Oxford Cambridge and RSA

Friday 13 June 2025 – Afternoon

A Level Further Mathematics B (MEI)

Y432/01 Statistics Minor

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

QP

1873515560

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.

© OCR 2025 [M/508/5596]
DC (SL) 354112/3
OCR is an exempt Charity

1 A discrete random variable, X, has a Poisson distribution with mean 4.

(a) Write down the value of $\mathrm{Var}(X)$. **[1]**

(b) Find $\mathrm{P}(X = 3)$. **[1]**

(c) Find $\mathrm{P}(X > 3)$. **[2]**

Another discrete random variable, Y, which is independent of X, has a Poisson distribution with mean 5.

(d) Find $\mathrm{P}(X + Y \leqslant 9)$. **[2]**

2 In an experiment, the speed of an object, $v\,\mathrm{m\,s^{-1}}$, at time t seconds, is recorded at intervals of 10 seconds from $t = 0$ to $t = 80$.

The summary statistics are as follows.

$n = 8$, $\Sigma t = 280$, $\Sigma t^2 = 14000$, $\Sigma v = 876.28$, $\Sigma v^2 = 133310.2$, $\Sigma tv = 43190$

A linear function of t is used as a model for v.

(a) By first determining the equation of an appropriate regression line, estimate the value of v when $t = 45$. **[3]**

(b) Comment on the reliability of your estimate in part **(a)**. **[1]**

(c) Explain why in this scenario it would be inappropriate to use a value of v to estimate a value of t using **any** regression line. **[1]**

3 An archer fires arrows one by one at a target until an arrow hits the target. The number of arrows fired up to and including the arrow that hits the target is denoted by X. The archer considers using a geometric model for the distribution of X.

(a) State **two** assumptions required for such a model to be appropriate. **[2]**

For the remainder of this question you may assume that a geometric model is appropriate for the distribution of X.

The archer keeps records of the values of X. He collects a sample of these values and uses them to estimate the variance of X.

(b) Give **two** features that the sample should have. **[2]**

The estimate of the variance of X calculated from the sample is 12.

(c) Determine an estimate of the probability that when the archer fires a single arrow at the target, the arrow hits the target. **[3]**

(d) Use your answer to part **(c)** to write down an estimate of the expected number of arrows that the archer must fire at the target up to and including the first arrow that hits the target. **[1]**

4 The discrete random variable, X, has a uniform distribution over the values $\{5, 6, ..., n\}$.

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

In part **(a)** you should assume that $n = 24$.

(i) Using the formula $\mathrm{E}(X) = \sum x_i p_i$, determine the value of $\mathrm{E}(X)$ by direct calculation. **[3]**

(ii) Use the substitution $Y = X - 4$ and the given formula in the formulae booklet to determine $\mathrm{Var}(X)$. **[2]**

(iii) Show that $\mathrm{P}(X > \mathrm{E}(X)) = \frac{1}{2}$. **[1]**

(b) A student states that $\mathrm{P}(X > \mathrm{E}(X)) = \frac{1}{2}$ for **all** $n \geqslant 24$.

Explain whether the student is correct. **[1]**

5 Long-term data collected at a certain farm indicates that 30% of eggs that it produces are large and the rest are categorised as medium.

Eggs are collected in batches of 4 and the number of large eggs in each batch, X, is recorded.

It is suggested that a binomial distribution might be appropriate to model X.

(a) **(i)** State **two** reasons, within the context of this question, why a binomial model might be appropriate to model X. **[2]**

(ii) State **two** assumptions, within the context of this question, that must be made for a binomial model to be appropriate to model X. You should not refer to the reasons given in part **(a)(i)**. **[2]**

(iii) Write down the binomial distribution which might be a suitable model for the distribution of X. **[1]**

The farm manager investigates whether the distribution in part **(a)(iii)** is a good model for X.

(b) State **one** reason why taking data from a sample of 6 batches would **not** be appropriate to carry out the investigation. **[1]**

The farm manager takes a random sample of 100 batches. The results are summarised in **Table 5.1**.

Table 5.1

x	0	1	2	$\geqslant 3$
Frequency	31	36	30	3

Table 5.2, shows expected frequencies for the values of X in 100 batches, assuming that X has the distribution from part **(a)(iii)**. The table is incomplete.

Table 5.2

x	0	1	2	$\geqslant 3$
Expected frequency	24.01			

(c) Complete the copy of **Table 5.2** in the **Printed Answer Booklet**. Give the expected frequencies to **2** decimal places. **[2]**

(d) Carry out a χ^2 test, at the 5% significance level, for goodness of fit of the model in part **(a)(iii)**. **[5]**

6 A factory discharges very low concentrations of an inert chemical through an outfall into a nearby body of water.

The concentration of the chemical is measured by sensors at seven known locations at fixed distances from the outfall. Each measurement is known to be affected by various environmental factors. None of these factors are measured or noted.

The manager of the factory needs to assess the relationship between the concentration of the chemical at each sensor and the distance of the sensor from the outfall.

The manager can carry out a hypothesis test **either** for any correlation **or** for any association between the concentration of the chemical at the sensor and the distance of the sensor from the outfall.

(a) Explain why the manager should carry out a test for association rather than one for correlation. **[1]**

The concentration of the chemical at each sensor and the distance of each sensor from the outfall, both in arbitrary units, for a particular day after the chemical has been discharged are given in the table below.

Sensor	A	B	C	D	E	F	G
Distance from outfall (arbitrary units)	1.2	6.6	8.9	11.0	15.6	23.2	30.5
Concentration of chemical (arb. units)	621	459	498	610	222	453	420

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

Carry out the appropriate test for association at the 5% significance level. A copy of the table, with spare, empty rows, is given in the **Printed Answer Booklet** to help with any calculations. **[8]**

Turn over for question 7

7 The probability function of the discrete random variable, X, is shown in the table. You are given that α, β and γ are constants.

x	-1	0	1
$\mathrm{P}(X = x)$	α	β	γ

(a) Find $\mathrm{E}(X)$ in terms of α and γ. **[1]**

(b) Find $\mathrm{Var}(X)$ in terms of α and γ. **[2]**

The random variable W is defined by $W = X_1 + X_2$ where X_1 and X_2 are independent random variables each representing an observation of X.

(c) Determine the probability distribution of W in terms of α, β and γ. Your answer should be entered in the empty table in the **Printed Answer Booklet**. You should not need all of the columns. **[4]**

(d) By forming an expression for $\mathrm{E}(W)$ from the table found in **(c)** and then eliminating β, show that $\mathrm{E}(W) = 2\mathrm{E}(X)$. **[4]**

(e) In the case where $\alpha = 0.1$ and $\beta = 0.7$, find $\mathrm{P}(W = 0 | X_1 = 1)$. **[1]**

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 OCR Copyright Team, The Triangle Building, Shaftesbury Road, Cambridge CB2 8EA.

OCR is part of Cambridge University Press & Assessment, which is itself a department of the University of Cambridge.

© OCR 2025