OCR RECOGNISING ACHIEVEMENT

ADVANCED GCE UNIT **4733/01**

MATHEMATICS

Probability & Statistics 2

FRIDAY 12 JANUARY 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.

1 The random variable H has the distribution $\mathrm{N}(\mu, 5^2)$. It is given that $\mathrm{P}(H < 22) = 0.242$. Find the value of μ. [4]

2 A school has 900 pupils. For a survey, Jan obtains a list of all the pupils, numbered 1 to 900 in alphabetical order. She then selects a sample by the following method. Two fair dice, one red and one green, are thrown, and the number in the list of the first pupil in the sample is determined by the following table.

		Score on green dice					
		1	2	3	4	5	6
Score on red dice	1, 2 or 3	1	2	3	4	5	6
	4, 5 or 6	7	8	9	10	11	12

For example, if the scores on the red and green dice are 5 and 2 respectively, then the first member of the sample is the pupil numbered 8 in the list.

Starting with this first number, every 12th number on the list is then used, so that if the first pupil selected is numbered 8, the others will be numbered 20, 32, 44, … .

(i) State the size of the sample. [1]

(ii) Explain briefly whether the following statements are true.

(a) Each pupil in the school has an equal probability of being in the sample. [1]

(b) The pupils in the sample are selected independently of one another. [1]

(iii) Give a reason why the number of the first pupil in the sample should not be obtained simply by adding together the scores on the two dice. Justify your answer. [2]

3 A fair dice is thrown 90 times. Use an appropriate approximation to find the probability that the number 1 is obtained 14 or more times. [6]

4 A set of observations of a random variable W can be summarised as follows:

$$n = 14, \quad \Sigma w = 100.8, \quad \Sigma w^2 = 938.70.$$

(i) Calculate an unbiased estimate of the variance of W. [4]

(ii) The mean of 70 observations of W is denoted by $\overline{W}$. State the approximate distribution of $\overline{W}$, including unbiased estimate(s) of any parameter(s). [3]

5 On a particular night, the number of shooting stars seen per minute can be modelled by the distribution $\text{Po}(0.2)$.

(i) Find the probability that, in a given 6-minute period, fewer than 2 shooting stars are seen. [3]

(ii) Find the probability that, in 20 periods of 6 minutes each, the number of periods in which fewer than 2 shooting stars are seen is exactly 13. [3]

(iii) Use a suitable approximation to find the probability that, in a given 2-hour period, fewer than 30 shooting stars are seen. [6]

6 The continuous random variable X has the following probability density function:

$$\mathrm{f}(x) = \begin{cases} a + bx & 0 \leqslant x \leqslant 2, \\ 0 & \text{otherwise,} \end{cases}$$

where a and b are constants.

(i) Show that $2a + 2b = 1$. [3]

(ii) It is given that $\mathrm{E}(X) = \frac{11}{9}$. Use this information to find a second equation connecting a and b, and hence find the values of a and b. [6]

(iii) Determine whether the median of X is greater than, less than, or equal to $\mathrm{E}(X)$. [4]

7 A television company believes that the proportion of households that can receive Channel C is 0.35.

(i) In a random sample of 14 households it is found that 2 can receive Channel C. Test, at the 2.5% significance level, whether there is evidence that the proportion of households that can receive Channel C is less than 0.35. [7]

(ii) On another occasion the test is carried out again, with the same hypotheses and significance level as in part **(i)**, but using a new sample, of size n. It is found that no members of the sample can receive Channel C. Find the largest value of n for which the null hypothesis is not rejected. Show all relevant working. [4]

[Question 8 is printed overleaf.]

8 The quantity, X milligrams per litre, of silicon dioxide in a certain brand of mineral water is a random variable with distribution $\mathrm{N}(\mu,\ 5.6^2)$.

(i) A random sample of 80 observations of X has sample mean 100.7. Test, at the 1% significance level, the null hypothesis $\mathrm{H}_0 : \mu = 102$ against the alternative hypothesis $\mathrm{H}_1 : \mu \neq 102$. [5]

(ii) The test is redesigned so as to meet the following conditions.

- The hypotheses are $\mathrm{H}_0 : \mu = 102$ and $\mathrm{H}_1 : \mu < 102$.
- The significance level is 1%.
- The probability of making a Type II error when $\mu = 100$ is to be (approximately) 0.05.

The sample size is n, and the critical region is $\overline{X} < c$, where $\overline{X}$ denotes the sample mean.

(a) Show that n and c satisfy (approximately) the equation $102 - c = \dfrac{13.0256}{\sqrt{n}}$. [3]

(b) Find another equation satisfied by n and c. [2]

(c) Hence find the values of n and c. [4]