CANDIDATE NAME	

CENTRE NUMBER				CANDIDATE NUMBER	

STATISTICS   **4040/23**

Paper 2   **October/November 2011**

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

UNIVERSITY *of* CAMBRIDGE
International Examinations

[Turn over

Section A [36 marks]

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

1 When a children's choir had ten members the averages (measures of central tendency) of their ages were as follows.

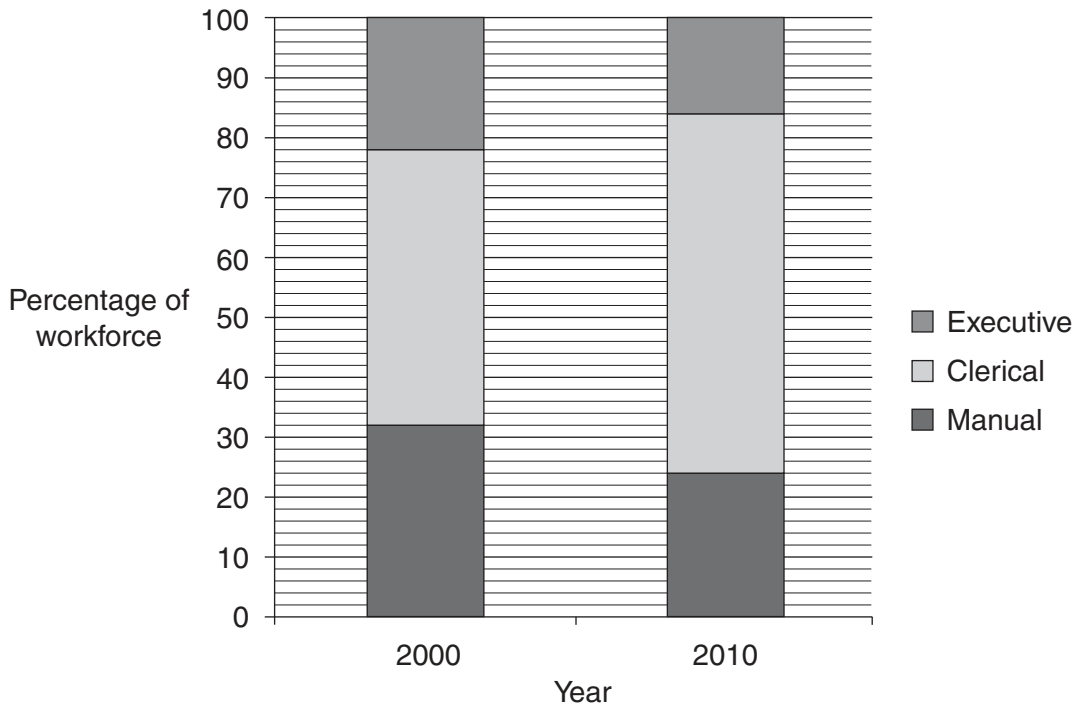Mean = 11.2 years   Median = 11.5 years   Mode = 13 years

Two new children aged 7 years and 13 years joined the choir. Find the mean, median and mode of the ages of the twelve members of the choir.

Mean = ... years

Median = ... years

Mode = ... years [5]

2 The component percentage bar chart below shows the breakdown of the workforce at a particular company in the years 2000 and 2010.

In the year 2000 there were 250 employees at the company.

(i) Find the number of each type of worker in the year 2000.

Executive ....

Clerical ....

Manual .... [3]

(ii) A student looking at the graph states that the number of clerical workers at the company has increased between the years 2000 and 2010.

State, giving a reason, whether the student's statement is correct.

....

....

.... [2]
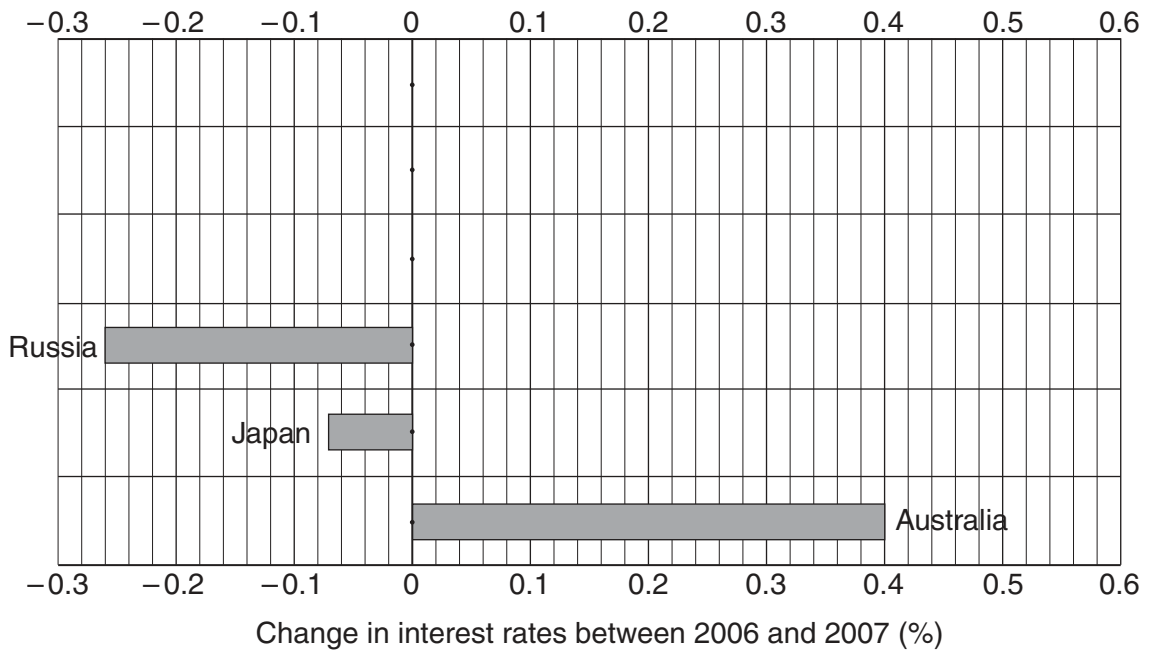
3 The table below shows the long-term interest rates for 2006 and 2007 in various countries.

	Interest rate (%)	
	2006	2007
Australia	5.62	6.02
Japan	1.74	1.67
Russia	6.98	6.72
South Africa	7.94	7.99
United Kingdom	4.50	5.01
United States	4.79	4.63

(i) Complete the change chart below.

Change in interest rates between 2006 and 2007 (%)

[2]

(ii) Give one advantage and one disadvantage of displaying data using a change chart.

Advantage ....

....

....

Disadvantage ....

....

.... [2]

4 **(i)** Explain, in words, the meaning of the term 'mutually exclusive events'.

....

....

.... [1]

(ii) Events *A* and *B* are mutually exclusive. P(*A*) = 0.3 and P(*A*∪*B*) = 0.8.

Find P(*B*).

.... [1]

(iii) Events *C* and *D* are independent. P(*C*) = 0.4 and P(*C*∩*D*) = 0.1.

Find

(a) P(*D*),

.... [2]

(b) P(*C*∪*D*).

.... [2]

[Turn over

5 A variable *V* with the frequency distribution below has mean *m* and standard deviation *s*.

V	1	2	3	4	5
Frequency	*g*	*h*	*i*	*j*	*k*

(i) Find the mean and standard deviation, in terms of *m* and *s*, of

(a) the variable *W* with the frequency distribution

W	5	6	7	8	9
Frequency	*g*	*h*	*i*	*j*	*k*

,

Mean = ...

Standard deviation = ... [2]

(b) the variable *X* with the frequency distribution

X	4	8	12	16	20
Frequency	*g*	*h*	*i*	*j*	*k*

.

Mean = ...

Standard deviation = ... [2]

(ii) *m* and *s* are found to be 3.18 and 1.12 (correct to 2 decimal places). Give an estimate of the mean and standard deviation of the continuous variable *Y* with the grouped frequency distribution

Y	$10 \leqslant y < 12$	$12 \leqslant y < 14$	$14 \leqslant y < 16$	$16 \leqslant y < 18$	$18 \leqslant y < 20$
Frequency	*g*	*h*	*i*	*j*	*k*

.

Mean = ...

Standard deviation = ... [4]

6   EXTRACT FROM A RANDOM NUMBER TABLE

15 37 72 50 37 02 49 91 21 82 16 94 54 09 65 98
73 11 02 30 33 91 47 28 19 28 56 04 22 51 83 24

(i) A population has a size of 60, and the individuals in that population are numbered 00 to 59. A sample is to be selected from the population and no individual may be selected more than once in any one sample.

  (a) Use the random number table above, starting at the beginning of the first row and working along the row, to find a **simple random sample** of size five.

.... [2]

  (b) (i) Use the second row of the table to find the first number for a **systematic sample** of size five.

.... [1]

  (ii) Write down the numbers of the other members of the sample.

.... [1]

  (c) The population consists of 19 males and 41 females. A **sample stratified by gender**, of size five, is to be taken. Calculate how many males and females will be required in the sample.

Males ....

Females .... [2]

(ii) Another population has a size of 600. Explain how the random number table above could be used to select a sample from this population.

....

....

....

.... [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 The owner of a small taxi firm classified the expenditure on her business into three categories, as shown in the table.

The price relatives, taking 2008 as the base year, are also given in the table.

	Price relatives		
	2008	2009	2010
Maintenance	100	96	112
Wages	100	$a =$	113
Fuel	100	$b =$	116

(i) Explain what the value 96 for maintenance in 2009 indicates.

....

.... [2]

(ii) Explain what the value 112 for maintenance in 2010 indicates.

....

.... [2]

(iii) Calculate, to the nearest whole number, the price relative for maintenance for 2010 taking **2009** as the base year.

.... [2]

(iv) The wage rates of her employees increased by 7% between 2008 and 2009. Insert the value of *a* in the table. [1]

(v) The price of fuel rose from $1.90 per litre in 2008 to $2.09 per litre in 2009. Calculate the value of *b* and insert it in the table.

[2]

In 2008, she spent 15% of her total expenditure on maintenance, 35% on wages and 50% on fuel.

(vi) Calculate, to one decimal place, a weighted aggregate cost index for 2010 using 2008 as the base year.

.... [3]

(vii) Her total expenditure on the business came to $6420 in 2008. Use your answer to part **(vi)** to estimate, to the nearest dollar, her total expenditure on the business in 2010.

$ .... [2]

(viii) Give two possible reasons why this estimate might be very inaccurate.

Reason 1 ....

....

Reason 2 ....

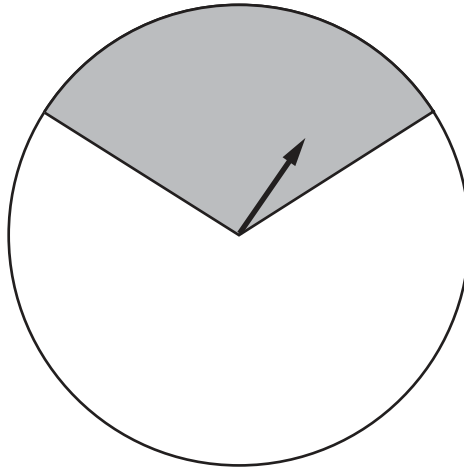... [2]

8 In this question give all probabilities as exact fractions.

A turn at a fairground game involves spinning an arrow once. Each turn costs $1. A player wins a prize of $2 if the arrow lands in the shaded area, otherwise the player loses. The probability of landing in the shaded area is $\frac{1}{3}$.

Three friends Amir, Baljit and Cody, each with $2, decide to play.

(i)   Amir decides to have two turns.

(a)   Find the probability that Amir wins on both turns.

.... [2]

(b)   Find the probability that Amir wins exactly once.

.... [3]

(c)   Let *X* represent the money possessed by Amir after playing. Find the expectation of *X*.

$.. [2]

(ii) The second friend, Baljit, decides to stop playing if he wins. Let *Y* represent the money possessed by Baljit after playing. Using *W* for win and *L* for lose, the possible outcomes are listed in the table below.

Outcome	W	LW	LL
Probability			
Y			

(a) Calculate the probability of each outcome and enter it in the above table.

[3]

(b) Remembering that Baljit arrived with $2, for each outcome enter the value of *Y* in the above table. [1]

(c) Which of Amir and Baljit had the better strategy for winning more money? Justify your answer.

....

.... [2]

(iii) The third friend, Cody, decides to keep playing for as long as he still has money. Find the probability that he has fewer than five turns.

.... [3]

   [Turn over

9 (a) Some students were asked about their journey to school on a particular day and the following data was collected.

(i) The length of time taken on their journey to school.

(ii) The method of transport they used for their journey.

(iii) The number of people they travelled with.

For each, state whether the data collected is qualitative or quantitative and whether it is discrete or continuous.

	Qualitative or *Quantitative*	*Discrete* or *Continuous*	
(i)			[1]
(ii)			[1]
(iii)			[1]

(b) The following data was collected regarding the length of time, **to the nearest minute**, that 120 students took on their journey to school.

Time taken (minutes)	*Number of pupils*	*Cumulative frequency*
0 – 9	1	
10 – 19	14	
20 – 24	30	
25 – 29	29	
30 – 34	21	
35 – 39	12	
40 – 49	7	
50 – 59	3	
60 – 69	3	

(i) Explain why the true class limits of the 20 – 24 class are 19.5 and 24.5.

....

.... [1]

(ii) Insert the cumulative frequencies in the table. [1]

(iii) State the class which contains the median.

.... [1]

(iv) Calculate, using linear interpolation, an estimate of the median, correct to one decimal place.

.... minutes [3]

(v) Explain why this value is an estimate.

....

.... [1]

(vi) Without any further calculation, state whether an estimate of the mean would be higher or lower than your result in part **(iv)**. Give a reason for your answer.

....

....

.... [2]

(vii) Calculate, using linear interpolation, an estimate of the number of pupils who took over 45 minutes to get to school.

.... [4]

10 In this question give all answers as exact fractions or as decimals correct to 3 decimal places.

The two-way table below shows the number of siblings (brothers and sisters) and the gender of the children in a class.

	Number of siblings				
	0	1	2	3 or more	TOTAL
Male	2	5	7	2	16
Female	1	7	3	1	12
TOTAL	3	12	10	3	28

For example, there are 3 females who each have 2 siblings.

(i) Find the probability that a child chosen at random from the class

(a) is male and has one sibling,

.... [1]

(b) has fewer than two siblings,

.... [1]

(c) is female, given that they have one sibling.

.... [1]

(ii) The members of a committee of two children are chosen at random. Find the probability that

(a) they both have two siblings,

.... [3]

4040/23/O/N/11

(b) one male and one female are chosen,

.... [3]

(c) they have altogether a total of two siblings.

.... [3]

(iii) A committee of three children is chosen at random. Find the probability that exactly two of them have no siblings.

.... [4]

11 In a particular country the academic year runs from September in one year to July in the following year, and is split into three terms.

The numbers of fiction books borrowed from a college library in that country are given in the table below.

		Year		
		2007/8	2008/9	2009/10
Term	1	242	254	269
	2	239	258	268
	3	184	193	204

(i) Describe the seasonal variation and suggest a possible reason for it.

....

....

.... [2]

(ii) Use the data above to plot a time series graph on the grid below, joining the plots with straight line segments.

Number of
fiction books
borrowed

270
260
250
240
230
220
210
200
190
180
0

1   2   3   1   2   3   1   2   3   1   2   3
  2007/8   2008/9   2009/10   2010/11

Term and year
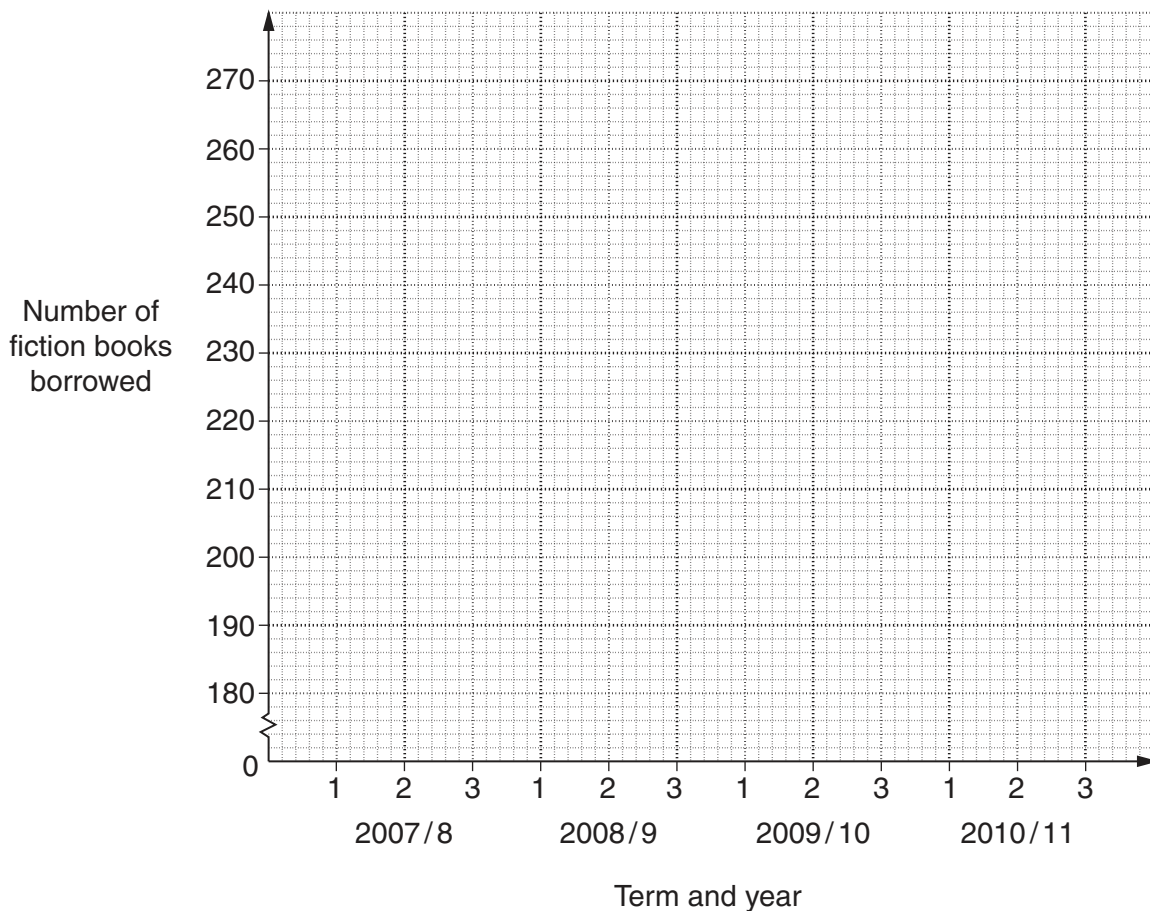
[2]

(iii) Explain the purpose of plotting moving average values.

....

....

.... [2]

(iv) Setting your work out in an appropriate table, find the 3-point moving average values for this data.

[3]

(v) Plot the 3-point moving average values on your time series graph and draw a trend line through the points. [3]

The seasonal components for this data are summarised in the table below.

	Term 1	Term 2	Term 3
Seasonal components	23.8	q	−44.3

(vi) Calculate the value of q.

q = ... [2]

(vii) Using your trend line and the appropriate seasonal component, estimate the number of fiction books that will be borrowed in term 2 of 2010/11.

.... [2]

BLANK PAGE

BLANK PAGE

20

BLANK PAGE

4040/23/O/N/11