RECOGNISING ACHIEVEMENT

ADVANCED GCE **4733/01**

MATHEMATICS

Probability & Statistics 2

THURSDAY 12 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 The head teacher of a school asks for volunteers from among the pupils to take part in a survey on political interests.

(i) Explain why a sample consisting of all the volunteers is unlikely to give a true picture of the political interests of all pupils in the school. [2]

(ii) Describe a better method of obtaining the sample. [3]

2 The annual salaries of employees in a company have mean £30 000 and standard deviation £12 000.

(i) Assuming a normal distribution, calculate the probability that the salary of one randomly chosen employee lies between £20 000 and £24 000. [4]

(ii) The salary structure of the company is such that a small number of employees earn much higher salaries than the others. Explain what this suggests about the use of a normal distribution to model the data. [2]

3 In a factory the time, T minutes, taken by an employee to make a single item is a normally distributed random variable with mean 28.0. A new ventilation system is installed, after which the times taken to produce a random sample of 40 items are measured. The sample mean is 26.44 minutes and it is given that $\frac{\Sigma t^2}{40} - 26.44^2 = 37.05$. Test, at the 10% significance level, whether there is evidence of a change in the mean time taken to make an item. [8]

4 The random variable U has the distribution $N(\mu, \sigma^2)$, where the value of σ is known. A test is carried out of the null hypothesis $H_0 : \mu = 50$ against the alternative hypothesis $H_1 : \mu > 50$. The test is carried out at the 1% significance level and is based on a random sample of size 10.

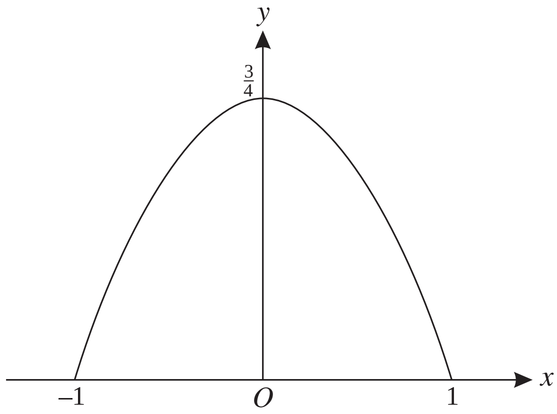
(i) The test is carried out once. The value of the sample mean is 53.0. The outcome of the test is that H_0 is not rejected. Show that $\sigma > 4.08$, correct to 3 significant figures. [4]

(ii) The test is carried out repeatedly. In each test the actual value of μ is 50. Find the probability that the first test to result in a Type I error is the fifth to be carried out. Give your answer correct to 2 significant figures. [3]

5 **(i)** A continuous random variable X has probability density function given by

$$f(x) = \begin{cases} \frac{3}{4}(1 - x^2) & -1 \leqslant x \leqslant 1, \\ 0 & \text{otherwise.} \end{cases}$$

The graph of $y = f(x)$ is shown in the diagram.



Calculate the value of $\text{Var}(X)$. [4]

(ii) A continuous random variable W has probability density function given by

$$g(x) = \begin{cases} k(9 - x^2) & -3 \leqslant x \leqslant 3, \\ 0 & \text{otherwise,} \end{cases}$$

where k is a constant.

(a) Sketch the graph of $y = g(x)$. [1]

(b) By comparing the graphs of $y = f(x)$ and $y = g(x)$, explain how you can tell without calculation that $9k < \frac{3}{4}$. [2]

(c) State with a reason, but without calculation, whether the standard deviation of W is greater than, equal to, or less than that of X. [2]

6 **(a)** On average I receive 19 e-mails per (8-hour) working day. Assuming that a Poisson distribution is a valid model, find the probability that in one randomly chosen hour I receive either 3 or 4 e-mails. [4]

(b) **(i)** State the conditions needed to use a Poisson distribution as an approximation to a binomial distribution. [2]

(ii) 108 people each throw a pair of fair six-sided dice. Use a Poisson approximation to find the probability that at least 4 people obtain a double six. [5]

[Questions 7 and 8 are printed overleaf.]

7 Wendy analyses the number of 'dropped catches' in international cricket matches. She finds that the mean number of dropped catches per day is 2. In a recent 5-day match she found that there was a total of c dropped catches. She tests, at the 5% significance level, whether the mean number of dropped catches per day has increased.

(i) State conditions needed for the number of dropped catches per day to be well modelled by a Poisson distribution. [2]

Assume now that these conditions hold.

(ii) Find the probability that the test results in a Type I error.

(iii) Given that $c = 14$, carry out the test.

[10]

8 A company sponsors a series of concerts. Surveys show that on average 40% of audience members know the name of the sponsor. As this figure is thought to be disappointingly low, the publicity material is redesigned.

(i) After the publicity material has been redesigned, a random sample of 12 audience members is obtained, and it is found that 9 members of this sample know the name of the sponsor. Test, at the 5% significance level, whether there is evidence that the proportion of audience members who know the name of the sponsor has increased. [7]

(ii) A more detailed 5% hypothesis test is carried out, based on a random sample of size 400. This test produces significant evidence that the proportion of audience members knowing the name of the sponsor has increased. Using an appropriate approximation, calculate the smallest possible number of audience members in the sample of 400 who know the name of the sponsor. [7]