Cambridge Assessment International Education

Cambridge O Level

CANDIDATE NAME

CENTRE NUMBER | | | | | |

CANDIDATE NUMBER | | | | |

8204869655

STATISTICS **4040/23**

Paper 2 **October/November 2020**

2 hours 15 minutes

You must answer on the question paper.

You will need: Calculator
Pair of compasses
Protractor

INSTRUCTIONS

- Answer **all** questions.
- Use a black or dark blue pen. You may use an HB pencil for any diagrams or graphs.
- Write your name, centre number and candidate number in the boxes at the top of the page.
- Write your answer to each question in the space provided.
- Do **not** use an erasable pen or correction fluid.
- Do **not** write on any bar codes.
- You should use a calculator where appropriate.
- You must show all necessary working clearly.

INFORMATION

- The total mark for this paper is 100.
- The number of marks for each question or part question is shown in brackets [ ].

This document has **16** pages. Blank pages are indicated.

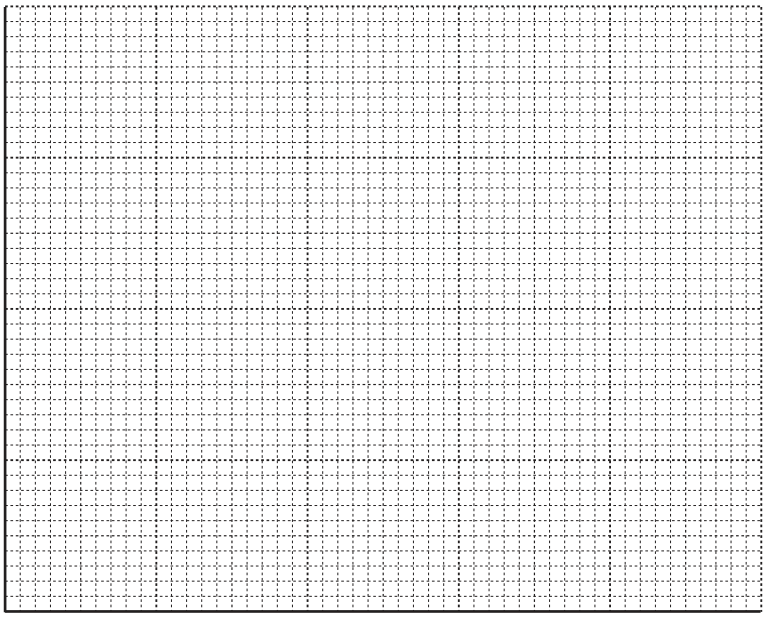
DC (MB) 199886

[Turn over

1 A researcher is interested in changes to the numbers of ospreys, eagles and hawks in Rocky Creek National Park. He finds data about the numbers of these three types of bird seen by visitors to the Park in Oct 2009 and in Oct 2019.

Type of bird	*Number seen in Oct 2009*	*Number seen in Oct 2019*
Osprey	70	60
Eagle	64	47
Hawk	122	192

On the grid below, display the data using a dual bar chart so that, for each type of bird, the number seen in Oct 2009 and the number seen in Oct 2019 can be compared easily.



[3]

2 The following information is collected from all the students in a class.

Gender Height Shoe size

(a) In each case below, use statistical language to describe fully the type of data that is being collected.

Gender .....

Height .....

Shoe size ..... [4]

The students in the class can decide what they want to find out from the data.

Larona wants to find out if taller students generally have bigger feet.

(b) Name the most appropriate statistical diagram that she could use.

..... [1]

Karim wants to find out if boys are generally taller than girls.

(c) Name the two alternatives for the statistical measure that he could use, and give an advantage that one of them has over the other one.

1 ..... 2 .....

.....

..... [3]

3 Five positive whole numbers have a mean of 5.4, a median of 5 and a mode of 8.

Find the five numbers.

..... [3]

4 The recorded distances, in metres, in the Javelin, Shot Put and Discus events in a competition are summarised below.

	Javelin	Shot Put	Discus
Mean	33.07	9.63	30.54
Standard deviation	4.64	1.84	3.72

Onalenna was one of the competitors, and her results, in metres, are shown below.

	Javelin	Shot Put	Discus
Onalenna	31.91	13.31	32.40

Find which of these events was her worst and which was her best, when compared with all the competitors. Justify your answer.

Worst .....

Best ..... [5]

5 A and B are two outcomes of the same experiment, such that

$$P(A) = 0.7 \qquad P(B) = 0.5 \qquad P(A \cup B) = 0.85 .$$

(a) Show whether or not A and B are mutually exclusive events.

[2]

(b) Show whether or not A and B are independent events.

[4]

6 A headteacher is considering making changes to the school uniform. He wishes to consult a sample of the 576 students at his school. He divides the student population by both gender (boys and girls) and age group (lower school and upper school). The number of students in each group is shown in the table below.

	Lower school	Upper school
Boys	192	96
Girls	192	96

He decides to select a sample of **size 6**, stratified by both gender and age.

(a) Find the number of lower school boys that should be in the sample.

..... [1]

Each student is allocated a 3-digit number as shown in the table below.

	Lower school	Upper school
Boys	001–192	385–480
Girls	193–384	481–576

(b) Use the random number table below, starting at the beginning of the table, to select a random sample of **size 6**, stratified by both gender and age. Use every number if the group to which it relates has not yet been fully sampled.

010 739 523 496 010 421 148 260 052 325 256 325 862 069

..... [3]

The deputy headteacher suggests selecting a systematic sample instead.
He selects the following sample of **size 8**:

010, 082, 154, 226, 298, 370, 442, 514

(c) Show whether or not this sample is representative of the lower school boys.

[3]

7 The owner of a publishing company divides her costs into four categories: Wages, Materials, Bills and Other.
The table shows weights, based on expenditure in 2016, and price relatives for 2019, taking 2016 as the base year.

Category	*Weight*	*Price relative in 2019*
Wages	9	.....
Materials	6	107
Bills	2	98
Other	3	104

In 2016 the average weekly wage for a worker at the company was $182.
In 2019 this had risen to $185.64 .

(a) Find the price relative for Wages in 2019, using 2016 as the base year, and insert your value in the table.

[2]

(b) Calculate a weighted aggregate cost index for 2019, taking 2016 as the base year.

..... [3]

(c) Explain what your answer to part **(b)** tells you.

.....

.....

..... [2]

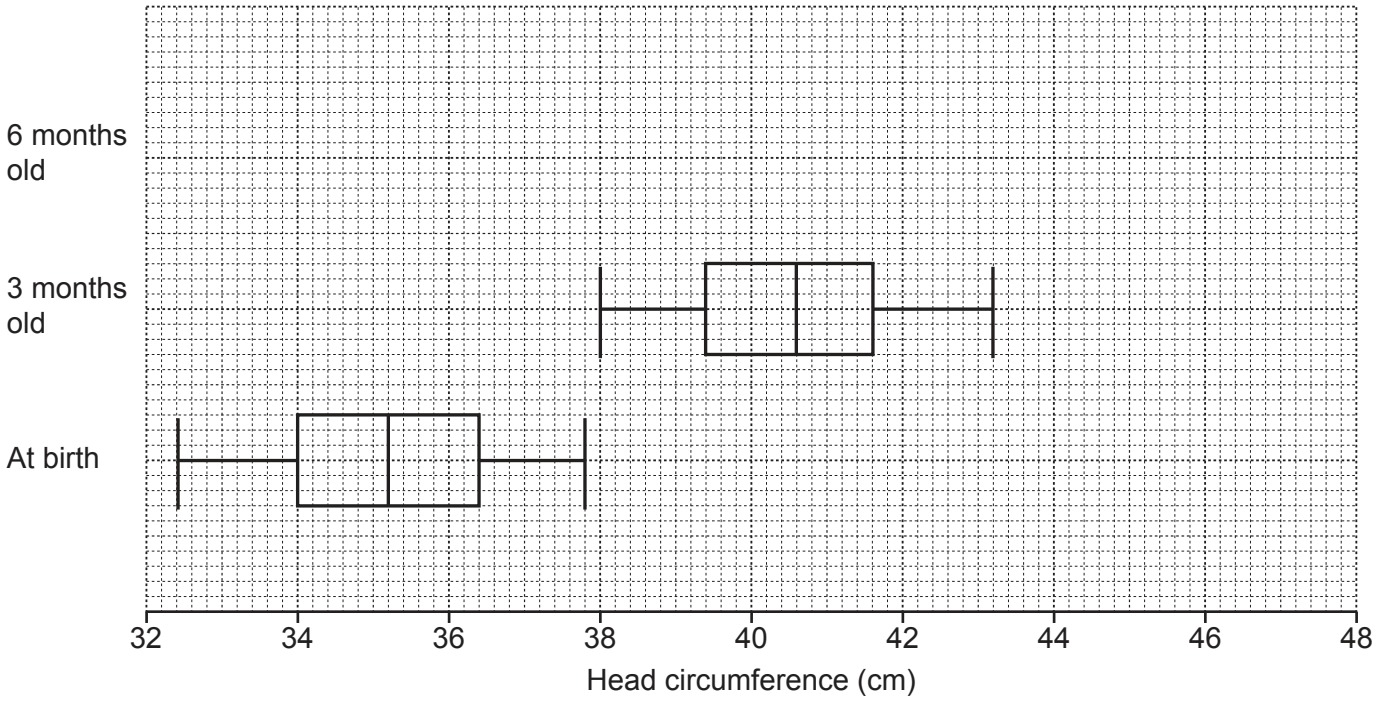
A forecast for expenditure in 2019 calculated using the index found in part **(b)** may not be accurate, if the weights have changed.

(d) Give one possible reason why the weights may have changed.

.....

..... [1]

8 A health worker measured the head circumferences, in cm, of some babies when they were born. She then measured the same babies again at 3 months old and at 6 months old.
Her results for the babies at birth and at 3 months old are summarised below.



(a) (i) Find the range of the head circumferences of the babies at birth.

..... [2]

(ii) Find the interquartile range of the head circumferences of the babies at 3 months old.

..... [2]

(iii) Make two comparisons between the head circumferences of the babies at birth and the head circumferences of the babies at 3 months old.

1 .....

.....

2 .....

.....[2]

The health worker has lost her raw data for the babies at birth. She knows that one particular baby had a head circumference of 39.4 cm at 3 months old.
She says, 'This baby would have had a head circumference of 34.0 cm when it was born'.

(b) Explain whether or not you think she is correct.

.....

..... [1]

The head circumferences for these babies at 6 months old are shown in the stem-and-leaf diagram below.

41	0	3	9	
42	1	6	7	9
43	0	5	8	9
44	7	7	9	
45	0	4	8	
46	0	2		

Key: 42 | 1 represents a head circumference of 42.1 cm

(c) Find the median, lower quartile, upper quartile and interquartile range of the head circumferences of the 6-month-old babies.

Median .....

Lower quartile .....

Upper quartile .....

Interquartile range ..... [4]

(d) Complete the diagram on page 8 by adding a box-and-whisker diagram for the head circumferences of the 6-month-old babies. [2]

One of the 6-month-old babies is to be selected at random for further measurements.

(e) Find the probability that this baby will have

(i) a head circumference that, when rounded to the nearest centimetre, is 42 cm,

..... [1]

(ii) a head circumference greater than 45.5 cm, given that the baby has a head circumference greater than 44.0 cm.

..... [1]

(f) The head circumference of each baby is expected to increase by 2.5 cm between 6 months old and 1 year old.

Use this information and your answer to part **(c)** to find estimates for the median and interquartile range of the head circumferences of these babies at 1 year old.

Median .....

Interquartile range ..... [1]

9 **(a)** A game consists of rolling an unbiased die and tossing an unbiased coin.
The die has six sides, labelled \$1, \$2, \$3, \$4, \$5 and \$6.
The coin has two sides, labelled 'win' and 'lose'.
If the coin lands on 'win' the player wins a prize equal to the amount shown on the die.
If the coin lands on 'lose' the player wins nothing.

(i) Complete the probability distribution of the prizes in dollars (\$).

Prize (\$)	0	1	2	3	4	5	6
Probability							

[3]

A charge is to be made for playing this game.

(ii) Find the amount that should be charged if it is to be a fair game.

..... [2]

(b) In another game, the probability distribution of the prizes is as shown below.

Prize ($)	0	1	2	3	4
Probability	$\frac{1}{4}$	$\frac{1}{2}$	$\frac{1}{8}$	p	p

(i) Given that the probability of a prize of $3 is equal to the probability of a prize of $4, find the value of p.

..... [2]

(ii) Tumelo plays this game twice.

Find the probability that he wins a total of $2 in prizes.

..... [3]

The owner of this game charges $1 each time someone plays it.

(iii) Find the expected profit or loss for the owner, if 20 people each play this game once.

..... [4]

10 A farmer finds the mass, in grams, of each of the 120 mangoes she picks one day. Some information about her results is shown below.

Mass, m (g)	*Frequency*	*Cumulative frequency*
$150 \leqslant m < 190$	7	
$190 \leqslant m < 225$	13	
$225 \leqslant m < 250$	15	
$250 \leqslant m < 270$	30	
$270 \leqslant m < 300$	27	
$300 \leqslant m < 330$	20	
$330 \leqslant m < 375$	8	

(a) Complete the cumulative frequency column in the table above. [1]

Of the mangoes she picks, the 10% with the smallest masses will be sold immediately to make chutney.

(b) Use linear interpolation to calculate an estimate of the mass of the largest of these mangoes that will be sold to make chutney.
Give your answer correct to 3 significant figures.

..... [4]

The remaining mangoes are to be classed as either small, medium or large, as defined in the table below.

Size	*Mass*
Small	less than 240 g
Medium	from 240 g to under 290 g
Large	290 g or more

(c) Use linear interpolation to find estimates for the numbers of the remaining mangoes that are classed as small, medium and large.

Small .....

Medium .....

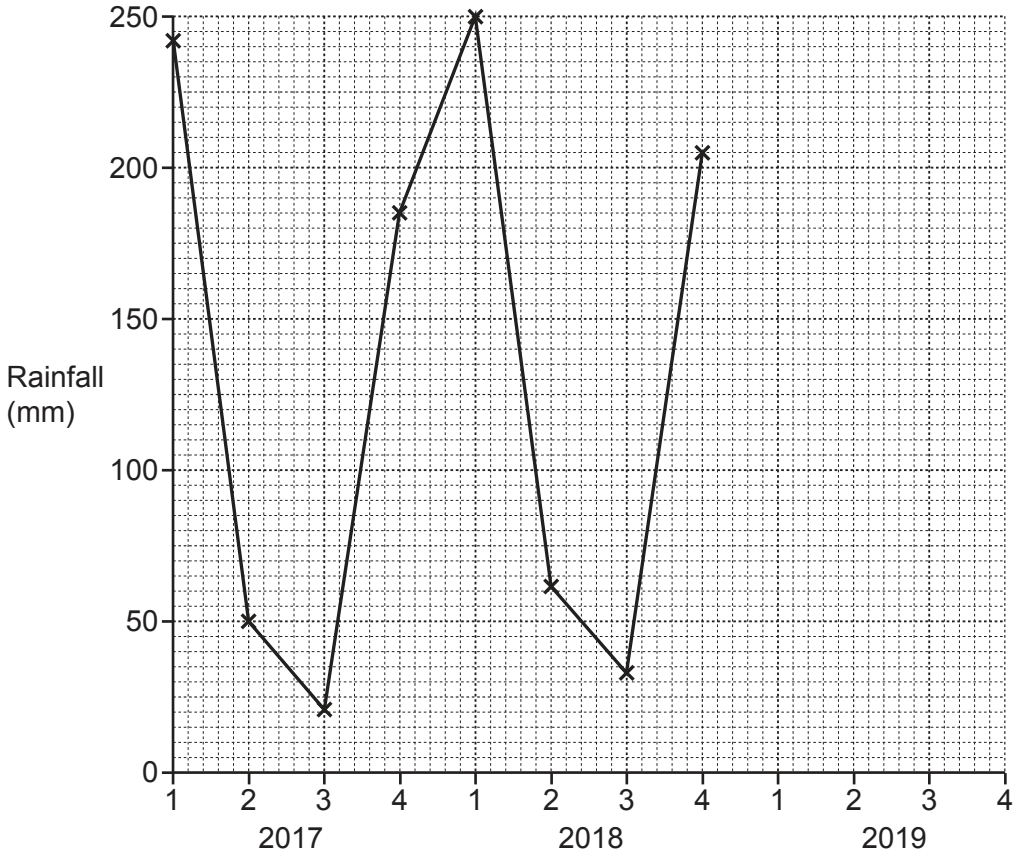
Large ..... [5]

4 of these remaining mangoes are selected at random and put into a bag.

(d) Use your answers to part **(c)** to find an estimate for the probability that the bag will contain 3 large mangoes and 1 small mango.

..... [4]

11 The quarterly rainfall totals, in mm, were recorded in a town from quarter 1 (Q1) of 2017 to quarter 4 (Q4) of 2018. They are shown in the time series graph and in the table below.



	Q1	Q2	Q3	Q4
2017	242	50	21	185
2018	250	62	33	205

(a) Describe the seasonal variation in the rainfall for this town.

.....

..... [1]

(b) State two purposes of finding moving average values.

1 .....

2 ..... [2]

In this situation, 4-point moving average values should be found.

(c) Explain why it will be necessary to centre the moving average values.

.....

..... [1]

(d) Calculate the set of centred 4-point moving average values.
Present your results using the table below.

Year and Quarter	*Rainfall* (mm)	
2017 Q1	242	
2017 Q2	50	
2017 Q3	21	
2017 Q4	185	
2018 Q1	250	
2018 Q2	62	
2018 Q3	33	
2018 Q4	205	

[5]

(e) Plot the centred moving average values on the grid containing the time series graph, on page 14, and draw an appropriate trend line. [3]

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

.....

..... [1]

The seasonal components for the first three quarters are shown in the table below.

Quarter	1	2	3
Seasonal component	120	–74	–106

(g) Use this table to find the seasonal component for quarter 4.

..... [1]

(h) Use your trend line and answer to part **(g)** to estimate the rainfall total for quarter 4 of 2019.

..... [2]

BLANK PAGE