OCR RECOGNISING ACHIEVEMENT

OXFORD CAMBRIDGE AND RSA EXAMINATIONS

Advanced Subsidiary General Certificate of Education
Advanced General Certificate of Education

MATHEMATICS 4735

Probability & Statistics 4

Wednesday **21 JUNE 2006** Afternoon 1 hour 30 minutes

Additional materials:
8 page answer booklet
Graph paper
List of Formulae (MF1)

TIME 1 hour 30 minutes

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.
- Questions carrying smaller numbers of marks are printed earlier in the paper, and questions carrying larger numbers of marks later in the paper.
- **You are reminded of the need for clear presentation in your answers.**

This question paper consists of 3 printed pages and 1 blank page.

1 **(i)** State whether the following are true or false for the random variables X and Y.

(a) X and Y are independent $\Longrightarrow \text{Cov}(X, Y) = 0$.

(b) $\text{Cov}(X, Y) = 0 \Longrightarrow X$ and Y are independent.

(c) $\text{Cov}(X, Y) \neq 0 \Longrightarrow X$ and Y are not independent.

[2]

(ii) Given that $\text{Var}(X) = 2$, $\text{Var}(Y) = 3$ and $\text{Var}(2X - Y) = 6$, find $\text{Cov}(X, Y)$. [4]

2 Out of a sample of 60 pairs of twin boys it was found that the first-born in 37 of the pairs was taller than the second-born. In the remaining 23 pairs the second-born was taller.

Stating clearly a necessary assumption, carry out a test at the 5% significance level of whether, in a majority of pairs of twin boys, the first-born is taller than the second-born. [9]

3 5% of valves manufactured in a certain factory are faulty. Each of the valves is tested by a machine which classifies 98% of faulty valves as faulty. It classifies 96% of non-faulty valves as non-faulty. F denotes the event 'a valve is faulty' and C denotes the event 'a valve is classified as faulty'.

(i) Show that $\text{P}(C) = 0.087$. [3]

(ii) Find $\text{P}(F \mid C)$. [3]

Each month 5000 valves are sold, all of which the machine classified as non-faulty.

(iii) Find the expected number of faulty valves sold each month. [4]

4 The continuous random variable X has a uniform distribution with probability density function given by

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

where a and b are constants. This distribution is denoted by $\text{U}(a, b)$.

(i) Show that the moment generating function of X is $\dfrac{\text{e}^{bt} - \text{e}^{at}}{t(b-a)}$. [3]

For the independent random variables X_1 and X_2, $X_1 \sim \text{U}(-1, 0)$ and $X_2 \sim \text{U}(0, 1)$.

(ii) Find the moment generating function of T, where $T = X_1 + X_2$. [2]

S denotes the sum of two independent observations of Y, where $Y \sim \text{U}\left(-\frac{1}{2}, \frac{1}{2}\right)$.

(iii) Show that S has the same moment generating function as T, and state what this indicates about the distributions of S and T. [4]

5 A random sample of 13 observations of a continuous random variable is taken, and the values ranked from 1 to 13. Four of these rankings are selected at random. The order in which the rankings are selected is irrelevant.

(i) Calculate how many possible different selections there are of the 4 rankings. [2]

The sum of the 4 rankings is denoted by R.

(ii) List all the selections of 4 rankings for which $R \leqslant 13$, and hence obtain the exact value of $P(R \leqslant 13)$. [3]

The distributions of the continuous random variables X and Y have the same shape. A random sample of 4 observations of X and a random sample of 9 observations of Y are taken and the 13 observations are ranked. The sum of the ranks of the 4 observations of X is 13.

(iii) Naming the test used, and stating the null and alternative hypotheses, show that the samples give evidence of a difference in the medians of X and Y at a significance level smaller than 2%. [5]

6 The discrete random variable Y has probability generating function given by

$$G(t) = \frac{0.8t}{1 - 0.2t}.$$

(i) Show that $E(Y) = \frac{5}{4}$. [3]

(ii) Express $P(Y = r)$ in terms of r, giving the possible values of r. [4]

(iii) By identifying the probability distribution of Y, or otherwise, find $\text{Var}(Y)$. [3]

(iv) Find $P(T \geqslant 8)$, where T is the sum of 6 independent observations of Y. [3]

7 During the Great War, *Invictor* tanks were used, but records of how many were manufactured have been lost. It may be assumed that the tanks had integer serial numbers ranging consecutively from 1 to n, where n is unknown. Suppose that a randomly selected tank has serial number denoted by X.

(i) Write down $E(X)$ and show that $\text{Var}(X) = \frac{1}{12}(n^2 - 1)$. [4]

The remains of two *Invictor* tanks have serial numbers denoted by X_1 and X_2.

(ii) Show that $N_1 = X_1 + X_2 - 1$ is an unbiased estimator of n. [2]

The larger of X_1 and X_2 is denoted by M.

(iii) Show that $P(M = r) = \dfrac{2(r-1)}{n(n-1)}$, for $r = 2, 3, \ldots, n$. [3]

(iv) Using the result $\sum_{r=2}^{n} r(r-1) = \frac{1}{3}(n^3 - n)$, find $E(M)$ and hence construct another unbiased estimator, N_2, of n. [3]

(v) Given that the variance of N_1 is $\frac{1}{6}(n^2 - n - 2)$ and that N_1 is a more efficient estimator than N_2, obtain an inequality for $\text{Var}(M)$. [3]

BLANK PAGE