OCR
Oxford Cambridge and RSA

Thursday 6 June 2019 – Afternoon

A Level Further Mathematics B (MEI)

Y432/01 Statistics Minor

Time allowed: 1 hour 15 minutes

7769553021

You must have:
- Printed Answer Booklet
- Formulae Further Mathematics B (MEI)

You may use:
- a scientific or graphical calculator

INSTRUCTIONS
- Use black ink. HB pencil may be used for graphs and diagrams only.
- Answer **all** the questions.
- **Write your answer to each question in the space provided in the Printed Answer Booklet.** If additional space is required, you should use the lined page(s) at the end of the Printed Answer Booklet. The question number(s) must be clearly shown.
- 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
- The total mark for this paper is **60**.
- The marks for each question are shown in brackets **[ ]**.
- 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 used. You should communicate your method with correct reasoning.
- The Printed Answer Booklet consists of **12** pages. The Question Paper consists of **8** pages.

Answer **all** the questions.

1 In a game at a charity fair, a spinner is spun 4 times.
On each spin the chance that the spinner lands on a score of 5 is 0.2.
The random variable X represents the number of spins on which the spinner lands on a score of 5.

(a) Find $P(X = 3)$. **[2]**

(b) Find each of the following.

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

One game costs £1 to play and, for each spin that lands on a score of 5, the player receives 50 pence.

(c) **(i)** Find the expected total amount of money gained by a player in one game. **[2]**

(ii) Find the standard deviation of the total amount of money gained by a player in one game. **[1]**

2 A market researcher wants to interview people who watched a particular television programme. Audience research data used by the broadcaster indicates that 12% of the adult population watched this programme. This figure is used to model the situation.
The researcher asks people in a shopping centre, one at a time, if they watched the programme. You should assume that these people form a random sample of the adult population.

(a) Find the probability that the fifth person the researcher asks is the first to have watched the programme. **[2]**

(b) Find the probability that the researcher has to ask at least 10 people in order to find one who watched the programme. **[1]**

(c) Find the probability that the twentieth person the researcher asks is the third to have watched the programme. **[3]**

(d) Find how many people the researcher would have to ask to ensure that there is a probability of at least 0.95 that at least one of them watched the programme. **[3]**

3 A company has been commissioned to make 50 very expensive titanium components.
A sample of the components needs to be tested to ensure that they are sufficiently strong. However, this is a test to destruction, so the components which are tested can no longer be used.

(a) Explain why it would not be appropriate to use a census in these circumstances. **[1]**

A manager suggests that the first 5 components to be manufactured should be tested.

(b) Explain why this would not be a sensible method of selecting the sample. **[1]**

A statistician advises the manager that the sample selected should be a random sample.

(c) Give two desirable features (other than randomness) that the sample should have. **[2]**

4 Zara uses a metal detector to search for coins on a beach.
She wonders if the numbers of coins that she finds in an area of $10\,m^2$ can be modelled by a Poisson distribution. The table below shows the numbers of coins that she finds in randomly chosen areas of $10\,m^2$ over a period of months.

Number of coins found	0	1	2	3	4	5	6	>6
Frequency	13	28	30	14	10	2	3	0

(a) Software gives the sample mean as 1.98 and the sample standard deviation as 1.4212.
Explain how these values suggest that a Poisson distribution may be an appropriate model for the numbers of coins found. **[2]**

Zara decides to carry out a chi-squared test to investigate whether a Poisson distribution is an appropriate model.
Fig. 4 is a screenshot showing part of the spreadsheet used to analyse the data. Some values in the spreadsheet have been deliberately omitted.

	A	B	C	D
1	Number of coins found	Observed frequency	Expected frequency	Chi-squared contribution
2	0	13	13.8069	0.0472
3	1	28		
4	2	30	27.0643	0.3184
5	3	14	17.8625	0.8352
6	4	10	8.8419	0.1517
7	$\geqslant 5$	5		0.0015
8				

Fig. 4

(b) Showing your calculations, find the missing values in each of the following cells.

- C3
- C7
- D3

[4]

(c) Explain why the numbers for 5, 6 and more than 6 coins found have been combined into the single category of at least 5 coins found, as shown in the spreadsheet. **[1]**

(d) Complete the hypothesis test at the 5% level of significance. **[6]**

For the rest of this question, you should assume that the number of coins that Zara finds in an area of $10\,m^2$ can be modelled by a Poisson distribution with mean 1.98.
Zara also finds pieces of jewellery independently of the coins she finds. The number of pieces of jewellery that she finds per $10\,m^2$ area is modelled by a Poisson distribution with mean 0.42.

(e) Find the probability that Zara finds a total of exactly 3 items (coins and/or jewellery) in an area of $10\,m^2$. **[2]**

(f) Find the probability that Zara finds a total of at least 30 items (coins and/or jewellery) in an area of $100\,m^2$. **[2]**

5 A student wants to know if there is a positive correlation between the amounts of two pollutants, sulphur dioxide and PM10 particulates, on different days in the area of London in which he lives; these amounts, measured in suitable units, are denoted by s and p respectively.
He uses a government website to obtain data for a random sample of 15 days on which the amounts of these pollutants were measured simultaneously. Fig. 5.1 is a scatter diagram showing the data.
Summary statistics for these 15 values of s and p are as follows.

$\Sigma s = 155.4$ $\Sigma p = 518.9$ $\Sigma s^2 = 2322.7$ $\Sigma p^2 = 21270.5$ $\Sigma sp = 6009.1$

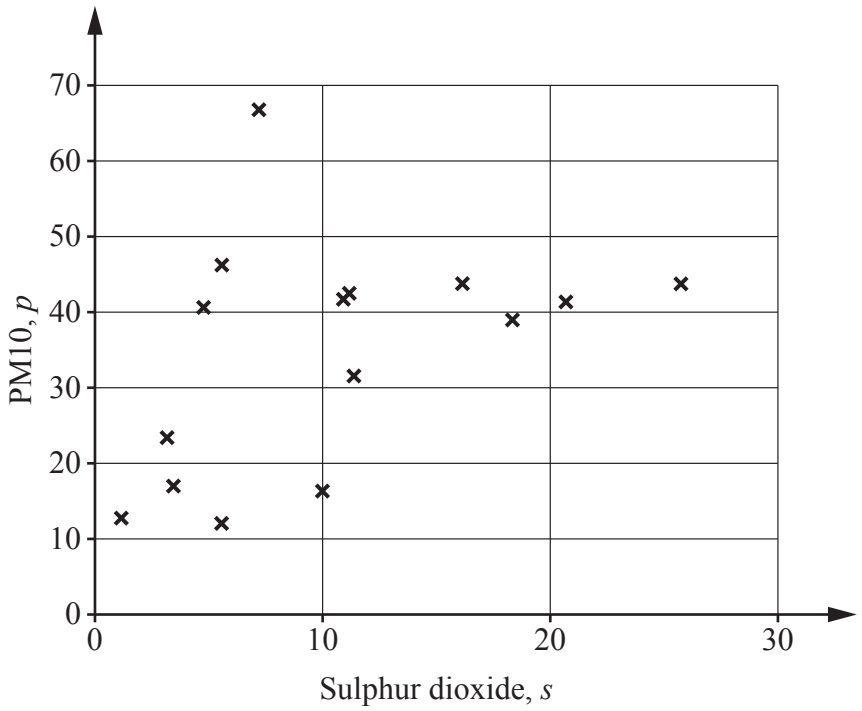

Fig. 5.1

(a) Explain why the student might come to the conclusion that a test based on Pearson's product moment correlation coefficient may be valid. **[2]**

(b) Find the value of Pearson's product moment correlation coefficient. **[4]**

(c) Carry out a test at the 5% significance level to investigate whether there is positive correlation between the amounts of sulphur dioxide and PM10 particulates. **[5]**

(d) Explain why the student made sure that the sample chosen was a random sample. **[2]**

The student also wishes to model the relationship between the amounts of nitrogen dioxide n and PM10 particulates p.
He takes a random sample of 54 values of the two variables, both measured at the same times. Fig. 5.2 is a scatter diagram which shows the data, together with the regression line of n on p, the equation of the regression line and the value of r^2.

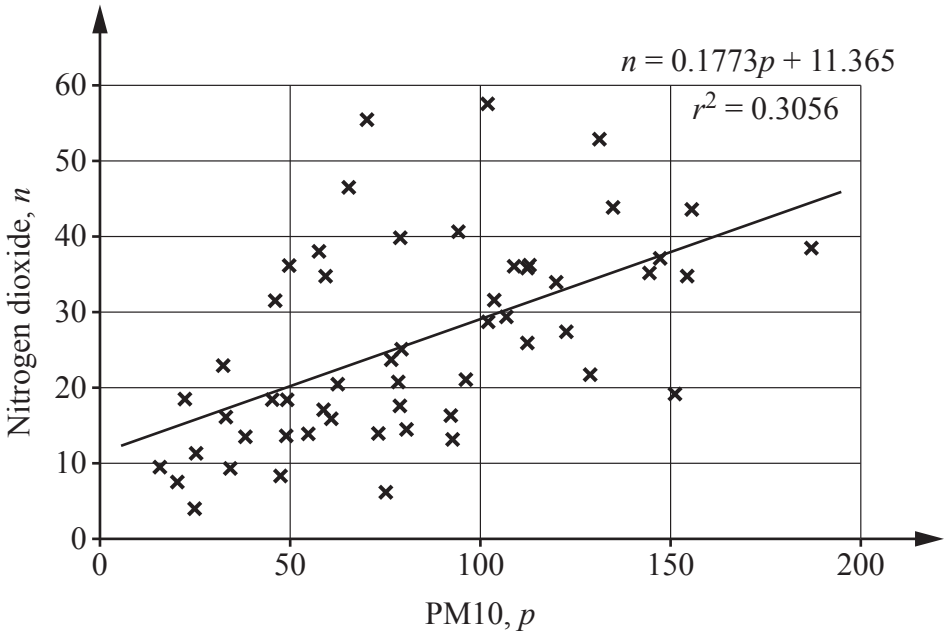

Fig. 5.2

(e) Predict the value of n for $p = 150$. **[1]**

(f) Discuss the reliability of your prediction in part **(e)**. **[2]**

6 The discrete random variable X has a uniform distribution over $\{n, n+1, ..., 2n\}$.

(a) Given that n is odd, find $P(X < \frac{3}{2}n)$. **[1]**

(b) Given instead that n is even, find $P(X < \frac{3}{2}n)$, giving your answer as a single algebraic fraction. **[3]**

(c) The sum of 6 independent values of X is denoted by Y.
Find $\text{Var}(Y)$. **[3]**

END OF QUESTION PAPER

BLANK PAGE

BLANK PAGE