OCR
RECOGNISING ACHIEVEMENT

OXFORD CAMBRIDGE AND RSA EXAMINATIONS

Advanced Subsidiary General Certificate of Education
Advanced General Certificate of Education

MATHEMATICS 4733

Probability & Statistics 2

Thursday **15 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.

[Turn over

1 Calculate the variance of the continuous random variable with probability density function given by

$$f(x) = \begin{cases} \frac{3}{37}x^2 & 3 \leqslant x \leqslant 4, \\ 0 & \text{otherwise.} \end{cases}$$ [6]

2 **(i)** The random variable R has the distribution $B(6, p)$. A random observation of R is found to be 6. Carry out a 5% significance test of the null hypothesis H_0: $p = 0.45$ against the alternative hypothesis H_1: $p \neq 0.45$, showing all necessary details of your calculation. [4]

(ii) The random variable S has the distribution $B(n, p)$. H_0 and H_1 are as in part **(i)**. A random observation of S is found to be 1. Use tables to find the largest value of n for which H_0 is not rejected. Show the values of any relevant probabilities. [3]

3 The continuous random variable T has mean μ and standard deviation σ. It is known that $P(T < 140) = 0.01$ and $P(T < 300) = 0.8$.

(i) Assuming that T is normally distributed, calculate the values of μ and σ. [6]

In fact, T represents the time, in minutes, taken by a randomly chosen runner in a public marathon, in which about 10% of runners took longer than 400 minutes.

(ii) State with a reason whether the mean of T would be higher than, equal to, or lower than the value calculated in part **(i)**. [2]

4 **(i)** Explain briefly what is meant by a random sample. [1]

Random numbers are used to select, with replacement, a sample of size n from a population numbered 000, 001, 002, …, 799.

(ii) If $n = 6$, find the probability that exactly 4 of the selected sample have numbers less than 500. [3]

(iii) If $n = 60$, use a suitable approximation to calculate the probability that at least 40 of the selected sample have numbers less than 500. [6]

5 An airline has 300 seats available on a flight to Australia. It is known from experience that on average only 99% of those who have booked seats actually arrive to take the flight, the remaining 1% being called 'no-shows'. The airline therefore sells more than 300 seats. If more than 300 passengers then arrive, the flight is over-booked. Assume that the number of no-show passengers can be modelled by a binomial distribution.

(i) If the airline sells 303 seats, state a suitable distribution for the number of no-show passengers, and state a suitable approximation to this distribution, giving the values of any parameters. [2]

Using the distribution and approximation in part **(i)**,

(ii) show that the probability that the flight is over-booked is 0.4165, correct to 4 decimal places, [2]

(iii) find the largest number of seats that can be sold for the probability that the flight is over-booked to be less than 0.2. [5]

6 Customers arrive at a post office at a constant average rate of 0.4 per minute.

(i) State an assumption needed to model the number of customers arriving in a given time interval by a Poisson distribution. [1]

Assuming that the use of a Poisson distribution is justified,

(ii) find the probability that more than 2 customers arrive in a randomly chosen 1-minute interval, [2]

(iii) use a suitable approximation to calculate the probability that more than 55 customers arrive in a given two-hour interval, [6]

(iv) calculate the smallest time for which the probability that no customers arrive in that time is less than 0.02, giving your answer to the nearest second. [5]

7 Three independent researchers, A, B and C, carry out significance tests on the power consumption of a manufacturer's domestic heaters. The power consumption, X watts, is a normally distributed random variable with mean μ and standard deviation 60. Each researcher tests the null hypothesis H_0: $\mu = 4000$ against the alternative hypothesis H_1: $\mu > 4000$.

Researcher A uses a sample of size 50 and a significance level of 5%.

(i) Find the critical region for this test, giving your answer correct to 4 significant figures. [6]

In fact the value of μ is 4020.

(ii) Calculate the probability that Researcher A makes a Type II error. [6]

(iii) Researcher B uses a sample bigger than 50 and a significance level of 5%. Explain whether the probability that Researcher B makes a Type II error is less than, equal to, or greater than your answer to part **(ii)**. [2]

(iv) Researcher C uses a sample of size 50 and a significance level bigger than 5%. Explain whether the probability that Researcher C makes a Type II error is less than, equal to, or greater than your answer to part **(ii)**. [2]

(v) State with a reason whether it is necessary to use the Central Limit Theorem at any point in this question. [2]

BLANK PAGE