OCR
RECOGNISING ACHIEVEMENT

Friday 24 May 2013 – Morning

A2 GCE MATHEMATICS (MEI)

4769/01 Statistics 4

QUESTION PAPER

Candidates answer on the Printed Answer Book.

OCR supplied materials:
- Printed Answer Book 4769/01
- MEI Examination Formulae and Tables (MF2)

Other materials required:
- Scientific or graphical calculator

Duration: 1 hour 30 minutes

4715880613

INSTRUCTIONS TO CANDIDATES

These instructions are the same on the Printed Answer Book and the Question Paper.

- The Question Paper will be found in the centre of the Printed Answer Book.
- Write your name, centre number and candidate number in the spaces provided on the Printed Answer Book. Please write clearly and in capital letters.
- **Write your answer to each question in the space provided in the Printed Answer Book**. 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.
- Read each question carefully. Make sure you know what you have to do before starting your answer.
- Answer any **three** questions.
- Do **not** write in the bar codes.
- 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 FOR CANDIDATES

This information is the same on the Printed Answer Book and the Question Paper.

- The number of marks is given in brackets **[ ]** at the end of each question or part question on the Question Paper.
- 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 being used.
- The total number of marks for this paper is **72**.
- The Printed Answer Book consists of **16** pages. The Question Paper consists of **4** pages. Any blank pages are indicated.

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.

Option 1: Estimation

1 Traffic engineers are studying the flow of vehicles along a road. At an initial stage of the investigation, they assume that the average flow remains the same throughout the working day. An automatic counter records the number of vehicles passing a certain point per minute during the working day. A random sample of these records is selected; the sample values are denoted by $x_1, x_2, \ldots, x_n$.

(i) The engineers model the underlying random variable X by a Poisson distribution with unknown parameter θ. Obtain the likelihood of $x_1, x_2, \ldots, x_n$ and hence find the maximum likelihood estimate of θ. **[10]**

(ii) Write down the maximum likelihood estimate of the probability that no vehicles pass during a minute. **[3]**

(iii) The engineers note that, in a sample of size 1000 with sample mean $\bar{x} = 5$, there are no observations of zero. Suggest why this might cast some doubt on the investigation. **[3]**

(iv) On checking the automatic counter, the engineers find that, due to a fault, no record at all is made if no vehicle passes in a minute. They therefore model X as a Poisson random variable, again with an unknown parameter θ, except that the value $x = 0$ cannot occur. Show that, under this model,

$$\mathrm{P}(X = x) = \frac{\theta^x}{\left(\mathrm{e}^\theta - 1\right)x!}, \quad x = 1, 2, \ldots,$$

and hence show that the maximum likelihood estimate of θ satisfies the equation

$$\frac{\theta \mathrm{e}^\theta}{\mathrm{e}^\theta - 1} = \bar{x}.$$ **[8]**

Option 2: Generating Functions

2 The random variable X takes values -2, 0 and 2, each with probability $\frac{1}{3}$.

(i) Write down the values of

(A) μ, the mean of X,

(B) $\mathrm{E}(X^2)$,

(C) σ^2, the variance of X. **[3]**

(ii) Obtain the moment generating function (mgf) of X. **[2]**

A random sample of n independent observations on X has sample mean $\overline{X}$, and the standardised mean is denoted by Z where

$$Z = \frac{\overline{X} - \mu}{\frac{\sigma}{\sqrt{n}}}.$$

(iii) Stating carefully the required general results for mgfs of sums and of linear transformations, show that the mgf of Z is

$$\mathrm{M}_Z(\theta) = \left\{\frac{1}{3}\left(1 + \mathrm{e}^{\frac{\theta\sqrt{3}}{\sqrt{2n}}} + \mathrm{e}^{-\frac{\theta\sqrt{3}}{\sqrt{2n}}}\right)\right\}^n. \quad \textbf{[8]}$$

(iv) By expanding the exponential functions in $\mathrm{M}_Z(\theta)$, show that, for large n,

$$\mathrm{M}_Z(\theta) \approx \left(1 + \frac{\theta^2}{2n}\right)^n. \quad \textbf{[7]}$$

(v) Use the result $\mathrm{e}^y = \lim_{n\to\infty}\left(1 + \frac{y}{n}\right)^n$ to find the limit of $\mathrm{M}_Z(\theta)$ as $n \to \infty$, and deduce the approximate distribution of Z for large n. **[4]**

Option 3: Inference

3 **(i)** Explain the meaning of the following terms in the context of hypothesis testing: Type I error, Type II error, operating characteristic, power. **[8]**

(ii) A test is to be carried out concerning a parameter θ. The null hypothesis is that θ has the particular value θ_0. The alternative hypothesis is $\theta \neq \theta_0$. Draw a sketch of the operating characteristic for a perfect test that never makes an error. **[3]**

(iii) The random variable X is distributed as $\mathrm{N}(\mu, 9)$. A random sample of size 25 is available. The null hypothesis $\mu = 0$ is to be tested against the alternative hypothesis $\mu \neq 0$. The null hypothesis will be accepted if $-1 < \bar{x} < 1$ where $\bar{x}$ is the value of the sample mean, otherwise it will be rejected. Calculate the probability of a Type I error. Calculate the probability of a Type II error if in fact $\mu = 0.5$; comment on the value of this probability. **[9]**

(iv) Without carrying out any further calculations, draw a sketch of the operating characteristic for the test in part **(iii)**. **[4]**

Option 4: Design and Analysis of Experiments

4 **(i)** Explain the advantages of randomisation and replication in a statistically designed experiment. **[6]**

(ii) The usual statistical model underlying the one-way analysis of variance is given, in the usual notation, by

$$x_{ij} = \mu + \alpha_i + e_{ij}$$

where x_{ij} denotes the jth observation on the ith treatment. Define carefully all the terms in this model and state the properties of the term that represents experimental error. **[7]**

(iii) A trial of five fertilisers is carried out at an agricultural research station according to a completely randomised design in which each fertiliser is applied to four experimental plots of a crop (so that there are 20 experimental units altogether). The sums of squares in a one-way analysis of variance of the resulting data on yields of the crop are as follows.

Source of variation	Sum of squares
Between fertilisers	219.2
Residual	304.5
Total	523.7

State the customary null and alternative hypotheses that are tested. Provide the degrees of freedom for each sum of squares. Hence copy and complete the analysis of variance table and carry out the test at the 5% level. **[11]**

OCR
RECOGNISING ACHIEVEMENT