OCR
RECOGNISING ACHIEVEMENT

ADVANCED GCE UNIT

4735/01

MATHEMATICS

Probability & Statistics 4

FRIDAY 22 JUNE 2007

Morning

Time: 1 hour 30 minutes

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

INSTRUCTIONS TO CANDIDATES

- Write your name, centre number and candidate number in the spaces provided on the answer booklet.
- 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.

ADVICE TO CANDIDATES

- Read each question carefully and make sure you know what you have to do before starting your answer.
- **You are reminded of the need for clear presentation in your answers.**

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

OCR is an exempt Charity

[Turn over

1 For the events A and B, $P(A) = 0.3$, $P(B) = 0.6$ and $P(A' \cap B') = c$, where $c \neq 0$.

(i) Find $P(A \cap B)$ in terms of c. [3]

(ii) Find $P(B \mid A)$ and deduce that $0.1 \leqslant c \leqslant 0.4$. [3]

2 Of 9 randomly chosen students attending a lecture, 4 were found to be smokers and 5 were non-smokers. During the lecture their pulse-rates were measured, with the following results in beats per minute.

Smokers	77	85	90	98	
Non-smokers	59	64	68	80	88

It may be assumed that these two groups of students were random samples from the student populations of smokers and non-smokers. Using a suitable Wilcoxon test at the 10% significance level, test whether there is a difference in the median pulse-rates of the two populations. [7]

3 The discrete random variables X and Y have the joint probability distribution given in the following table.

		X		
		-1	0	1
Y	1	0.24	0.22	0.04
	2	0.26	0.18	0.06

(i) Show that $\text{Cov}(X, Y) = 0$. [5]

(ii) Find the conditional distribution of X given that $Y = 2$. [2]

4 The levels of impurity in a particular alloy were measured using a random sample of 20 specimens. The results, in suitable units, were as follows.

3.00	2.05	3.15	2.65	3.50	3.25	2.85	3.35	2.65	2.75
2.90	2.20	2.95	3.05	3.65	3.45	2.55	2.15	2.80	2.60

(i) Use the sign test, at the 5% significance level, to decide if there is evidence that the population median level of impurity is greater than 2.70. [7]

(ii) State what other test might have been used, and give one advantage and one disadvantage this other test has over the sign test. [3]

5 The continuous random variable X has probability density function given by

$$\mathrm{f}(x) = \begin{cases} \dfrac{1}{(\alpha-1)!}x^{\alpha-1}\mathrm{e}^{-x} & x \geqslant 0, \\ 0 & x < 0, \end{cases}$$

where α is a positive integer.

(i) Explain how you can deduce that $\int_0^{\infty} x^{\alpha-1}\mathrm{e}^{-x}\,\mathrm{d}x = (\alpha-1)!$. [1]

(ii) Write down an integral for the moment generating function $\mathrm{M}_X(t)$ of X and show, by using the substitution $x = \dfrac{u}{1-t}$, that $\mathrm{M}_X(t) = (1-t)^{-\alpha}$. [5]

(iii) Use the moment generating function to find, in terms of α,

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

(b) $\mathrm{Var}(X)$. [3]

6 The discrete random variable X takes the values 0 and 1 with $\mathrm{P}(X=0) = q$ and $\mathrm{P}(X=1) = p$, where $p + q = 1$.

(i) Write down the probability generating function of X. [1]

The sum of n independent observations of X is denoted by S.

(ii) Write down the probability generating function of S, and name the distribution of S. [2]

(iii) Use the probability generating function of S to find $\mathrm{E}(S)$ and $\mathrm{Var}(S)$. [6]

(iv) The independent random variables Y and Z are such that Y has the distribution $\mathrm{B}(10, \frac{1}{2})$, and Z has probability generating function $\mathrm{e}^{-(1-t)}$. Find the probability that the sum of one random observation of Y and one random observation of Z is equal to 2. [6]

[Question 7 is printed overleaf.]

7 The continuous random variable X has a uniform distribution over the interval $[0, \theta]$ so that the probability density function is given by

$$f(x) = \begin{cases} \dfrac{1}{\theta} & 0 \leqslant x \leqslant \theta, \\ 0 & \text{otherwise,} \end{cases}$$

where θ is a positive constant. A sample of n independent observations of X is taken and the sample mean is denoted by $\overline{X}$.

(i) The estimator T_1 is defined by $T_1 = 2\overline{X}$. Show that T_1 is an unbiased estimator of θ. [2]

It is given that the probability density function of the largest value, U, in the sample is

$$g(u) = \begin{cases} \dfrac{nu^{n-1}}{\theta^n} & 0 \leqslant u \leqslant \theta, \\ 0 & \text{otherwise.} \end{cases}$$

(ii) Find $E(U)$ and show that $\text{Var}(U) = \dfrac{n\theta^2}{(n+1)^2(n+2)}$. [6]

(iii) The estimator T_2 is defined by $T_2 = \dfrac{n+1}{n}U$. Given that T_2 is also an unbiased estimator of θ, show that T_2 is a more efficient estimator than T_1 for $n > 1$. [7]