OCR
RECOGNISING ACHIEVEMENT

ADVANCED GCE **4735/01**

MATHEMATICS

Probability & Statistics 4

WEDNESDAY 18 JUNE 2008

Morning

Time: 1 hour 30 minutes

Additional materials (enclosed): None

Additional materials (required):
Answer Booklet (8 pages)
List of Formulae (MF1)

INSTRUCTIONS TO CANDIDATES

- Write your name in capital letters, your Centre Number and Candidate Number in the spaces provided on the Answer Booklet.
- Read each question carefully and make sure you know what you have to do before starting your answer.
- Answer **all** the questions.
- 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.
- You are permitted to use a graphical calculator in this paper.

INFORMATION FOR CANDIDATES

- The number of marks is given in brackets [ ] at the end of each question or part question.
- The total number of marks for this paper is **72**.
- **You are reminded of the need for clear presentation in your answers.**

This document consists of **4** printed pages.

1 For the mutually exclusive events A and B, $\mathrm{P}(A) = \mathrm{P}(B) = x$, where $x \neq 0$.

(i) Show that $x \leqslant \frac{1}{2}$. [1]

(ii) Show that A and B are not independent. [2]

The event C is independent of A and also independent of B, and $\mathrm{P}(C) = 2x$.

(iii) Show that $\mathrm{P}(A \cup B \cup C) = 4x(1 - x)$. [4]

2 Part of Helen's psychology dissertation involved the reaction times to a certain stimulus. She measured the reaction times of 30 randomly selected students, in seconds correct to 2 decimal places. The results are shown in the following stem-and-leaf diagram.

14	1 2
15	2 4
16	0 3 6
17	1 5 7
18	3 4 5 7 9
19	2 4 6 7 8 9
20	0 1 3 4 5 7 8 9
21	7

Key: 18 | 3 means 1.83 seconds

Helen wishes to test whether the population median time exceeds 1.80 seconds.

(i) Give a reason why the Wilcoxon signed-rank test should not be used. [1]

(ii) Carry out a suitable non-parametric test at the 5% significance level. [7]

3 From the records of Mulcaster United Football Club the following distribution was suggested as a probability model for future matches. X and Y denoted the numbers of goals scored by the home team and the away team respectively.

		X = 0	1	2	3
Y	0	0.11	0.04	0.06	0.08
	1	0.08	0.05	0.12	0.05
	2	0.05	0.08	0.07	0.03
	3	0.03	0.06	0.07	0.02

Use the model to find

(i) $\mathrm{E}(X)$, [3]

(ii) the probability that the away team wins a randomly chosen match, [2]

(iii) the probability that the away team wins a randomly chosen match, given that the home team scores. [4]

One of the directors, an amateur statistician, finds that $\mathrm{Cov}(X, Y) = 0.007$. He states that, as this value is very close to zero, X and Y may be considered to be independent.

(iv) Comment on the director's statement. [2]

4 William takes a bus regularly on the same journey, sometimes in the morning and sometimes in the afternoon. He wishes to compare morning and afternoon journey times. He records the journey times on 7 randomly chosen mornings and 8 randomly chosen afternoons. The results, each correct to the nearest minute, are as follows, where M denotes a morning time and A denotes an afternoon time.

M	A	A	M	M	M	M	M	M	A	A	A	A	A	A
19	20	22	24	25	26	28	30	31	33	35	37	38	39	42

William wishes to test for a difference between the average times of morning and afternoon journeys.

(i) Given that $s_M^2 = 16.5$ and $s_A^2 = 64.5$, with the usual notation, explain why a t-test is not appropriate in this case. [1]

(ii) William chooses a non-parametric test at the 5% significance level. Carry out the test, stating the rejection region. [6]

5 The discrete random variable X has moment generating function $\frac{1}{4}\mathrm{e}^{2t} + a\mathrm{e}^{3t} + b\mathrm{e}^{4t}$, where a and b are constants. It is given that $\mathrm{E}(X) = 3\frac{3}{8}$.

(i) Show that $a = \frac{1}{8}$, and find the value of b. [6]

(ii) Find $\mathrm{Var}(X)$. [4]

(iii) State the possible values of X. [1]

6 The continuous random variable Y has cumulative distribution function given by

$$\mathrm{F}(y) = \begin{cases} 0 & y < a, \\ 1 - \dfrac{a^3}{y^3} & y \geqslant a, \end{cases}$$

where a is a positive constant. A random sample of 3 observations, Y_1, Y_2, Y_3, is taken, and the smallest is denoted by S.

(i) Show that $\mathrm{P}(S > s) = \left(\dfrac{a}{s}\right)^9$ and hence obtain the probability density function of S. [5]

(ii) Show that S is not an unbiased estimator of a, and construct an unbiased estimator, T_1, based on S. [4]

It is given that T_2, where $T_2 = \frac{2}{9}(Y_1 + Y_2 + Y_3)$, is another unbiased estimator of a.

(iii) Given that $\mathrm{Var}(Y) = \frac{3}{4}a^2$ and $\mathrm{Var}(S) = \frac{9}{448}a^2$, determine which of T_1 and T_2 is the more efficient estimator. [3]

(iv) The values of Y for a particular sample are 12.8, 4.5 and 7.0. Find the values of T_1 and T_2 for this sample, and give a reason, unrelated to efficiency, why T_1 gives a better estimate of a than T_2 in this case. [3]

7 The probability generating function of the random variable X is given by

$$\mathrm{G}(t) = \frac{1 + at}{4 - t},$$

where a is a constant.

(i) Find the value of a. [2]

(ii) Find $\mathrm{P}(X = 3)$. [4]

The sum of 3 independent observations of X is denoted by Y. The probability generating function of Y is denoted by $\mathrm{H}(t)$.

(iii) Use $\mathrm{H}(t)$ to find $\mathrm{E}(Y)$. [5]

(iv) By considering $\mathrm{H}(-1) + \mathrm{H}(1)$, show that $\mathrm{P}(Y \text{ is an even number}) = \frac{62}{125}$. [2]