Oxford Cambridge and RSA

Thursday 14 October 2021 – Afternoon

AS Level Further Mathematics B (MEI)

Y412/01 Statistics a

Time allowed: 1 hour 15 minutes

904833442

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

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 The random variable X represents the clutch size (the number of eggs laid) by female birds of a particular species. The probability distribution of X is given in the table.

r	2	3	4	5	6	7
$P(X = r)$	0.03	0.07	0.27	0.49	0.13	0.01

(a) Find each of the following.

- $E(X)$
- $Var(X)$ **[2]**

On average 65% of eggs laid result in a young bird successfully leaving the nest.

(b) **(i)** Find the mean number of young birds that successfully leave the nest. **[1]**

(ii) Find the standard deviation of the number of young birds that successfully leave the nest. **[1]**

2 A football player is practising taking penalties. On each attempt the player has a 70% chance of scoring a goal. The random variable X represents the number of attempts that it takes for the player to score a goal.

(a) Determine $P(X = 4)$. **[2]**

(b) Find each of the following.

- $E(X)$
- $Var(X)$ **[2]**

(c) Determine the probability that the player needs exactly 4 attempts to score 2 goals. **[2]**

(d) The player has n attempts to score a goal.

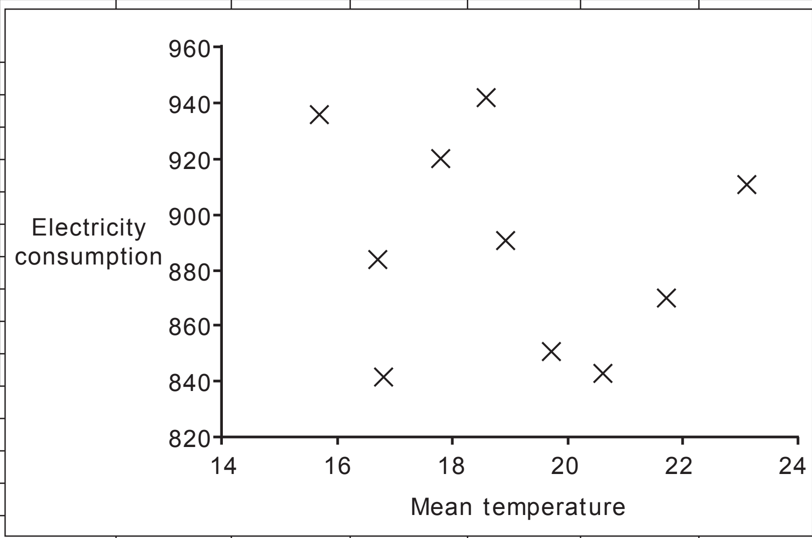
(i) Determine the least value of n for which the probability that the player first scores a goal on the nth attempt is less than 0.001. **[2]**

(ii) Determine the least value of n for which the probability that the player scores at least one goal in n attempts is at least 0.999. **[2]**

3 A student is investigating the link between temperature (in degrees Celsius) and electricity consumption (in Gigawatt-hours) in the country in which he lives.

The student has read that there is strong negative correlation between daily mean temperature over the whole country and daily electricity consumption during a year. He wonders if this applies to an individual season. He therefore obtains data on the mean temperature and electricity consumption on ten randomly selected days in the summer. The spreadsheet output below shows the data, together with a scatter diagram to illustrate the data.

	A	B
1	Mean temperature	Electricity consumption
2	15.7	936
3	19.7	851
4	21.7	870
5	18.6	942
6	16.7	884
7	23.1	911
8	16.8	842
9	18.9	891
10	20.6	843
11	17.8	920



(a) Calculate Pearson's product moment correlation coefficient between daily mean temperature and daily electricity consumption. **[1]**

The student decides to carry out a hypothesis test to investigate whether there is negative correlation between daily mean temperature and daily electricity consumption during the summer.

(b) Explain why the student decides to carry out a test based on Pearson's product moment correlation coefficient. **[2]**

(c) Show that the test at the 5% significance level does not result in the null hypothesis being rejected. **[4]**

(d) The student concludes that there is no correlation between the variables in the summer months.

Comment on the student's conclusion. **[2]**

4 It is known that in an electronic circuit, the number of electrons passing per nanosecond can be modelled by a Poisson distribution. In a particular electronic circuit, the mean number of electrons passing per nanosecond is 12.

(a) **(i)** Determine the probability that there are more than 15 electrons passing in a randomly selected nanosecond. **[2]**

(ii) Determine the probability that there are fewer than 50 electrons passing in a randomly selected period of 5 nanoseconds. **[2]**

(b) Explain what you can deduce about the electrons passing in the circuit from the fact that a Poisson distribution is a suitable model. **[2]**

5 A fair spinner has five faces, labelled 0, 1, 2, 3, 4.

(a) State the distribution of the score when the spinner is spun once. **[2]**

(b) Determine the probability that, when the spinner is spun twice, one of the scores is less than 2 and the other is at least 2. **[2]**

(c) Find the variance of the total score when the spinner is spun 5 times. **[3]**

6 A health researcher is investigating the relationship between age and maximum heart rate.

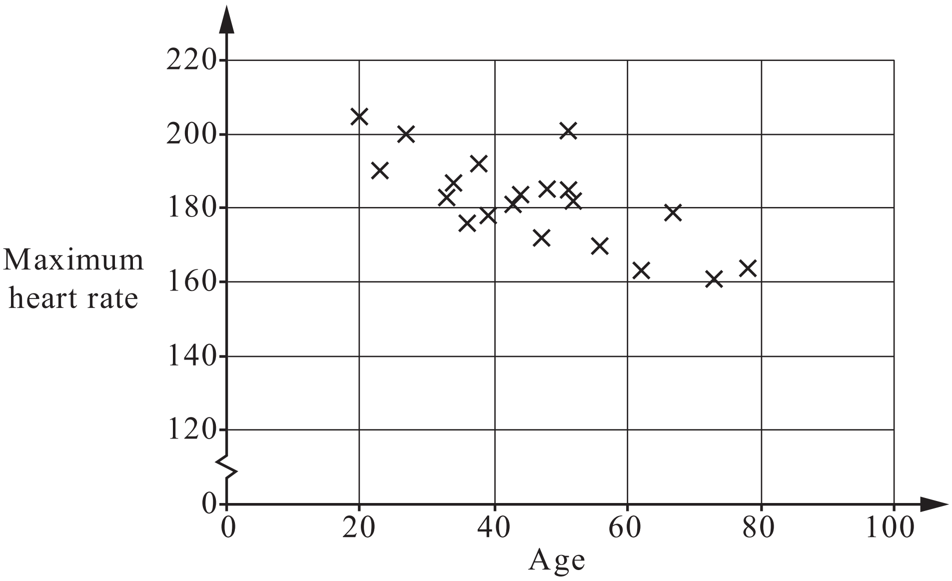
A commonly quoted formula states that 'maximum heart rate = 220 − age in years'. The researcher wants to check if this formula is a satisfactory model for people who work in the large hospital where she is employed. The researcher selects a random sample of 20 people who work in her hospital, and measures their maximum heart rates.

(a) Explain why the researcher selects a sample, rather than using all of the people who work in the hospital. **[1]**

The ages, x years, and maximum heart rates, y beats per minute, of the people in the researcher's sample are summarised as follows.

$n = 20 \quad \sum x = 922 \quad \sum y = 3638 \quad \sum x^2 = 47\,250 \quad \sum y^2 = 664\,610 \quad \sum xy = 164\,998$

These data are illustrated below.



(b) (i) Draw the line which represents the formula 'maximum heart rate = 220 − age in years' on the copy of the scatter diagram in the Printed Answer Booklet. **[1]**

(ii) Comment on how well this model fits the data. **[1]**

(c) Determine the equation of the regression line of maximum heart rate on age. **[4]**

(d) Use the equation of the regression line to predict the values of the maximum heart rate for each of the following ages.

- 40 years
- 5 years **[2]**

(e) Comment on the reliability of your predictions in part **(d)**. **[2]**

7 A biologist is investigating migrating butterflies. **Fig. 7.1** shows the numbers of migrating butterflies passing her location in 100 randomly chosen one-minute periods.

Number of butterflies	0	1	2	3	4	5	6	7	$\geqslant 8$
Frequency	6	9	18	26	13	16	9	3	0

Fig. 7.1

(a) **(i)** Use the data to show that a suitable estimate for the mean number of butterflies passing her location per minute is 3.3. **[1]**

(ii) Explain how the value of the variance estimate calculated from the sample supports the suggestion that a Poisson distribution may be a suitable model for these data. **[2]**

The biologist decides to carry out a test to investigate whether a Poisson distribution may be a suitable model for these data.

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

Complete the copy of **Fig. 7.2** of expected frequencies and contributions for a chi-squared test in the Printed Answer Booklet. **[4]**

Number of butterflies	Frequency	Probability	Expected frequency	Chi-squared contribution
0	6	0.0369	3.6883	1.4489
1	9	0.1217	12.1714	0.8264
2	18			0.2160
3	26			0.6916
4	13	0.1823	18.2252	1.4981
5	16	0.1203	12.0286	
6	9	0.0662	6.6158	0.8593
$\geqslant 7$	3	0.0510	5.0966	0.8625

Fig. 7.2

(c) Complete the chi-squared test at the 5% significance level. **[6]**

END OF QUESTION PAPER

BLANK PAGE

