



Oxford Cambridge and RSA

A Level Further Mathematics B (MEI)

Y432/01 Statistics Minor

Practice Paper – Set 1

Time allowed: 1 hour 15 minutes

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.
- Complete the boxes provided on the Printed Answer Booklet with your name, centre number and candidate number.
- 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.
- 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

- The total number of marks 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 (i) State the conditions required for a Poisson distribution to be a suitable model for the number of lightning strikes detected during a fixed interval of time in a thunderstorm. [2]

In a particular thunderstorm the number of lightning strikes detected per second is modelled by a Poisson distribution with mean 0.3.

- (ii) Find the probability that

(A) no lightning strikes are detected in a randomly chosen second, [1]

(B) at least 20 lightning strikes are detected in a randomly chosen period of 1 minute. [2]

- 2 In a TV quiz game, there are 4 questions which a contestant has to answer. For each question, the contestant has to choose the one correct answer from a number of possibilities. For the first question there are 2 possible answers. For each subsequent question there is 1 more possible answer than for the previous question, so that for the final question there are 5 possible answers. The random variable X represents the number of questions correctly answered by a contestant who randomly guesses each answer.

- (i) Show that $P(X = 4) = \frac{1}{120}$. [1]

The table below shows the probability distribution of X .

r	0	1	2	3	4
$P(X = r)$	$\frac{1}{5}$	$\frac{5}{12}$	$\frac{7}{24}$	$\frac{1}{12}$	$\frac{1}{120}$

- (ii) In this question, you must show detailed reasoning.

Find each of the following.

- $E(X)$
- $\text{Var}(X)$ [5]

Contestants are given £1000 for taking part in the quiz game, and an additional £500 for each answer which they get correct.

- (iii) (A) Find the expected total amount of money gained by a contestant who guesses each answer. [1]
- (B) Find the standard deviation of the total amount of money gained by a contestant who guesses each answer. [1]

- 3** A fair 12-sided dice has faces labelled 1, 2, 3, ..., 12. The random variable X represents the score when the dice is rolled once.
- (i)** Write down $P(X > 5)$. [1]
 - (ii)** Find the probability that a score greater than 5 occurs for the first time on the third roll. [1]
 - (iii)** The dice is rolled 10 times. Find the probability that the score is greater than 5 in at least 4 of the rolls. [2]
 - (iv)** The dice is rolled 4 times. Find the probability that at least 1 of the scores is greater than 5 and at least 1 of the scores is less than or equal to 5. [2]
 - (v)** The dice is rolled repeatedly until a score greater than 5 has occurred twice. Find the probability that at most 4 rolls are required. [2]
- 4** As part of an investigation into whether there is any link between income and well-being, a researcher selects a random sample of 10 people. The annual income £ x of each of the people is recorded, together with a well-being index y , based on a questionnaire which each of them fills in. Summary statistics for x and y are as follows.

$$\sum x = 263.1 \quad \sum y = 5.61 \quad \sum x^2 = 7545.91 \quad \sum y^2 = 3.4637 \quad \sum xy = 153.539$$

- (i)** Find the value of Pearson's product moment correlation coefficient. [4]
- (ii)** Given that the conditions required for a test based on this coefficient are valid, carry out a hypothesis test to examine at the 5% significance level whether there is positive correlation between income and well-being index. [5]
- (iii)** Briefly explain why it would not be appropriate to use a sample that consisted of people who all work for the same company. [1]

- 5 When steel is welded, there are often very small cracks in the weld. When steel is held at a high temperature for a long time, these cracks can gradually increase in length until eventually the weld may fail. The length l mm of a particular crack in a piece of steel was measured when the steel had been at a high temperature for t hours. The data, together with a scatter diagram which illustrates the data, are shown in Fig. 5. A scientist wants to construct a regression line to model the relationship between time and length.

Time t hours	500	1000	1500	2000	2500	3000	3500
Crack length l mm	0.10	0.21	0.34	0.42	0.55	0.72	1.19

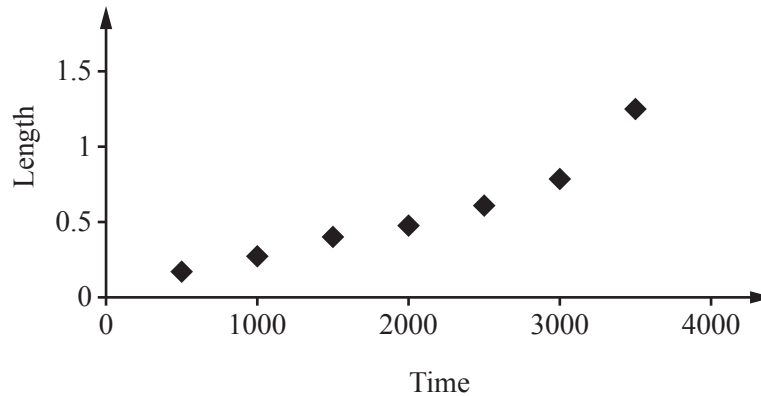


Fig. 5

- (i) Explain which of the two variables t and l is the independent variable. [1]
- (ii) Explain why it would be sensible to exclude the final data point (3500, 1.19). [1]
- (iii) Find the equation of the regression line of l on t excluding the final data point. [2]
- (iv) Find the value of the residual for $t = 2000$. [2]
- (v) Explain why it would not be sensible to make predictions of the values of l for values of t which are greater than 3000. [2]

- 6 In an investigation into bicycle use by children and adults in a particular country, a researcher selected a random sample of 300 people. The people in the sample were classified according to age and sex and whether or not they used a bicycle regularly.

Fig. 6 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	E
1	Observed frequencies				
2		Uses bicycle	Does not use bicycle	Totals	
3	Female child	18	27	45	
4	Male child	38	30	68	
5	Female adult	24	68	92	
6	Male adult	37	58	95	
7	Totals	117	183	300	
8					
9	Expected frequencies				
10	Female child	17.5500	27.4500		
11	Male child	26.5200			
12	Female adult	35.8800	56.1200		
13	Male adult	37.0500	57.9500		
14					
15	Contributions to the test statistic				
16	Female child	0.0115	0.0074		
17	Male child	4.9695	3.1772		
18	Female adult		2.5149		
19	Male adult	0.0001	0.0000		
20					

Fig. 6

- (i) State the null and alternative hypotheses for a test to investigate whether there is any association between the type of person and bicycle use. [1]
- (ii) Showing your calculations, find the missing values in each of the following cells. [3]
- C11
 - B18
- (iii) Complete the hypothesis test at the 1% level of significance. [5]
- (iv) Discuss briefly what the data suggest about bicycle use for different types of people. You should make a comment for each type of person. [3]

7 A biased coin has faces labelled with scores of 0 and 1. When the coin is spun, the probability that it lands showing a score of 1 is 0.4.

(i) Find the variance of the score when the coin is spun once. [2]

The coin is spun 10 times and the scores are added together to find the total score.

(ii) Using the result of part (i), find the variance of the total score, stating any assumption that you need to make. [2]

(iii) Explain how your answer to part (ii) relates to a binomial distribution. [2]

When the coin is spun n times, the probability that the total score is less than n is greater than 0.999.

(iv) Find the smallest possible value of n . [3]

END OF QUESTION PAPER

BLANK PAGE

OCR

Oxford Cambridge and RSA

Copyright Information

OCR is committed to seeking permission to reproduce all third-party content that it uses in its assessment materials. OCR has attempted to identify and contact all copyright holders whose work is used in this paper. To avoid the issue of disclosure of answer-related information to candidates, all copyright acknowledgements are reproduced in the OCR Copyright Acknowledgements Booklet. This is produced for each series of examinations and is freely available to download from our public website (www.ocr.org.uk) after the live examination series.

If OCR has unwittingly failed to correctly acknowledge or clear any third-party content in this assessment material, OCR will be happy to correct its mistake at the earliest possible opportunity.

For queries or further information please contact the Copyright Team, First Floor, 9 Hills Road, Cambridge CB2 1GE.

OCR is part of the Cambridge Assessment Group; Cambridge Assessment is the brand name of University of Cambridge Local Examinations Syndicate (UCLES), which is itself a department of the University of Cambridge.