OCR
Oxford Cambridge and RSA

# Wednesday 14 June 2017 – Morning

## A2 GCE MATHEMATICS (MEI)

**4767/01** Statistics 2

**QUESTION PAPER**



Candidates answer on the Printed Answer Book.

**OCR supplied materials:**
- Printed Answer Book 4767/01
- MEI Examination Formulae and Tables (MF2)

**Other materials required:**
- Scientific or graphical calculator

**Duration:** 1 hour 30 minutes

## INSTRUCTIONS TO CANDIDATES

These instructions are the same on the Printed Answer Book and the Question Paper.

- The Question Paper will be found inside the Printed Answer Book.
- Write your name, centre number and candidate number in the spaces provided on the Printed Answer Book. Please write clearly and in capital letters.
- **Write your answer to each question in the space provided in the Printed Answer Book**. 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.
- Use black ink. HB pencil may be used for graphs and diagrams only.
- Read each question carefully. Make sure you know what you have to do before starting your answer.
- Answer **all** the questions.
- Do **not** write in the barcodes.
- You are permitted to use a scientific or graphical calculator in this paper.
- Final answers should be given to a degree of accuracy appropriate to the context.

## INFORMATION FOR CANDIDATES

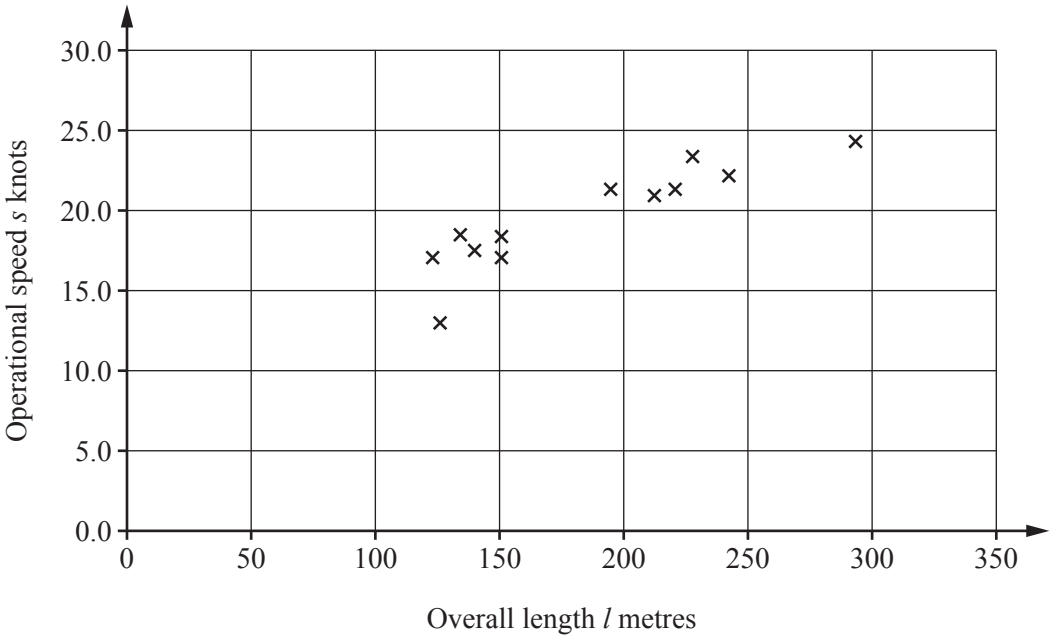
This information is the same on the Printed Answer Book and the Question Paper.

- The number of marks is given in brackets **[ ]** at the end of each question or part question on the Question Paper.
- 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.
- The total number of marks for this paper is **72**.
- The Printed Answer Book consists of **12** pages. The Question Paper consists of **8** pages. Any blank pages are indicated.

## INSTRUCTION TO EXAMS OFFICER/INVIGILATOR

- Do not send this Question Paper for marking; it should be retained in the centre or recycled. Please contact OCR Copyright should you wish to re-use this document.

**1** The scatter diagram below illustrates the overall lengths $l$ metres and the typical operational speeds $s$ knots (nautical miles per hour) of 12 container ships. The length of a ship is one of the factors which determines its typical operational speed.




Summary statistics for these data are as follows:

$n = 12, \quad \Sigma l = 2219, \quad \Sigma s = 234.6, \quad \Sigma l^2 = 443\,867, \quad \Sigma s^2 = 4700.56, \quad \Sigma ls = 45\,149.0.$

**(i)** State which of the two variables $l$ and $s$ is the independent variable and which is the dependent variable. Briefly explain your answers. **[2]**

**(ii)** Calculate the equation of the regression line of $s$ on $l$. **[5]**

**(iii)** Interpret the coefficient of $l$ in terms of the relationship between overall length and speed in the equation of the regression line found in part **(ii)**. **[2]**

**(iv)** Calculate the value of the residual for the data point where $l = 126$ and $s = 13.0$. **[3]**

**(v)** Use the equation of the regression line in part **(ii)** to calculate an estimate of the operational speed of a ship with overall length of 100 metres. Comment on the reliability of this estimate. **[2]**

If the data point in part **(iv)** is removed from the data set, the equation of the new regression line is

$$s = 0.0453l + 11.5.$$

**(vi)** Recalculate the estimate for an overall length of 100 metres using this new equation. Discuss which of these two estimates you think is likely to be more reliable. **[3]**

**2** At a small hospital maternity department, there is an average of 1.3 births per day. Throughout this question, you should assume that births occur independently, at random times, and at a uniform average rate.

**(i)** Briefly explain the meaning of each of the three terms 'independently', 'random' and 'uniform', in the context of births at the maternity department. **[3]**

$X$ represents the number of births at the hospital on a randomly chosen day.

**(ii)** State the distribution of $X$ and also the variance of $X$. **[2]**

**(iii)** Find $P(X > 3)$. **[2]**

**(iv)** Find the probability that there are exactly 3 births in a period of 3 days. **[2]**

There is an average of 0.4 home births per **week** in the area served by the hospital.

**(v)** Find the probability that the total number of births in a week (at home and in hospital) is at least 10. **[2]**

**(vi)** Use a suitable approximating distribution to find the probability that there is a total of at least 50 births in a period of 4 weeks. **[5]**

**(vii)** How realistic do you think the assumption of independence made at the start of this question is? **[2]**

**3** The random variable $X$ represents the weight, in grams, of a particular type of chocolate bar. It is known that $X$ is Normally distributed with mean 50.7 and variance 0.72. On the wrapper it states that the bar weighs 50 grams.

**(i)** Find the proportion of these chocolate bars that actually weigh at least 50 grams. **[3]**

**(ii)** A quality control manager wishes to increase this proportion to 95%.

($A$) Find the required value of the mean if the variance remains unchanged. **[4]**

($B$) Find the required value of the variance if the mean remains unchanged. **[2]**

The weights of another type of chocolate bar are also Normally distributed. On the wrapper it states that the bar weighs 25 grams. It is known that 99% of these bars weigh at least 25.0 grams and 75% of them weigh at least 25.4 grams.

**(iii)** Find the probability that one of these bars weighs at least 26.0 grams. **[6]**

**(iv)** One bar of the first type (with the original mean and standard deviation) and 2 bars of the second type are selected at random. Find the probability that at least one of the bars has a weight less than that stated on its wrapper. **[2]**

**4** **(a)** In an investigation into dietary supplements, a random sample of 200 adults was selected. Each of them was asked whether or not they regularly take dietary supplements. The 200 adults were categorised as 'Male under 50', 'Male 50 or older', 'Female under 50', 'Female 50 or older'. Their answers to the question are summarised in the table below.

| | Yes | No |
|---|---|---|
| Male under 50 | 13 | 33 |
| Male 50 or older | 18 | 31 |
| Female under 50 | 24 | 37 |
| Female 50 or older | 24 | 20 |

**(i)** Write down null and alternative hypotheses for a test to examine whether there is any association between category of adult and whether or not they regularly take dietary supplements. **[1]**

The expected frequencies under the null hypothesis for the usual $\chi^2$ test are shown in the table below.

| Expected frequency | Yes | No |
|---|---|---|
| Male under 50 | 18.17 | 27.83 |
| Male 50 or older | 19.36 | 29.65 |
| Female under 50 | 24.10 | 36.91 |
| Female 50 or older | 17.38 | 26.62 |

**(ii)** Verify the expected frequency of 18.17 for Male under 50 answering Yes. **[2]**

The contributions to the test statistic are shown in the table below.

| Contribution | Yes | No |
|---|---|---|
| Male under 50 | 1.4710 | 0.9604 |
| Male 50 or older | 0.0949 | 0.0619 |
| Female under 50 | 0.0004 | 0.0002 |
| Female 50 or older | 2.5215 | 1.6463 |

**(iii)** Verify the contribution to the test statistic of 1.4710 for Male under 50 answering Yes. **[2]**

**(iv)** Given that the total of the contributions is 6.757, correct to 3 decimal places, complete the test at the 10% significance level. **[4]**

**(v)** For each category of adult, comment briefly on how their taking of dietary supplements compares with what would be expected if there were no association. **[3]**

**(b)** The breaking strength of a particular type of rope is known to be Normally distributed with mean 562 kg and a standard deviation of 27.4 kg when the rope is dry. A researcher believes that the breaking strength may be different when the rope is wet. He selects a random sample of 12 pieces of wet rope and finds that the mean of their breaking strengths is 547 kg. Carry out a test at the 10% significance level to investigate the researcher's belief. You may assume that the breaking strengths of wet ropes of this type are still Normally distributed with standard deviation 27.4 kg. **[8]**

**END OF QUESTION PAPER**

**BLANK PAGE**

**BLANK PAGE**

**BLANK PAGE**