OCR
Oxford Cambridge and RSA

# Thursday 25 May 2023 – Afternoon

## Level 3 Certificate Core Maths B (MEI)

**H869/02** Statistical Problem Solving

**Time allowed: 2 hours**

**You must have:**
- the Insert (inside this document)
- the Statistical Tables (ST1) (inside this document)

**You can use:**
- a scientific or graphical calculator

*H869/02*

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

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

First name(s) _____

Last name _____

**INSTRUCTIONS**
- Use black ink. You can use an HB pencil, but only for graphs and diagrams.
- Write your answer to each question in the space provided. If you need extra space use the lined pages at the end of this booklet. The question numbers must be clearly shown.
- Answer **all** the questions.
- Where appropriate, your answer should be supported with working.
- Give your final answers to a degree of accuracy that is appropriate to the context.

**INFORMATION**
- The total mark for this paper is **60**.
- The marks for each question are shown in brackets **[ ]**.
- This document has **20** pages.

**ADVICE**
- Read each question carefully before you start your answer.

**BLANK PAGE**

**PLEASE DO NOT WRITE ON THIS PAGE**

## Section A (30 marks)

**1** The article below was written by a journalist and appeared in the local newspaper of a small town. The town is on a main railway line and has a station.

### More parking needed

If you want to travel by train you'll most likely have to think again. There are just 35 parking spaces at our station and on most days you won't be able to find an empty one to park your car in.

Many readers have told me that parking is a problem. So I decided to see for myself and went to the station at 9 am on each of the ten working days in a fortnight. Every day I checked whether each of the 35 parking spaces was occupied or free. So overall I carried out 350 checks. Here are the numbers of parking spaces I found free on the different days.

**Free parking spaces**

| | **Mo** | **Tu** | **We** | **Th** | **Fr** |
|---|---|---|---|---|---|
| **Week 1** | 0 | 2 | 0 | 0 | 5 |
| **Week 2** | 1 | 0 | 0 | 0 | 6 |

So the readers are right. My checks showed that, at the times I was there, over 95% of the parking spaces were occupied. Most days there was nowhere to park at all. While I was at the station several cars drove round the car park looking for a space that wasn't there and so they continued their journey by road.

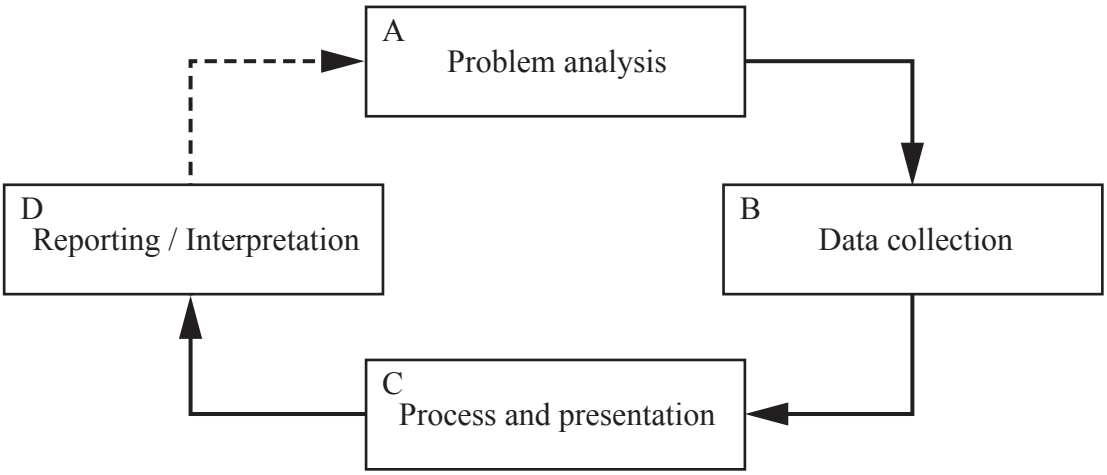
It is time for those responsible to wake up. The car park needs to be enlarged.

**(a)** Show that overall for over 95% of the checks that the journalist made the parking spaces were occupied. **[2]**

| | |
|---|---|
| **1(a)** | |
| | |
| | |
| | |
| | |

The diagram below shows the four stages of the Statistical Problem Solving Cycle.

The piece of work described in the journalist's article fits this cycle.




**(b)** The following sentences are taken from the journalist's article. In each case say which of the stages, A, B, C or D, the work described fits into. **[4]**

| | |
|---|---|
| **1(b)(i)** | over 95% of the parking spaces were occupied.<br><br>**Stage** |
| **(ii)** | The car park needs to be enlarged.<br><br>**Stage** |
| **(iii)** | Every day I checked whether each of the 35 parking spaces was occupied or free.<br><br>**Stage** |
| **(iv)** | Many readers have told me that parking is a problem.<br><br>**Stage** |

**(c)** A member of the railway company's staff said "The article is not fair because the times sampled were not random". Comment critically on this statement. **[1]**

| | |
|---|---|
| **1(c)** | |
| | |
| | |
| | |
| | |
| | |
| | |
| | |

**2** Jamila is a sociology student. She is carrying out a project on the reactions of men, women and young people to global warming.

She draws up the short questionnaire in **Fig. 2.1**.

**Fig. 2.1**

**1.** Please choose the **one** of the following actions you are most likely to take to reduce global warming. Circle A, B, C or D.

**A** I should stop eating meat

**B** I should spend at least one day a month planting trees

**C** For long journeys I should travel by train or bus and for short ones I should walk or cycle

**D** The climate will look after itself and so I don't need to do anything special

**2.** Describe yourself by ringing one of the following.

A young person (under 18) An adult woman An adult man

Thank you.

Jamila asks her classmates and friends to complete the questionnaire. She also asks them to get their parents and their friends to do it.

The results are summarised in **Fig. 2.2**.

**Fig. 2.2**

| **Observed frequency, $f_o$** | **Young people** | **Adult women** | **Adult men** | **Total** |
|---|---|---|---|---|
| **Statement A** | 25 | 12 | 12 | **49** |
| **Statement B** | 10 | 13 | 12 | **35** |
| **Statement C** | 13 | 23 | 23 | **59** |
| **Statement D** | 0 | 4 | 13 | **17** |
| **Total** | **48** | **52** | **60** | **160** |

Jamila uses the data to carry out a $\chi^2$ hypothesis test at the 1% significance level.

**(a)** Which one of these is the null hypothesis for Jamila's test? **[1]**
Circle **one** of the options in the answer space.

**P** Young people are better informed about global warming than adults.

**Q** The proportions of people holding particular opinions about global warming are independent of whether they are young or adult (men or women).

**R** Men and women hold the same views about global warming.

**S** All of the courses of action stated are equally important.

**(b)** Complete the table of Expected frequencies given as **Fig. 2.3** in the answer space. **[2]**

**(c)** Complete the calculation of $X^2$ in **Fig. 2.4** in the answer space. **[2]**

**2(a)** **P** **Q** **R** **S**

**2(b)** **Fig. 2.3**

| **Expected frequency, $f_e$** | **Young people** | **Adult women** | **Adult men** | **Total** |
|---|---|---|---|---|
| **Statement A** | 14.7 | 15.925 | 18.375 | **49** |
| **Statement B** | 10.5 | 11.375 | 13.125 | **35** |
| **Statement C** | 17.7 | 19.175 | 22.125 | **59** |
| **Statement D** | | | | |
| **Total** | | | | |

**2(c)** **Fig. 2.4**

$X^2$ = ..... + 0.967 ... + 2.211 ...

\+ 0.023 ... + 0.232 ... + 0.096 ...

\+ 1.248 ... + 0.763 ... + 0.034 ...

\+ 5.100 ... + 0.420 ... + 6.884 ...

$X^2$ = ..... (to 1 d.p.)

**(d)** Show that the test is significant at the 1% level. **[4]**

**(e)** Jamila is very excited by the result. She says

"The overall result and the contributions to $X^2$ prove that young people everywhere have the best ideas and adult men's views can't be trusted".

Give **two** different criticisms of Jamila's statement. **[2]**

| | |
|---|---|
| **2(d)** | |
| | |
| | |
| | |
| | |
| | |
| | |
| | |
| | |
| | |
| **2(e)** | **Criticism 1** |
| | |
| | |
| | |
| | |
| | |
| | **Criticism 2** |
| | |
| | |
| | |
| | |
| | |

**BLANK PAGE**

**PLEASE DO NOT WRITE ON THIS PAGE**

**3** **Fig. 3** shows a Normal distribution with mean 176 and standard deviation 7.

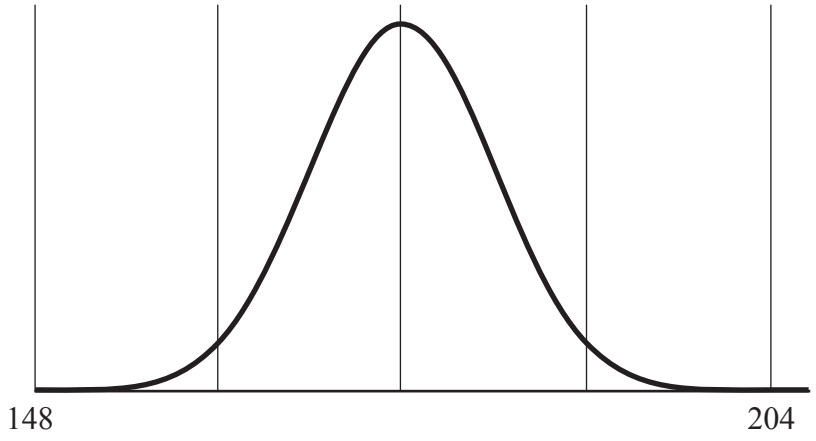
**(a)** **(i)** Show that the $z$-values of 204 and 148 are $+4$ and $-4$. **[2]**

**(ii)** Mark the horizontal scale in integer values of $z$ from $-4$ to $+4$. **[1]**

**(b)** Find the area of the region between $z = -1$ and $z = +2$. **[3]**

**3(a)(i)**

**(ii)** **Fig. 3**




**3(b)**

The graph in **Fig. 3** shows the distribution of the heights in centimetres of adult men in the USA.

**(c)** A recent survey in Chicago showed that 1 054 753 adult men lived there.
Estimate, to the nearest 100, how many of them were at least 1.9 metres tall. **[3]**

**(d)** The mean height of adult males in Peru is 5 feet $4\frac{1}{2}$ inches.

In which country, the USA or Peru, is the mean height of adult men greater, and by how much?

[1 metre is 3 feet 3.4 inches and 1 foot is 12 inches.] **[3]**

**3(c)**

**3(d)**

## 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** Riley says "I expect that the larger a country, the greater will be the length of its roads".

Taylor disagrees. "Many large countries have open spaces where nothing much happens".

They decide to carry out a test on a sample of countries. They choose one country at random from each region in the pre-release data set. (See **Fig. 4**.)

Then they calculate Spearman's rank correlation coefficient and use it to carry out a hypothesis test at the 5% significance level on Riley's theory.

**(a)** Complete the table in the answer space and calculate $r_s$. **[4]**

**(b)** Write down the null and alternative hypotheses for the test. **[1]**

**(c)** Complete the test and say whether the outcome supports Riley's theory. **[3]**

**(d)** Give one reason why the result should be treated with caution. **[1]**

**4(a)** **Fig. 4**

| Country | Area ($km^2$) | Area rank, $x$ | Roads (km) | Roads rank, $y$ | $d = x - y$ | $d^2$ |
|---|---|---|---|---|---|---|
| **Tunisia** | 163 610 | 5 | 19 418 | 6 | −1 | 1 |
| **Botswana** | 581 730 | | 17 916 | | | |
| **Panama** | 75 420 | 6 | 15 137 | 8 | −2 | 4 |
| **Canada** | 9 984 670 | | 1 042 300 | | | |
| **Paraguay** | 406 752 | 3 | 32 059 | 4 | −1 | 1 |
| **Laos** | 236 800 | 4 | 39 586 | 3 | 1 | 1 |
| **Jamaica** | 10 991 | 8 | 22 121 | 5 | 3 | 9 |
| **Denmark** | 43 094 | 7 | 74 497 | 2 | 5 | 25 |
| **Samoa** | 2 831 | 9 | 2 337 | 9 | 0 | 0 |
| | | | | $\sum$ | 0 | |

$r_s$ = .....

| | |
|---|---|
| 4(b) | **Null hypothesis, $H_0$:** |
| | |
| | **Alternative hypothesis, $H_1$:** |
| | |
| 4(c) | |
| | |
| | |
| | |
| | |
| | |
| | |
| | |
| | |
| 4(d) | |
| | |
| | |
| | |

The pre-release data cover 219 countries for which the values of land area and total road length are given. Riley uses the spreadsheet to find the product moment correlation coefficient for these countries. It is 0.6332.

**(e)** What should Riley and Taylor conclude from this figure? **[1]**

| | |
|---|---|
| 4(e) | |
| | |
| | |
| | |
| | |
| | |
| | |
| | |

**5** **(a)** Write down the population and the birth rate of Ethiopia.

How many babies are born in Ethiopia in a typical year? **[2]**

**(b)** On average, is a baby born every second in Ethiopia? **[2]**

| | | |
|---|---|---|
| **5(a)** | **Ethiopia Population** | **Birth rate** |
| | **Babies in a typical year** | |
| | | |
| | | |
| | | |
| | | |
| | | |
| | | |
| **5(b)** | | |
| | | |
| | | |
| | | |
| | | |
| | | |
| | | |
| | | |
| | | |

Jack calculates the birth rate across the world. He works with the spreadsheet containing the pre-release data.

He starts by deleting the ten countries for which data on population and birth rate are not both available.

Then he enters =(C2/1000)*J2 into cell P2 and copies P2 down to P237.

He enters =SUM(C2:C237) into C238 and gets 7 403 176 022.

He enters =SUM(P2:P237) into P238 and gets 135 877 586.

Finally he enters =P238/(C238/1000) into P239 and gets 18.35 (to 2 decimal places).

**(c)** Name one of the countries that Jack deletes. **[1]**

**(d)** State what the numbers in C238, P238 and P239 represent. **[3]**

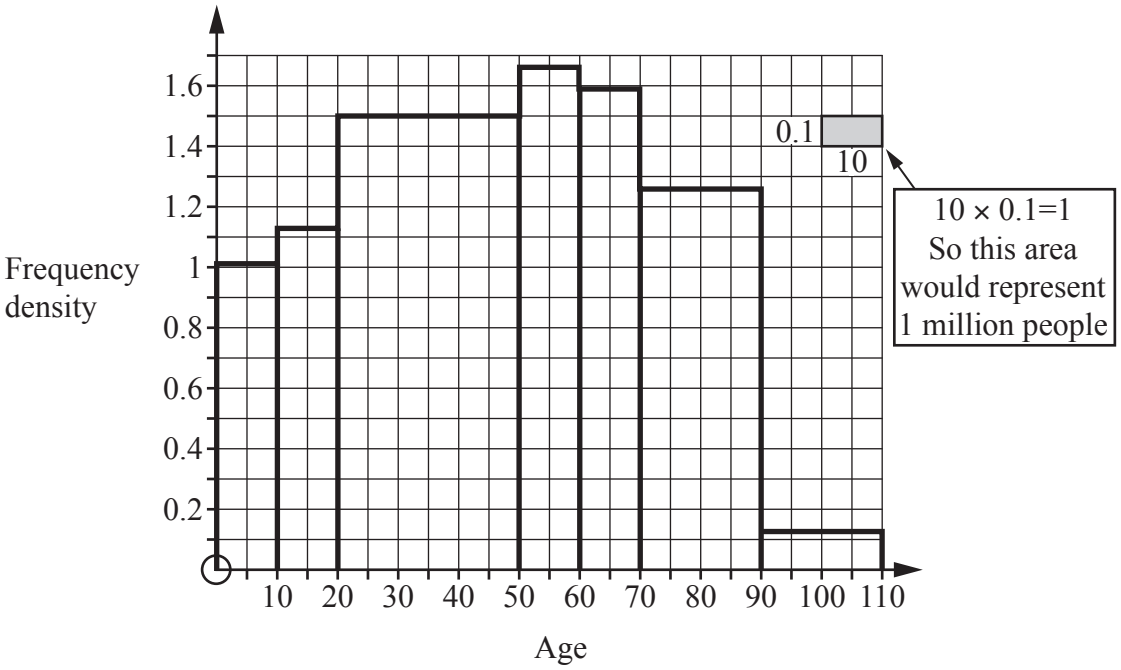
**(e)** On average, how many babies are born around the world every second? **[1]**

| | |
|---|---|
| **5(c)** | |
| | |
| **5(d)** | **C238** |
| | |
| | |
| | **P238** |
| | |
| | |
| | **P239** |
| | |
| | |
| **5(e)** | |
| | |
| | |
| | |
| | |
| | |
| | |
| | |
| | |

**6** **(a)** Write down the total population and the median age of Japan. **[1]**

The histogram in **Fig. 6** below shows the distribution of the population of Japan in seven age groups at a recent time.

**Fig. 6**




The heights of the bars, from left to right, are 1.01, 1.13, 1.50, 1.66, 1.59, 1.26 and 0.12. The units for frequency density are millions of people per year of age.

**(b)** Show that the number of people under 20 is approximately 21.4 million. **[1]**

**(c)** Estimate the number of people over 65. **[3]**

| | | |
|---|---|---|
| **6(a)** | **Japan** | **Population** |
| | | **Median age** |

| | |
|---|---|
| **6(b)** | |
| **6(c)** | |

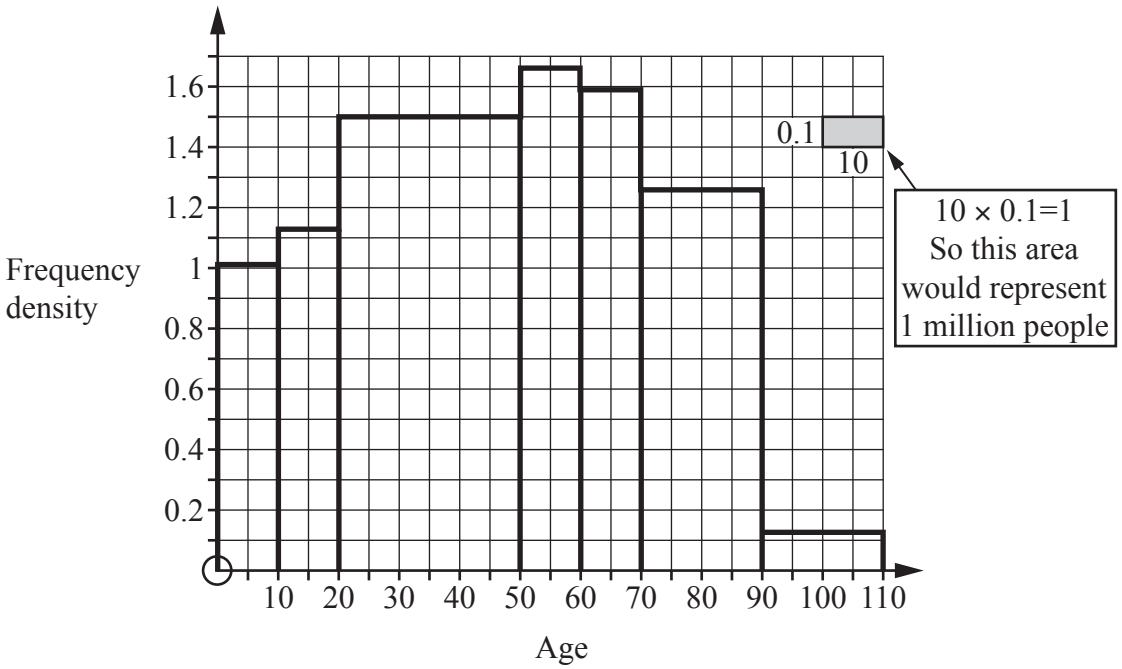
Planners describe the population by dividing it into three categories.

| Under 20 | 20 to 65 | Over 65 |
|---|---|---|
| In education | Working | Retired |

**(d)** Use the information in the histogram to calculate estimates for the numbers of people in each of these categories. **[1]**

| | |
|---|---|
| **6(d)** | **In education** |
| | **Working** |
| | **Retired** |

**Fig. 6 (repeated)**




The heights of the bars, from left to right, are 1.01, 1.13, 1.50, 1.66, 1.59, 1.26 and 0.12. The units for frequency density are millions of people per year of age.

In a planning model everyone from 20 to 65 is assumed to be working, but no one else. Those in work provide financial support for people who are in education or who have retired. The support may be as family members and/or via the government through taxes.

**(e)** Use the model to estimate the mean number of people that a person in work is supporting financially, excluding himself or herself. **[2]**

In 2021 the Japanese parliament voted to raise the retirement age from 65 to 70.

**(f)** Using the numbers illustrated in the histogram, show that, if the retirement age is 70, the mean number of people supported by a working person is 0.63.

Comment on your answer. **[2]**

**(g)** The Netherlands are also planning to raise their retirement age. Suggest, with a brief explanation, one data value in the pre-release data for the Netherlands that could provide a reason for this. **[1]**

| | |
|---|---|
| 6(e) | |
| | |
| | |
| | |
| | |
| | |
| | |
| | |
| | |
| 6(f) | |
| | |
| | |
| | |
| | |
| | |
| | **Comment** |
| | |
| | |
| | |
| 6(g) | **The Netherlands** |
| | |
| | |
| | |

**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).