OCR
Oxford Cambridge and RSA

Wednesday 22 May 2019 – Morning

Level 3 Certificate Quantitative Problem Solving (MEI)

H867/02 Statistical Problem Solving

Time allowed: 2 hours

You must have:
- the Insert (inserted)
- the Statistical Tables (ST1) (inserted)

You may use:
- a scientific or graphical calculator

Please write clearly in black ink. **Do not write in the barcodes.**

Centre number | | | | | | Candidate number | | | | |

First name(s) _____

Last name _____

INSTRUCTIONS
- The Insert will be found inside this document.
- Use black ink. You may use an HB pencil for graphs and diagrams.
- Answer **all** the questions.
- Write your answer to each question in the space provided. If additional space is required, you should use the lined page(s) at the end of this booklet. The question number(s) must be clearly shown.
- You are advised that an answer may receive **no marks** unless you show sufficient detail of the working to indicate that a correct method is being used.

INFORMATION
- The total mark for this paper is **60**.
- The marks for each question are shown in brackets **[ ]**.
- This document consists of **20** pages.
- Final answers should be given to a degree of accuracy appropriate to the context.

Answer **all** the questions.

Section A (30 marks)

1 A town has a rail users' group. They want to investigate what improvements to the present service people in the town would most value and produce the survey in Fig. 1.1 below.

Five of the group's members agree to interview people for the survey. They are given the instructions in Fig. 1.2.

The results are summarised in Table 1.3.

Rail improvement survey

Please look at the following list of possible improvements to the railway system and ring the **two** that you think are the most important.

A Cheaper fares
B Greater reliability
C More frequent trains
D Cleaner trains
E Less overcrowding on the trains
F Better information on the stations
G A simpler ticket structure.

For office use only

Category of respondent .....

Fig. 1.1

Instructions

Select people in the following categories
5 adult women
5 adult men
10 students. (Do not include any mature students.)

Give them the list of possible improvements and ask them to ring the **two** that they consider most important.

Collect their responses and fill in the category of each person.

Fig. 1.2

Improvement	A	B	C	D	E	F	G
Adult men	23	9	2	2	10	1	1
Adult women	20	2	0	17	9	0	4
Students	50	15	10	1	14	1	9
Total	**93**	**26**	**12**	**20**	**33**	**2**	**14**

Table 1.3

(i) Which one of the following terms best describes their sample?

Opportunity, simple random, stratified, quota, cluster, self-selected. **[1]**

(ii) Make three statements describing different things that the rail users' group can learn from the survey. **[3]**

(iii) One of the interviewers did not carry out the instructions properly, as they did not interview the correct number of people.

What mistake did the interviewer make? **[1]**

1 (i)	
1 (ii)	**Statement 1**
	Statement 2
	Statement 3
1 (iii)	

2 A local authority is responsible for a long coastline including 15 beaches that are used by holiday makers. The authority is planning to clean up the beaches before the start of the holiday season. The total length of the beaches is 45 km. The authority wants to estimate the weight of rubbish that they will have to take away and dispose of.

They choose 4 of the beaches and on each of them select, at random, 10 stretches of length 20 metres. They collect the rubbish from each of these stretches and weigh it.

(i) Give one reason why an estimate based on this sample may not be very accurate. **[1]**

2 (i)	

The histogram in Fig. 2.1 shows the distribution of the weights from the stretches in the sample.

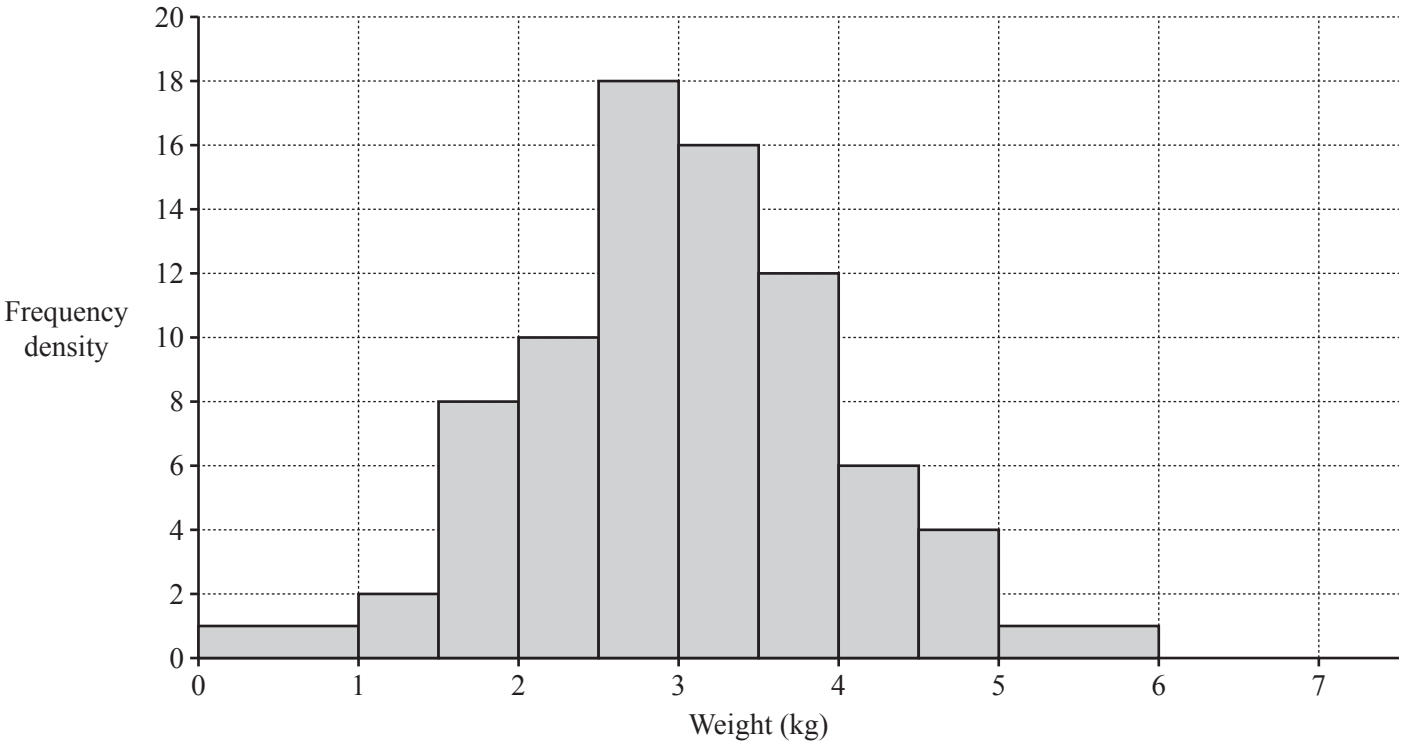

Fig. 2.1

(ii) Complete Table 2.2 in the answer space giving the frequencies in the different groups. **[2]**

2 (ii)

Weight, w kg	$0 \leqslant w < 1$	$1 \leqslant w < 1.5$	$1.5 \leqslant w < 2$	$2 \leqslant w < 2.5$	$2.5 \leqslant w < 3$
Frequency	**1**			**5**	**9**

$3 \leqslant w < 3.5$	$3.5 \leqslant w < 4$	$4 \leqslant w < 4.5$	$4.5 \leqslant w < 5$	$5 \leqslant w < 6$	**Total**
					40

Table 2.2

(iii) **(*A*)** Identify one feature of the distribution shown in Fig. 2.1 that suggests that the Normal distribution would be a good model. **[1]**

(*B*) State an appropriate value for the mean of the Normal model. **[1]**

(*C*) It is suggested that a value of 1.0 is appropriate for the standard deviation of the Normal distribution. Show that this value is consistent with the frequency in the sample data for $1 \leqslant w < 5$. **[3]**

2 (iii) **(*A*)**

(*B*)

(*C*)

2 cont

(iv) Using the value for the mean you gave in part **(iii)(*B*)**, estimate the total weight of rubbish that the authority will have to remove from the 15 beaches used by holiday makers, giving your answer in tonnes. **[3]**

The authority is planning a campaign called "Beachclean". During one day local volunteers will collect all the rubbish in sacks and leave them in piles, one per beach, ready to be taken away.

(v) Given that each stretch takes one person about a quarter of an hour, make a rough estimate of the number of volunteers who would be needed, explaining your reasoning. **[2]**

2 (iv)

2 (v)

BLANK PAGE

PLEASE DO NOT WRITE ON THIS PAGE

3 A new type of mouse is being found in woodlands in the UK. It is believed to have come from Asia. A naturalist fears that it may endanger two native species: the wood mouse and the yellow-necked mouse.

A friend expresses the hope that the three species may be able to live alongside each other. He suggests that this would be indicated if the proportions of the three different types of mice in various places are about the same.

They trap mice in three different woods, A, B and C, and record how many of each type they get. The results are given in Table 3.1. They then carry out a χ^2 test on the proportions.

Observed frequency, f_o	**New mouse**	**Wood mouse**	**Yellow-necked**	**Total**
Wood A	16	30	14	**60**
Wood B	16	28	16	**60**
Wood C	20	10	10	**40**
Total	**52**	**68**	**40**	**160**

Table 3.1

(i) State the null and alternative hypotheses for their test. **[1]**

(ii) Complete the tables in the answer space.

Show that, at the 5% significance level, the result is not significant. **[7]**

3 (i)	

3 (ii)

Expected frequency, f_e	New mouse	Wood mouse	Yellow-necked	Total
Wood A	19.5	25.5		**60**
Wood B				**60**
Wood C	13			**40**
Total	**52**	**68**	**40**	**160**

Table 3.2

Contributions	New mouse	Wood mouse	Yellow-necked
Wood A	$\frac{(16-19.5)^2}{19.5} = 0.6282$	0.7941	
Wood B	0.6282	0.2451	
Wood C	3.7692	2.8824	

Table 3.3

3 cont The naturalist then says "This proves that there is nothing to worry about. The test shows conclusively that the new mice pose no threat to the native species."

(iii) Give two criticisms of the naturalist's statement. **[2]**

3 (iii)	**Criticism 1**
	Criticism 2

Unknown to the naturalist and his friend, a group of scientists have been monitoring the situation in a different wood for 5 years. Each year they have trapped a sample of mice, using exactly the same methods. Their results are given in Table 3.4.

Observed frequency, f_o	**New mouse**	**Wood mouse**	**Yellow-necked**	**Total**
Year 1	2	27	31	**60**
Year 2	5	22	27	**54**
Year 3	12	23	21	**56**
Year 4	19	21	12	**52**
Year 5	27	26	8	**61**

Table 3.4

(iv) State two conjectures that the scientists might make from their data. **[2]**

3 (iv)	**Conjecture 1**
	Conjecture 2

BLANK PAGE

PLEASE DO NOT WRITE ON THIS PAGE

Section B (30 marks)

The questions in this section are based on the pre-release data. A hard copy of this is provided with this examination paper.

4 (a) (i) The land area of China is 9 326 410 km^2.

Show that, to 3 significant figures, its population density is 145 people per km^2. **[2]**

(ii) To 3 significant figures, the population density of India is 416 people per km^2.

Estimate the land area of India. **[2]**

(b) (i) Show that, to 3 significant figures, the total GDP for Malaysia is 5.26×10^{11} US$. **[2]**

(ii) Find the mean GDP per capita of the combined population of Malaysia and Singapore. **[4]**

4 (a)(i)	
4 (a)(ii)	

4 (b)(i)	
4 (b)(ii)	

5 Solveig is a student in Norway. She is interested in the relationship between GDP per capita and life expectancy. She carries out a pilot investigation using the Nordic countries as a sample. She carries out a Spearman's rank correlation test on her data, using a 5% significance level.

(i) Complete the missing cells in Table 5.1.

State suitable null and alternative hypotheses, carry out the test and state the conclusion. **[7]**

5 (i)

Country	GDP per capita	GDP rank, x	LE	LE rank, y	$d = x - y$	d^2
Denmark	37 800	4	79.09	6	−2	4
Estonia	22 400	8	74.07	8	0	0
Faroe Islands	30 500	6	80.11	4		
Finland	35 900	5	79.69	5		
Iceland	40 700		81.22			
Latvia	19 100		73.44			
Lithuania	22 600	7	75.98	7		
Norway	55 400	1	81.60	2	−1	1
Sweden	40 900	2	81.89	1	1	1
				Total		

Table 5.1

Solveig draws the scatter diagram for all the countries in the world. It is shown in Fig. 5.2.

(ii) Circle the point on the scatter diagram corresponding to Equatorial Guinea. **[1]**

(iii) Make two comments on what the scatter diagram shows. **[2]**

5 (ii)

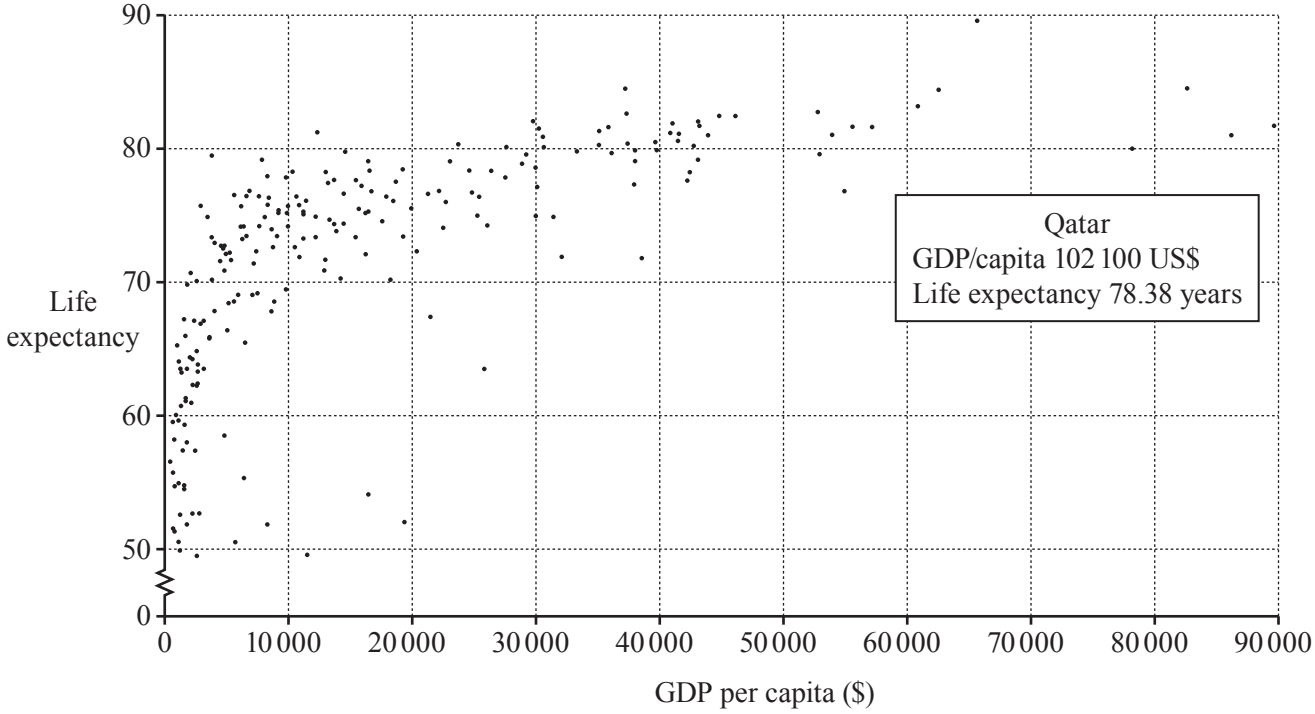

Fig. 5.2

5 (iii)	**Comment 1**
	Comment 2

6 This question is about modelling the size of the world population. Its rapid growth in recent years is a cause of concern over the future of the planet.

Table 6.1 gives the years in which the population is estimated to have reached whole numbers of billions. It also gives the lengths of the intervals involved, their mid-points and average year-on-year percentage increases over the intervals.

Size (billions)	1	2	3	4	5	6	7
Year	1804	1927	1960	1974	1987	1999	2011
Interval (years)	123	33	14	13	12	12	
Mid-point of interval	1865.5	1943.5	1967	1980.5	1993	2005	
% increase per year	0.565	1.236	2.076	1.731	1.531	1.293	

Table 6.1

(i) Using your calculator, show that, to 3 significant figures, the value of 1.00565^{123} is 2.00.

Explain briefly how this relates to the information in Table 6.1. **[2]**

6 (i)

(ii) The last row of Table 6.1 shows that the percentage rates of increase are going down. Fig. 6.2 shows the four most recent values from Table 6.1. (They are plotted at the mid-points of their intervals.)

By drawing a suitable line on this graph, estimate the year when the world population may stop increasing.

State one reason why this estimate should be treated with caution. **[3]**

6 (ii)

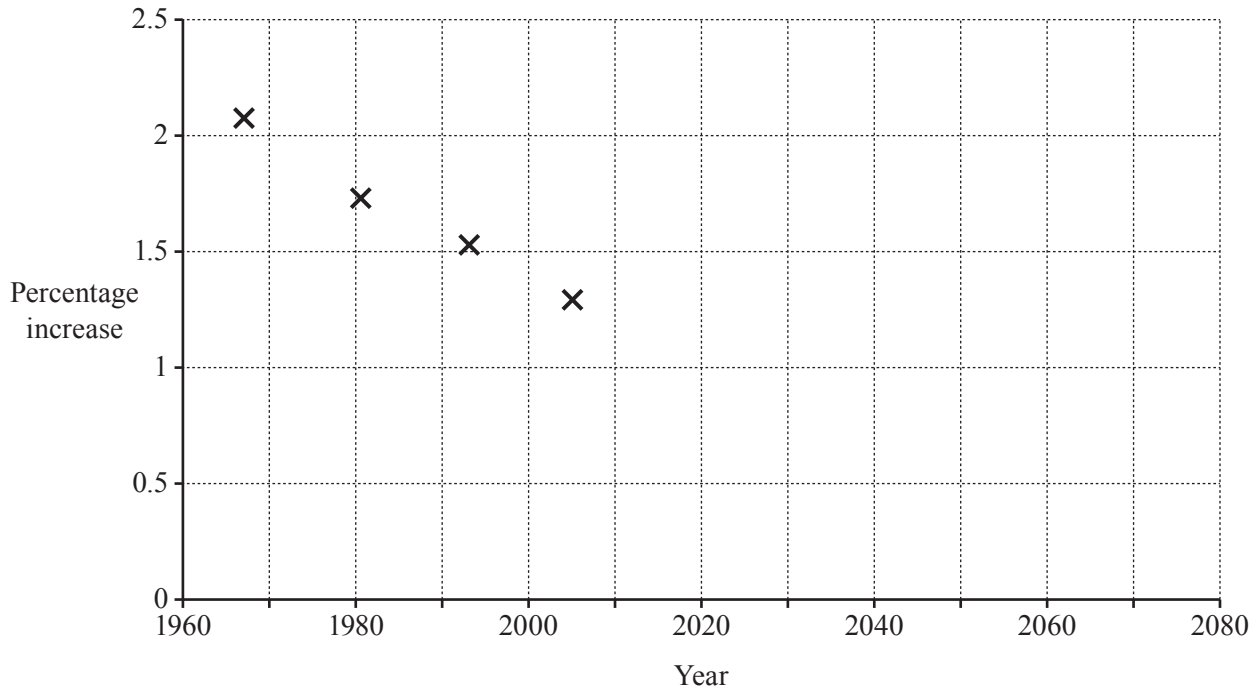

Fig. 6.2

Year

Reason

6 cont The next two parts of this question involve a birth rate of 12.5. This is the theoretical birth rate of a country with a stable population and a life expectancy of 80 years.

The frequency chart in Fig. 6.3 shows the distribution of the birth rates of all the countries in the world.

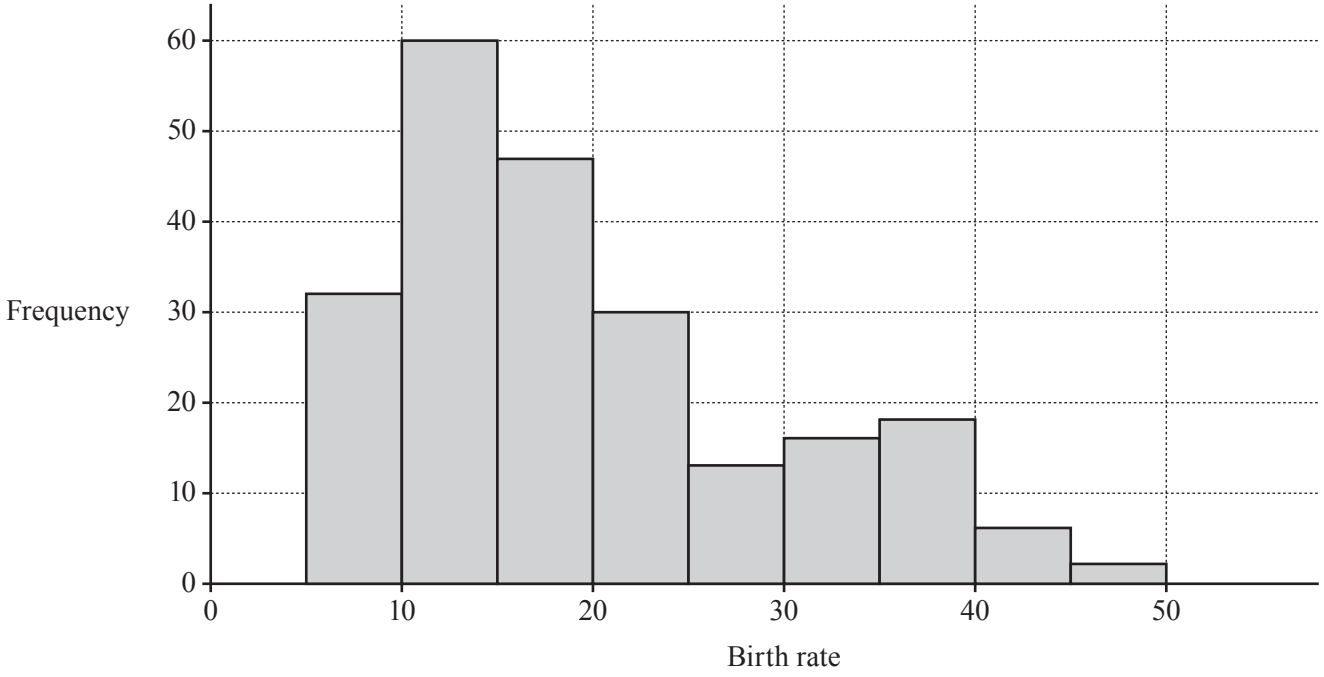

Fig. 6.3

(iii) Use Fig. 6.3 to estimate the number of countries with birth rates less than 12.5. **[2]**

(iv) The pre-release data includes the birth rates of 32 countries in Western Europe. For how many of these countries is the birth rate more than 12.5? **[1]**

Table 6.4 gives the birth rates for the UK for the 100 years from 1901 to 2001. (Data were not collected in 1941 when the country was at war). The overall pattern is typical of countries in Europe and many other parts of the world.

Year	1901	1911	1921	1931	1951	1961	1971	1981	1991	2001
Birth rate	28.5	23.2	18.7	17.2	16.7	18.2	13.2	13.4	12.7	12.2

Table 6.4

(v) Give two comments on your answer to part **(ii)** considering the data in Table 6.4 and the rest of the question. **[2]**

6 (iii)	
6 (iv)	
6 (v)	**Comment 1**
	Comment 2

END OF QUESTION PAPER

ADDITIONAL ANSWER SPACE

If additional space is required, you should use the following lined page(s). The question number(s) must be clearly shown in the margin(s).