

CANDIDATE
NAME

--

CENTRE
NUMBER

--	--	--	--	--

CANDIDATE
NUMBER

--	--	--	--



STATISTICS

Paper 2

4040/23

October/November 2015

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.

Section A [36 marks]

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

- 1 The ages, in completed years, of the children attending two swimming clubs, Fun Swim and Aqua Splash, are shown in the table below.

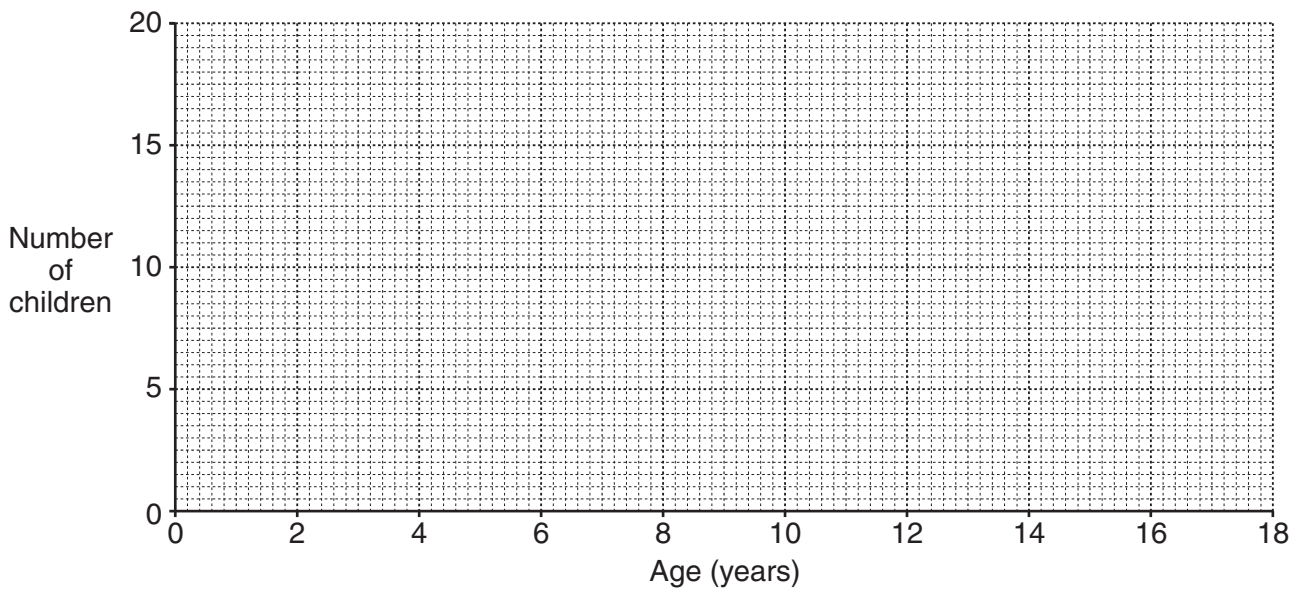
<i>Age (years)</i>	<i>Fun Swim</i>	<i>Aqua Splash</i>
0 – 2	0	0
3 – 5	8	4
6 – 8	16	11
9 – 11	12	14
12 – 14	6	13
15 – 17	0	0

- (i) State the lower and upper class boundaries of the 9 – 11 class.

Lower class boundary

Upper class boundary[1]

- (ii) On the grid below, draw frequency polygons to show the children’s ages at each swimming club.



[4]

- (iii) Use the frequency polygons to compare the children’s ages at each swimming club.

.....
[1]

- 2 Each student in a class took a Physics test, a Chemistry test and a Biology test. The scores for the class are summarised in the table below.

	<i>Physics</i>	<i>Chemistry</i>	<i>Biology</i>
<i>Mean</i>	45.1	62.7	48.5
<i>Standard deviation</i>	8.2	7.4	<i>a</i>

The students' scores on each test were then scaled to a mean of 50 and a standard deviation of 10.

- (i) Salma had a scaled mark of 45 in the Chemistry test. Find her actual score in the Chemistry test.

.....[2]

- (ii) Peter had the same scaled mark in Physics and Chemistry. He had an actual score of 82 in the Physics test. Find his actual score in the Chemistry test.

.....[2]

- (iii) In the Biology test Kumar had an actual score of 39 and a scaled mark of 37.5. Find *a*, the standard deviation of the actual scores in the Biology test.

.....[2]

3 (a) (i) Given that $P(A') = 0.7$, $P(B) = 0.6$ and $P(A \cup B) = 0.7$, find $P(A \cap B)$.

.....[3]

(ii) Hence explain whether or not A and B are independent events.

.....
.....[2]

(b) Four cards are numbered 1, 2, 3 and 4. The following are possible events when one card is selected at random.

C : an even number is chosen

D : an odd number is chosen

E : the 3 or the 4 is chosen

F : the 4 is chosen

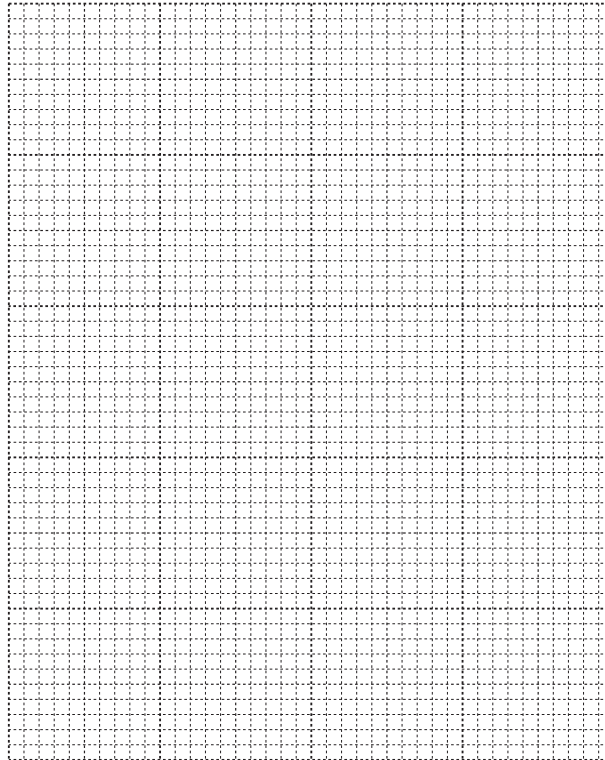
From the above, list all the pairs of mutually exclusive events.

.....[1]

- 4 The areas in km² of various types of land in two countries *A* and *B* are shown in the table below. The types of land have been categorised as being Urban, Farmland or Other (including forest, desert, lakes etc.).

	Area (km ²)	
	Country A	Country B
<i>Urban</i>	300	250
<i>Farmland</i>	1200	850
<i>Other</i>	1900	700
TOTAL	3400	1800

- (i) Display the data in a percentage sectional (component) bar chart to allow the proportions of the area of each type of land in each country to be compared.



[4]

- (ii) Use the table and your percentage sectional bar chart to make two statements comparing the area of urban land in the two countries.

1

.....

2

.....[2]

- 5 (a) Give one advantage and one disadvantage of taking a sample rather than obtaining data from the whole population.

Advantage

.....

Disadvantage

.....[2]

- (b) For each of the following statements, state whether it is always true, sometimes true, or never true. Give a reason for each of your answers.

- (i) A representative sample of people should contain equal numbers of males and females.

.....

.....

.....[2]

- (ii) A sampling method which produces a sample consisting of the people numbered 000 to 029 from a population listing numbered 000 to 599 is biased.

.....

.....

.....[2]

- 6** A museum monitored the number of visitors it had.
In a particular week, for the six days from Monday to Saturday, the mean daily attendance was 38 and the variance was 71.
When the visitor numbers for the Sunday of that week were also included, the mean for all seven days became 39.

Find

- (i)** the number of visitors to the museum on the Sunday of that week,

.....[2]

- (ii)** the variance of the number of visitors for all seven days of that week, correct to 1 decimal place.

.....[4]

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 number of bags each contain 2 white beads and 3 black beads.

(i) Two such bags are taken and a bead is selected at random from each bag.

Find the probability of selecting one bead of each colour.

.....[4]

(ii) Three such bags are taken and a bead is selected at random from each bag.

Find the probability of selecting at least two black beads.

.....[5]

(iii) One such bag is chosen and three beads are selected at random, without replacement.

Find the probability of selecting both white beads.

.....[4]

(iv) Another such bag is chosen. All the beads are selected one at a time and threaded onto a string.

Find the probability that the beads on the string alternate in colour.

.....[3]

8 A small village has 120 residents. The table below summarises their ages.

<i>Age, x years</i>	<i>Number of people</i>	<i>Cumulative frequency</i>
$0 \leq x < 10$	16	
$10 \leq x < 20$	21	
$20 \leq x < 30$	33	
$30 \leq x < 40$	22	
$40 \leq x < 60$	16	
$60 \leq x < 80$	8	
$80 \leq x < 100$	4	

(i) Calculate the cumulative frequencies of the data and insert them in the final column of the table. [1]

(ii) Use linear interpolation to calculate an estimate of the median age of the residents, correct to one decimal place.

.....[4]

(iii) 20% of the residents are aged p years or more. Use linear interpolation to calculate an estimate of p .

.....[4]

- (iv) Use linear interpolation to calculate an estimate of the percentage of people in the village aged 18 and over but less than 67.

.....[5]

- (v) State why your answers to parts (ii), (iii) and (iv) are only estimates and explain what assumption you are making when calculating these estimates.

.....
.....
.....[2]

- 9 A company employs three grades of worker: skilled, semi-skilled and unskilled. The manager of the company wishes to calculate a weighted aggregate cost index for the total wage bill of the company.

The incomplete table 1 below is to show the rates of pay per hour for each of these grades for three consecutive years.

The incomplete table 2 below is to show the price relatives of the rates of hourly pay, for these three years, taking 2011 as base year.

Table 1

	<i>Rate of pay per hour (\$)</i>		
	2011	2012	2013
<i>Skilled</i>	8.10	8.52	8.36
<i>Semi-skilled</i>	7.50		
<i>Unskilled</i>	7.20	7.01	7.01

Table 2

	<i>Price relative</i>		
	2011	2012	2013
<i>Skilled</i>			
<i>Semi-skilled</i>		98	106
<i>Unskilled</i>			

- (i) Use the information provided in the tables to find the rates of pay missing from table 1. Insert them in the appropriate places in table 1.

[2]

- (ii) Use the rates of pay from table 1 and the fact that 2011 is the base year to find the price relatives missing from table 2, correct to the nearest whole number. Insert them in the appropriate places in table 2.

[4]

(iii) The table below shows the number of workers at each grade at the company in 2011.

	<i>Number of workers</i>
<i>Skilled</i>	10
<i>Semi-skilled</i>	6
<i>Unskilled</i>	5

By considering the total cost of employing each grade of worker in 2011, show that weights of 9, 5 and 4 should be assigned to the skilled, semi-skilled and unskilled workers respectively. State what assumption you have made in this calculation.

.....
[3]

(iv) When the earnings of the manager for 2011 are included, the ratio of the weights becomes 9 : 5 : 4 : 2 for the workers who are skilled, semi-skilled, unskilled and the manager respectively. The price relative for the manager in 2013, taking 2011 as base, is 108.

Calculate a weighted aggregate cost index for the total wage bill in 2013 with 2011 as base year.

.....[4]

(v) Explain what this figure tells you. State what additional assumption you have made in giving this answer.

.....

[3]

10 The table below shows the number of marriages (in thousands) in a particular country each quarter for a period of 3 years.

Year	Quarter	Number of marriages (thousands)	Four-quarter total	Centred total	Centred moving average value
2011	I	27.9			
	II	65.0			
			233.5		
	III	97.0		465.6	$c = \dots\dots\dots$
			232.1		
	IV	43.6		463.3	57.9
			231.2		
2012	I	26.5		461.3	57.7
			230.1		
	II	64.1		459.8	57.5
			229.7		
	III	$a = \dots\dots\dots$		457.9	57.2
			228.2		
	IV	43.2		455.5	56.9
			227.3		
2013	I	25.0		453.5	56.7
			$b = \dots\dots\dots$		
	II	63.2		451.3	56.4
			225.1		
	III	94.8			
	IV	42.1			

(i) Give a reason for finding moving average values and explain why it is useful to do this.

.....

[2]

(ii) Explain clearly why it is necessary to centre the moving average values in this table.

.....
[2]

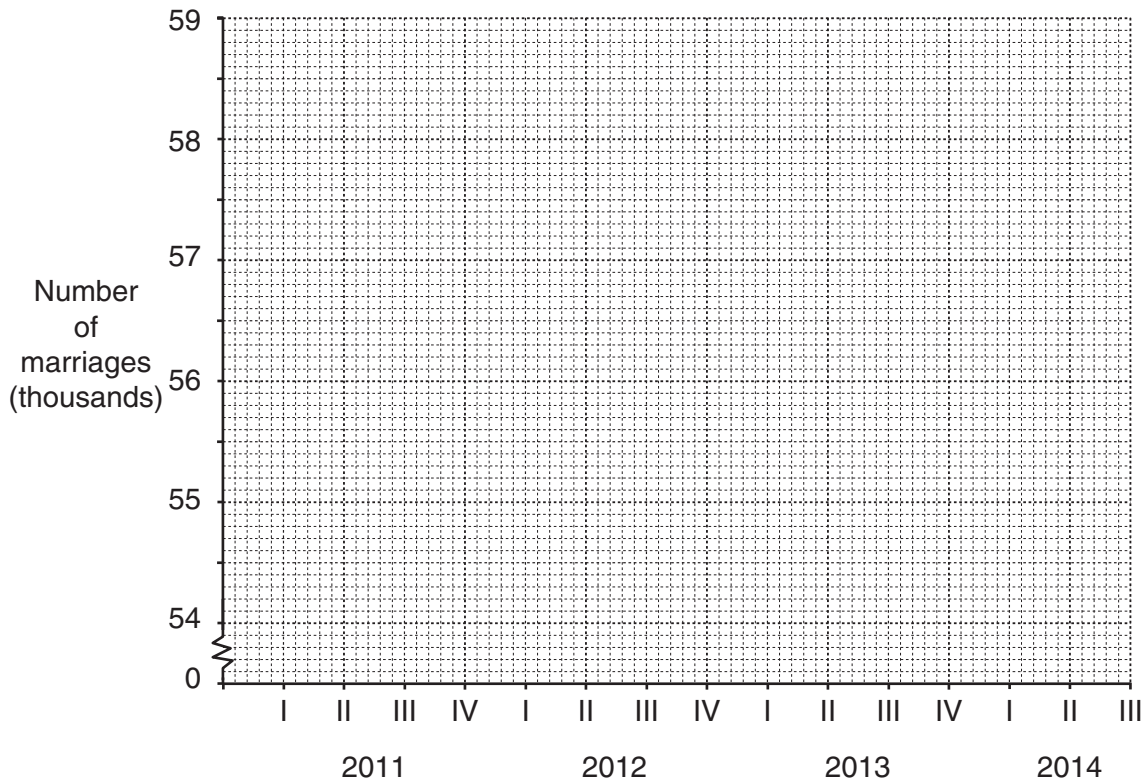
(iii) Calculate the values of a , b and c and insert them in the table.

[3]

(iv) Use the number of marriages and centred moving average values for quarter II of 2012 and 2013 to find an estimate of the seasonal component for quarter II.

.....[3]

(v) Plot all the centred moving average values on the grid below and draw the trend line.



[3]

(vi) Comment on what the trend line tells you.

.....
[1]

(vii) Use your trend line and answer to part (iv) to estimate the number of marriages in quarter II of 2014.

.....[2]

11 (i) The faces of a fair 6-sided die are numbered 1, 1, 1, 2, 2, 3.

(a) A game costs \$2 to play. The die is thrown and a player wins the number of dollars equal to the number on the uppermost face.

Find the expected profit or loss, correct to the nearest cent, for someone playing the game.

.....[5]

(b) In another game two such dice are thrown. Again it costs \$2 to play. A player wins \$ x if they score a double (the same number on the uppermost face of both dice), otherwise they get nothing.

(i) Find, **correct to the nearest dollar**, the value of x which would make this game as fair as possible.

.....[5]

(ii) 90 people are to play this game and the prize for winning is awarded as you have calculated in part (i)(b)(i).

By finding the number of people expected to win, calculate the expected profit or loss for the game owner.

.....[3]

- (ii) Another fair 6-sided die has 4 faces each with one spot and 2 faces each with two spots. The die is thrown once and \$2 is charged to play.

If it is to be a fair game and the prizes are to be proportional to the number of spots on the uppermost face, find how much the prizes should be.

Prize for one spot

Prize for two spots[3]

BLANK PAGE

Permission to reproduce items where third-party owned material protected by copyright is included has been sought and cleared where possible. Every reasonable effort has been made by the publisher (UCLES) to trace copyright holders, but if any items requiring clearance have unwittingly been included, the publisher will be pleased to make amends at the earliest possible opportunity.

To avoid the issue of disclosure of answer-related information to candidates, all copyright acknowledgements are reproduced online in the Cambridge International Examinations Copyright Acknowledgements Booklet. This is produced for each series of examinations and is freely available to download at www.cie.org.uk after the live examination series.

Cambridge International Examinations is part of the Cambridge Assessment Group. Cambridge Assessment is the brand name of University of Cambridge Local Examinations Syndicate (UCLES), which is itself a department of the University of Cambridge.