each box.

For Examiner's Use

(a) In this part of the question give all probabilities as exact fractions.

Katya plays one turn of the game. Find the probability that

(i) all three counters are red,

.... [2]

(ii) all three counters are the same colour,

.... [2]

(iii) exactly two of the counters are red,

.... [3]

(iv) exactly two of the counters are the same colour.

.... [2]

A player wins $10 if all three counters are the same colour, and $2 if exactly two counters are the same colour, otherwise they win nothing.

For Examiner's Use

(v) Find the cost of a turn if this is to be a fair game.

$.. [2]

(b) **In this part of the question give all probabilities as exact fractions or decimals correct to three decimal places.**

In another game, all 18 counters are put into one bag. Three counters are drawn at random from the bag, without replacement.

(i) Find the probability that the three counters are different colours.

.... [3]

(ii) Find the probability that two or more of the counters are the same colour.

.... [2]

For Examiner's Use

10 **(a)** Using 2009 as base year, the price relative of petrol in 2010 was 108.

(i) Explain what the price relative of 108 tells you.

....

....

....[3]

The cost of one litre of petrol in 2009 was \$0.64.

(ii) Find the cost, to the nearest cent, of one litre of petrol in 2010.

....[2]

Using 2009 as base year, the price relative of diesel in 2010 was 110.
The cost of one litre of diesel in 2010 was \$0.78.

(iii) Find the cost, to the nearest cent, of one litre of diesel in 2009.

....[2]

For Examiner's Use

(b) The pay of a company's employees falls into one of three grades *A*, *B* or *C*.
The table below shows the hourly rate of pay for each grade in each of three years.

	2009	2010	2011
Grade A	\$8.10	\$9.40	\$9.80
Grade B	\$9.50	\$10.50	\$10.50
Grade C	\$12.60	\$12.20	\$12.60

(i) Using 2009 as base year, calculate the price relatives for 2010 and 2011 for each grade and complete the table below.

	Price relative		
	2009	2010	2011
Grade A			
Grade B			
Grade C			

[5]

The table below shows the expenditure, in thousands of dollars to the nearest \$1000, on the wage bills of the company's employees at each grade in 2009.

	Grade A	*Grade B*	*Grade C*
Expenditure (\$'000)	350	270	90

(ii) Using the expenditure in 2009 as weights, find a weighted aggregate cost index for the company's wage bill in 2011, using 2009 as base year.

.... [3]

(iii) The result in part **(ii)** could be inaccurate if the weights have changed. Give one reason why the weights may have changed.

....

.... [1]

For Examiner's Use

11 (a) A six-sided dice with faces numbered 1, 2, 3, 4, 5 and 6 is rolled, and the score on the upper face is noted. Some of the possible outcomes are listed below.

A the score is 4

B the score is odd

C the score is 2 or 4

D the score is not 4

From this list, state all the possible pairs of mutually exclusive events.

.... [2]

(b) Events E and F are two outcomes of an experiment such that

$$P(E) = 0.8\,, \quad P(F) = 0.45 \quad \text{and} \quad P(E \cap F) = 0.36\,.$$

(i) State, with a reason, whether or not events E and F are mutually exclusive.

.... [2]

(ii) State, with a reason, whether or not events E and F are independent.

.... [2]

(iii) Find

(a) $P(E \cup F)$,

.... [2]

(b) $P(E' \cap F)$.

.... [3]

(c) An athlete's chances of winning a race depend on whether or not she eats breakfast. If the athlete eats breakfast the probability that she wins is $\frac{3}{5}$, whereas if she does not eat breakfast the probability is $\frac{1}{3}$. On any particular day the probability that she eats breakfast is $\frac{3}{4}$.

The athlete is going to run a race tomorrow. Find the probability that she

(i) eats breakfast and then does not win the race,

.... [2]

(ii) wins the race.

.... [3]

BLANK PAGE

BLANK PAGE

BLANK PAGE