OXFORD CAMBRIDGE AND RSA EXAMINATIONS

A2 GCE

4733/01

MATHEMATICS

Probability & Statistics 2

QUESTION PAPER

FRIDAY 21 JUNE 2013: Morning

DURATION: 1 hour 30 minutes plus your additional time allowance

MODIFIED ENLARGED

Candidates answer on the Printed Answer Book or any suitable paper provided by the centre. The Printed Answer Book may be enlarged by the centre.

OCR SUPPLIED MATERIALS:

Printed Answer Book 4733/01
List of Formulae (MF1)

OTHER MATERIALS REQUIRED:

Scientific or graphical calculator

READ INSTRUCTIONS OVERLEAF

DC (SLM) 1644937118 80778/1

INSTRUCTIONS TO CANDIDATES

- Write your name, centre number and candidate number in the spaces provided on the Printed Answer Book or on the paper provided by the centre. Please write clearly and in capital letters.
- IF YOU USE THE PRINTED ANSWER BOOK WRITE YOUR ANSWER TO EACH QUESTION IN THE SPACE PROVIDED. Additional paper may be used if necessary but you must clearly show your candidate number, centre number and question number(s).
- Use black ink. HB pencil may be used for graphs and diagrams only.
- Answer ALL the questions.
- Read each question carefully. Make sure you know what you have to do before starting your answer.
- You are permitted to use a scientific or graphical calculator in this paper.
- Give non-exact numerical answers correct to 3 significant figures unless a different degree of accuracy is specified in the question or is clearly appropriate.

INFORMATION FOR CANDIDATES

- **The number of marks is given in brackets [ ] at the end of each question or part question on the Question Paper.**
- **YOU ARE REMINDED OF THE NEED FOR CLEAR PRESENTATION IN YOUR ANSWERS.**
- **The total number of marks for this paper is 72.**

INSTRUCTION TO EXAMS OFFICER / INVIGILATOR

- **Do not send this Question Paper for marking; it should be retained in the centre or recycled. Please contact OCR Copyright should you wish to re-use this document.**

1 **It is required to select a random sample of 30 pupils from a school with 853 pupils. A student suggests the following method.**
"Give each pupil sequentially a three-digit number from 001 to 853. Use a calculator to generate random three-digit numbers from 0.000 to 0.999 inclusive, multiply the answer by 853, add 1 and round off to the nearest whole number. Select the corresponding pupil, and repeat as necessary".

(i) **Determine which pupil would be picked for each of the following calculator outputs:**

0.103, 0.104, 0.105, 0.106, 0.107. [2]

(ii) **Use your answers to part (i) to show that this method is biased, and suggest an improvement. [2]**

2 **The number of neutrinos that pass through a certain region in one second is a random variable with the distribution $\mathrm{Po}\left(5 \times 10^4\right)$. Use a suitable approximation to calculate the probability that the number of neutrinos passing through the region in 40 seconds is less than 1.999×10^6. [4]**

3 **The mean of a sample of 80 independent observations of a continuous random variable Y is denoted by $\overline{Y}$. It is given that $\mathrm{P}\left(\overline{Y} \leqslant 157.18\right) = 0.1$ and $\mathrm{P}\left(\overline{Y} \geqslant 164.76\right) = 0.7$.**

(i) Calculate $\mathrm{E}(Y)$ and the standard deviation of Y. [6]

(ii) State

(a) where in your calculations you have used the Central Limit Theorem,

(b) why it was necessary to use the Central Limit Theorem,

(c) why it was possible to use the Central Limit Theorem. [3]

4 **The number of floods in a certain river plain is known to have a Poisson distribution. It is known that up until 10 years ago the mean number of floods per year was 0.32. During the last 10 years there were 6 floods. Test at the 1% significance level whether there is evidence of an increase in the mean number of floods per year. [7]**

5 **Two random variables S and T have probability density functions given by**

$$\mathrm{f}_S(x) = \begin{cases} \dfrac{3}{a^3}(x-a)^2 & 0 \leqslant x \leqslant a, \\ 0 & \text{otherwise,} \end{cases}$$

$$\mathrm{f}_T(x) = \begin{cases} c & 0 \leqslant x \leqslant a, \\ 0 & \text{otherwise,} \end{cases}$$

where a and c are constants.

(i) On a single diagram sketch both probability density functions. [3]

(ii) Calculate the mean of S, in terms of a. [5]

(iii) Use your diagram to explain which of S or T has the bigger variance. (Answers obtained by calculation will score no marks.) [2]

6 **The random variable X denotes the yield, in kilograms per acre, of a certain crop. Under the standard treatment it is known that $\mathrm{E}(X) = 38.4$. Under a new treatment, the yields of 50 randomly chosen regions can be summarised as**

$n = 50,$

$\sum x = 1834.0,$

$\sum x^2 = 70027.37.$

Test at the 1% level whether there has been a change in the mean crop yield. [11]

7 **Past experience shows that 35% of the senior pupils in a large school know the regulations about bringing cars to school. The head teacher addresses this subject in an assembly, and afterwards a random sample of 120 senior pupils is selected. In this sample it is found that 50 of these pupils know the regulations. Use a suitable approximation to test, at the 10% significance level, whether there is evidence that the proportion of senior pupils who know the regulations has increased. Justify your approximation. [11]**

8 **The random variable R has the distribution B(14, p). A test is carried out at the α% significance level of the null hypothesis H_0: $p = 0.25$, against H_1: $p > 0.25$.**

(i) Given that α is as close to 5 as possible, find the probability of a Type II error when the true value of p is 0.4. [4]

(ii) State what happens to the probability of a Type II error as

(a) p increases from 0.4,

(b) α increases, giving a reason. [2]

9 **The managers of a car breakdown recovery service are discussing whether the number of breakdowns per day can be modelled by a Poisson distribution. They agree that breakdowns occur randomly. Manager *A* says, "it must be assumed that breakdowns occur at a constant rate throughout the day".**

(i) Give an improved version of Manager *A*'s statement, and explain why the improvement is necessary. [2]

(ii) Explain whether you think your improved statement is likely to hold in this context. [1]

Assume now that the number *B* of breakdowns per day can be modelled by the distribution Po(λ).

(iii) Given that $\lambda = 9.0$ and $P(B > B_0) < 0.1$, use tables to find the smallest possible value of B_0, and state the corresponding value of $P(B > B_0)$. [2]

(iv) Given that $P(B = 2) = 0.0072$, show that λ satisfies an equation of the form $\lambda = 0.12e^{k\lambda}$, for a value of *k* to be stated. Evaluate the expression $0.12e^{k\lambda}$ for $\lambda = 8.5$ and $\lambda = 8.6$, giving your answers correct to 4 decimal places. What can be deduced about a possible value of λ? [5]