Oxford Cambridge and RSA

Friday 22 October 2021 – Afternoon

AS Level Further Mathematics B (MEI)

Y416/01 Statistics b

Time allowed: 1 hour 15 minutes

You must have:
- the Printed Answer Booklet
- the Formulae Booklet for Further Mathematics B (MEI)
- a scientific or graphical calculator

8311773518

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 in the **Printed Answer Booklet**. If you need extra space use the lined pages at the end of the Printed Answer Booklet. The question numbers must be clearly shown.
- Fill in the boxes on the front of the Printed Answer Booklet.
- Answer **all** the questions.
- Where appropriate, your answer should be supported with working. Marks might be given for using a correct method, even if your answer is wrong.
- Give your final answers to a degree of accuracy that is appropriate to the context.
- Do **not** send this Question Paper for marking. Keep it in the centre or recycle it.

INFORMATION
- The total mark for this paper is **60**.
- The marks for each question are shown in brackets **[ ]**.
- This document has **8** pages.

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

Answer **all** the questions.

1 Over time LED light bulbs gradually lose brightness. For a particular type of LED bulb, it is known that the mean reduction in brightness after 10 000 hours is 2.6%. A manufacturer produces a new version of this bulb, which costs less to make, but is claimed to have the same reduction in brightness after 10 000 hours as the previous version.

In order to check this claim, a random sample of 10 bulbs is selected. For each bulb, the original brightness and the brightness after 10 000 hours are measured, in suitable units. A spreadsheet is used to produce a 95% confidence interval for the mean percentage reduction in brightness. A screenshot of the spreadsheet is shown in Fig. 1.

	A	B	C	D	E	F	G	H	I	J	K
1	**Original brightness**	1075	1121	1106	1095	1101	1109	1114	1123	1108	1115
2	**After 10 000 hours**	1042	1084	1076	1065	1070	1079	1081	1091	1080	1082
3	**Percentage reduction**	3.07	3.30	2.71	2.74	2.82	2.71	2.96	2.85	2.53	2.96
4											
5											
6	Sample mean (%)	2.8650									
7	Sample sd (%)	0.2179									
8	SE	0.0689									
9	DF	9									
10	*t* value	2.262									
11	Lower limit	2.709									
12	Upper limit	3.021									
13											

Fig. 1

(a) State the confidence interval in the form $a < \mu < b$. **[1]**

(b) Explain whether the confidence interval suggests that the mean percentage reduction in brightness after 10 000 hours is different from 2.6%. **[2]**

(c) Explain how the value in cell B8 was calculated. **[2]**

(d) State an assumption necessary for this confidence interval to be calculated. **[2]**

(e) Explain the advantage of using the same bulbs for both measurements. **[2]**

2 Natasha is investigating binomial distributions. She constructs the spreadsheet in Fig. 2 which shows the first 3 and last 4 rows of a simulation involving two independent variables, X and Y, with distributions B(10, 0.3) and B(50, 0.3) respectively. The spreadsheet also shows the corresponding value of the random variable Z, defined by $Z = 5X - Y$, for each pair of values of X and Y.

There are 100 simulated values of each of X, Y and Z. The spreadsheet also shows whether each value of Z is greater than 6, and cells D103 and D104 show the number of values of Z which are greater than 6 and not greater than 6 respectively.

	A	B	C	D	E
1	***X***	***Y***	***Z* = 5*X* − *Y***	***Z* > 6**	
2	4	13	7	Y	
3	4	17	3	N	
4	3	21	−6	N	
5					
6					
98	1	14	−9	N	
99	5	12	13	Y	
100	3	18	−3	N	
101	3	15	0	N	
102					
103			Number of Y	19	
104			Number of N	81	
105					

Fig. 2

(a) Use the information in the spreadsheet to write down an estimate of $P(Z > 6)$. **[1]**

(b) Explain how a more reliable estimate of $P(Z > 6)$ could be obtained. **[1]**

(c) **(i)** State the greatest possible value of Z. **[1]**

(ii) Explain why it is very unlikely that Z would have this value. **[1]**

(d) Use the Central Limit Theorem to calculate an estimate of the probability that the mean of 20 independent values of Z is greater than 2. **[6]**

3 The weights in kg of male otters in a large river system are known to be Normally distributed with mean 8.3 and standard deviation 1.8. A researcher believes that weights of male otters in another river are higher because of what he suspects is better availability of food. The researcher records the weights of a random sample of 9 male otters in this other river. The sum of these 9 weights is 83.79 kg.

(a) **In this question you must show detailed reasoning.**

You should assume that:

- the weights of otters in the other river are Normally distributed,
- the standard deviation of the weights of otters in the other river is also 1.8 kg.

Show that a test at the 5% significance level provides sufficient evidence to conclude that the mean weight of male otters in the other river is greater than 8.3 kg. **[7]**

(b) Explain whether the result of the test suggests that the weights are higher due to better availability of food. **[2]**

(c) If the standard deviation of the weights of otters in the other river could not be assumed to be 1.8 kg, name an alternative test that the researcher could carry out to investigate otter weights. **[1]**

(d) Explain why, even if a test at the 5% significance level results in the rejection of the null hypothesis, you cannot be sure that the alternative hypothesis is true. **[2]**

4 John regularly downloads podcasts onto his mobile phone. From past experience he knows that the average time to download one 30-minute podcast is 12.7 s. He believes that this time has recently increased. At each of 12 randomly chosen times, he downloads a 30-minute podcast. The times in seconds to download the 12 podcasts are as follows.

12.63 13.24 11.73 14.91 13.17 13.53 12.33 14.27 11.48 13.51 13.05 13.83

(a) Given that it is not known whether the times are Normally distributed, carry out a suitable test at the 5% significance level to investigate whether the average download time has increased. **[8]**

(b) What assumption is required to carry out the test in part **(a)**? **[1]**

5 A food company makes mini apple pies. The weight of pastry in a pie is Normally distributed with mean 75 g and standard deviation 4 g. The weight of filling in a pie is Normally distributed with mean 130 g and standard deviation 8 g. You should assume that the weights of pastry and filling in a pie are independent.

(a) Find the probability that the weight of pastry in a randomly chosen pie is between 70 g and 80 g. **[1]**

(b) Find the probability that the mean weight of filling in 10 randomly chosen pies is at least 125 g. **[3]**

The pies are sold in packs of 4. The weight of the packaging is Normally distributed with mean 165 g and standard deviation 6 g.

(c) In order to find the probability that the total weight of a pack of 4 pies is less than 1 kg, you must assume that the weight of the packaging is independent of the weight of the pies.

(i) State another necessary assumption. **[1]**

(ii) Given that the assumptions are valid, calculate this probability. **[4]**

6 The probability density function of the continuous random variable X is given by

$$f(x) = \begin{cases} 2(1+ax) & 0 \leqslant x \leqslant 1, \\ 0 & \text{otherwise,} \end{cases}$$

where a is a constant.

(a) Show that $a = -1$. **[2]**

(b) Find the cumulative distribution function of X. **[3]**

(c) Find $P(X < 0.5)$. **[2]**

(d) Show that $E(X)$ is greater than the median of X. **[4]**

END OF QUESTION PAPER

BLANK PAGE

BLANK PAGE