UNIVERSITY OF CAMBRIDGE INTERNATIONAL EXAMINATIONS
General Certificate of Education Ordinary Level

CANDIDATE NAME			
CENTRE NUMBER		CANDIDATE NUMBER	

3373524824

STATISTICS **4040/23**

Paper 2 **October/November 2013**

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 **20** printed pages.

DC (LEG/CGW) 66936/4

UNIVERSITY *of* CAMBRIDGE
International Examinations

Section A [36 marks]

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

For Examiner's Use

1 State, for each of the following variables, whether it is discrete or continuous:

(i) the number of items of mail delivered each day to a particular address;

....[1]

(ii) the distances run by a number of athletes during 1 hour.

....[1]

The variables described above are each grouped into classes labelled 0 – 4, 5 – 9, 10 – 14 etc.

State the true lower and upper class limits for the 5 – 9 class for

(iii) the variable described in **(i)**,

....[2]

(iv) the variable described in **(ii)**, after the distances have been rounded to the nearest integer.

....[2]

2 Give a brief explanation of the meaning of each of the following terms when used in the calculation of index numbers:

For Examiner's Use

(i) base year;

....

....

.... [2]

(ii) weight;

....

....

.... [2]

(iii) price relative.

....

....

.... [2]

For Examiner's Use

3 The body lengths (including the tail) of a sample of 45 white-footed Texas mice were measured in millimetres. 25 of the mice were found to be male and 20 female. The following table summarises the data obtained on mouse length.

	Number of mice	*Sum of lengths*	*Sum of squares of lengths*
Male	25	4325	748 369
Female	20	3060	468 252

(i) Explain why the mean length of the total sample of 45 mice is not just given by (mean length of male mice + mean length of female mice) / 2.

....

....[1]

(ii) Calculate, to 1 decimal place, the mean and the standard deviation of the lengths of the total sample of 45 mice.

Mean = ...

Standard deviation = ...[5]

4 Values of experimental readings taken by different people are to be scaled for purposes of comparison. The readings have a mean of 37 and a standard deviation of 5.
The scaled values are to have a mean of 100 and a standard deviation of 10.

For Examiner's Use

Calculate

(i) the scaled value corresponding to a reading of 55,

.... [2]

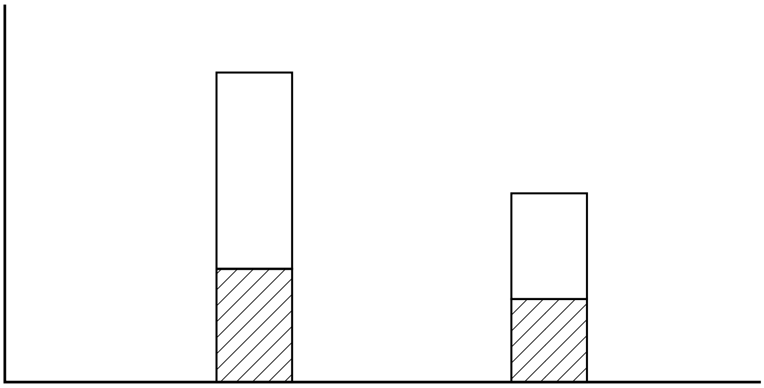
(ii) the reading corresponding to a scaled value of 87.5,

.... [2]

(iii) the reading which is unaltered when scaled.

.... [2]

5



The bar chart above is intended to illustrate information about how many boys and girls attend each of two schools, *A* and *B*.

(i) The bar chart is incomplete. List three items of detail which are missing.

....

....

.... [2]

(ii) State the name of this type of bar chart.

.... [1]

(iii) Explain how you know that the bar chart illustrates the actual number of boys and girls, and not percentages.

....

.... [1]

(iv) Another type of diagram which could be used to illustrate the data is a pictogram. State a disadvantage of pictograms, compared with bar charts, when illustrating frequencies such as the number of pupils at a school.

....

.... [1]

(v) Give a reason why a change chart could not be used to illustrate these data.

....

.... [1]

For Examiner's Use

For Examiner's Use

6 A farmer classifies the expenditure in running his farm under four headings: Animal Feed, Labour, Fuel and Professional Services (e.g. veterinary services). The price relatives for each of these headings for the year 2011, taking 2006 as base year, and the weight allocated by the farmer to each heading are given in the following table.

	Price relative	*Weight*
Animal Feed	104	14
Labour	110	6
Fuel	107	4
Professional Services	102	3

(i) Calculate, correct to 2 decimal places, the overall percentage increase in the farmer's weighted cost index from 2006 to 2011.

.... [4]

(ii) In 2011 the farmer's income was 7% greater than it had been in 2006. State, with a reason, whether or not the farm was more profitable than it had been five years earlier.

....

.... [2]

Section B [64 marks]

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

Each question in this section carries 16 marks.

7 This question must be answered by calculation. An answer using a graphical method will not be awarded any marks.

The following table summarises the heights, in centimetres, of a sample of 8585 adult males in the United Kingdom.

Height (cm)	*Frequency*	*Cumulative frequency*
150 – under 160	144	
160 – under 165	1232	
165 – under 170	2213	
170 – under 175	2559	
175 – under 180	1709	
180 – under 190	705	
190 – under 200	23	

(i) Calculate the cumulative frequencies and insert them in the table. [2]

(ii) (a) State the class in which the median height lies.

.... [1]

(b) Estimate, to 1 decimal place, the median height.

.... cm [3]

(iii) (a) State the class in which the lower quartile height lies.

.... [1]

(b) Estimate, to 1 decimal place, the lower quartile height.

.... cm [3]

The upper quartile height, correct to 1 decimal place, is 175.9 cm.

For Examiner's Use

(iv) **(a)** Estimate the interquartile range of the heights.

.... cm [1]

(b) Compare the distances of the quartiles from the median, and comment on whether this is what you would expect in a distribution of the heights of a large number of adult males.

....

....

....

.... [3]

(v) If a cumulative frequency curve were drawn to illustrate this distribution, state, with a reason, in which part of the graph the curve would be at its steepest.

....

.... [2]

8 **In this question give all answers as fractions in their lowest terms.**

For Examiner's Use

Two identical bags each contain a number of coloured balls.
Bag X contains 4 white and 7 blue balls. Bag Y contains 3 blue and 8 red balls.

(i) A bag is chosen at random and a ball selected at random from it. Find the probability that the selected ball is blue.

.... [3]

(ii) Two balls are chosen at random from bag Y. Find the probability that they are of the same colour.

.... [3]

(iii) One ball is chosen at random from each bag. Find the probability that the chosen balls are of the same colour.

.... [4]

(iv) A bag is chosen at random and two balls are selected at random from it. Find the probability that both selected balls are white.

.... [3]

(v) All the balls from both bags are emptied into a third bag, bag *Z*. Two balls are then chosen at random from bag *Z*. Find the probability that both selected balls are white.

For Examiner's Use

.... [2]

(vi) Explain briefly why the answer to part **(iv)** is greater than the answer to part **(v)**.

....

.... [1]

9 Three unbiased six-sided dice, each with faces numbered 1, 2, 3, 4, 5 and 6, are rolled simultaneously.

Find the probability that the numbers on the uppermost faces will be

(i) three 1s,

.... [1]

(ii) three of the same number except 1,

.... [1]

(iii) exactly two 1s and some other number.

.... [3]

A game, in which three such dice are rolled simultaneously and for which the entry fee is $1, is organised. Prizes are paid for certain outcomes on the uppermost faces, as given in the following table.

Outcome	*Prize paid* ($)
Three 1s	6
Three of the same number except 1	4
Exactly two 1s and some other number	3

(iv) Calculate, to the nearest cent, the organiser's expected profit each time the game is played.

.... [3]

In another game, a contestant chooses three cards at random from a set of ten. The numbers on the cards are 1, 1, 1, 2, 2, 2, 3, 4, 5 and 6. Prizes are again paid as given in the previous table.

For Examiner's Use

(v) By first calculating the appropriate probabilities, calculate, to the nearest cent, the entry fee which should be charged to make this a fair game.

.... [8]

10 **(a)** A large housing estate contains approximately equal numbers of three types of dwelling: detached houses (D), semi-detached houses (S) and bungalows (B). A research organisation wishes initially to get some idea of how many occupants there tend to be in each type of dwelling. It has instructed an interviewer to call at four of each type of dwelling to ask how many people live there – but the choice of exactly which dwellings is up to the interviewer.

(i) State the name of the method of sampling being used.

.... [1]

(ii) Give a reason why the research organisation could not just simply use a list of registered voters for the estate.

....

.... [1]

The interviewer labelled his chosen dwellings 1 to 12, and the following is a copy of the notes he made during a number of visits to the estate:

ad = adult(s) ch = child(ren)

1	B	2ad
2	S	3ch 2ad
3	S	2ad 4ch
4	D	7ch 2ad
5	B	call again later
6	D	2ad 5ch
7	D	2ad 5ch
8	B	no reply
9	S	4ch 2ad
10	S	no reply
11	D	2ad
12	B	1ch 1ad
5		call again
8		still no reply
10		2ad
5		2ad
8		still no reply
8		2ad

(iii) For the twelve dwellings chosen, find the total number of

(a) adults,

.... [1]

(b) bungalows with no children.

.... [1]

(iv) Draw up and complete a table showing the number of dwellings, classified by their type and by the number of children who live in them.

For Examiner's Use

[3]

(v) The research organisation is to carry out a survey on behalf of a manufacturer of children's clothes. If it only has sufficient funding to investigate the expenditure on such clothes by the inhabitants of one type of dwelling, state, with a reason, which type it should choose.

....

....

....[2]

[Question 10 continues on the next page]

(b) A group of 60 people are each allocated a different two-digit random number in the range 01 to 60. The 20 men are numbered 01 to 20 and the 40 women are numbered 21 to 60.
A sample of **size six** is to be selected by different sampling methods using the following random number table, starting at the beginning of the row for each sample. No person may be selected more than once in any one sample.

RANDOM NUMBER TABLE

21 32 07 42 98 81 21 57 81 59 31 17 36

Select

(i) a simple random sample,

....[2]

(ii) a systematic sample,

....[3]

(iii) a sample stratified by gender, using every number if the gender to which it relates has not yet been fully sampled.

....[2]

For Examiner's Use

11 **(a)** **(i)** A company's sales are recorded every month over a period of several years. Use this example to explain briefly the meaning of the term

(a) trend,

....

....

....[1]

(b) seasonal variation,

....

....

....[1]

(c) cyclic variation.

....

....

....[1]

(ii) State which one of trend, seasonal variation and cyclic variation the method of moving averages removes from a time series, and explain briefly how this is achieved.

....

....

....[2]

For Examiner's Use

(b) The following table gives the number of properties sold during each quarter of the years 2006 to 2009 by a small estate agent, together with values of relevant totals and moving averages.

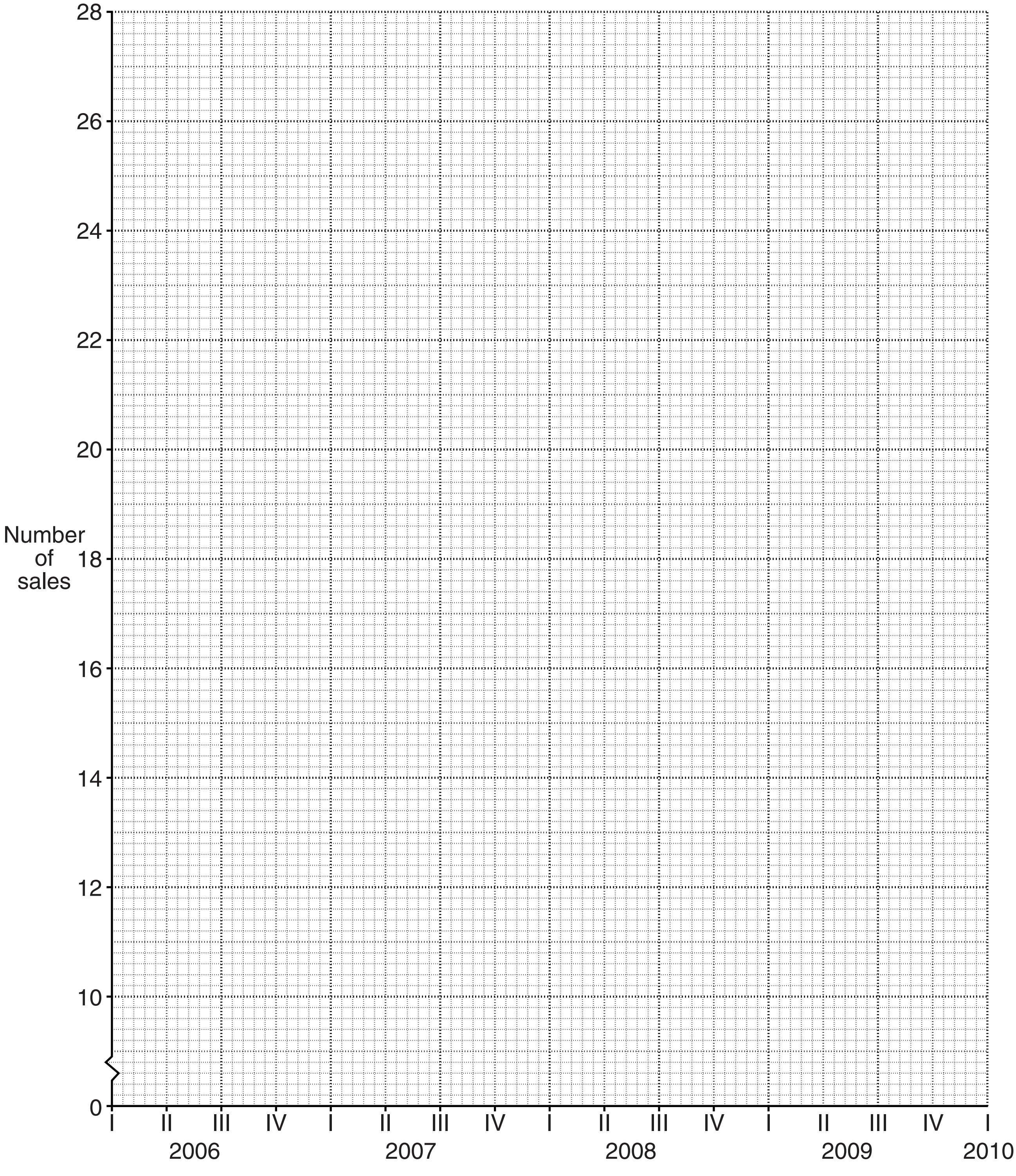
Year	*Quarter*	*Number of sales*	*4-quarter total*	*8-quarter total*	*8-quarter moving average value*
2006	I	18			
	II	24			
			96		
	III	28		197	24.625
			101		
	IV	26		205	25.625
			104		
2007	I	23		209	26.125
			$x =$		
	II	27		209	26.125
			104		
	III	29		$y =$	25.625
			101		
	IV	25		197	24.625
			96		
2008	I	20		188	23.5
			92		
	II	22		180	$z =$
			88		
	III	25		167	20.875
			79		
	IV	21		151	18.875
			72		
2009	I	11		137	17.125
			65		
	II	15		121	15.125
			56		
	III	18			
	IV	12			

(i) Calculate the values of x, y and z and insert them in the table.

[3]

(ii) On the grid below plot the 8-quarter moving average values. [2]

For Examiner's Use



(iii) Describe what your plotted points show about the sales of properties during this time period.

....

.... [2]

[Question 11 continues on the next page]

(iv) State, with a reason, whether or not it would be meaningful to draw a single straight trend line through the plotted points.

....

....[1]

(v) Draw a straight trend line which would be useful for estimating the number of properties sold in the first quarter of 2010. [1]

The seasonal components for the number of sales are given in the following table.

Quarter	I	II	III	IV
Seasonal component	-4.4	-0.1	3.5	1.0

(vi) State, with a reason, whether the actual sales of properties in the first quarter of 2010 would be likely to be greater or smaller than the value indicated by your trend line.

....

....[2]