**OCR**
RECOGNISING ACHIEVEMENT

**ADVANCED GCE**

**MATHEMATICS 4733**

Probability & Statistics 2

**Friday 15 January 2010**
**Afternoon**

**Duration:** 1 hour 30 minutes

*4733*

**INSTRUCTIONS TO CANDIDATES**

* Write your name clearly in capital letters, your Centre Number and Candidate Number in the spaces provided on the Answer Booklet.
* Use black ink. Pencil may be used for graphs and diagrams only.
* Read each question carefully and make sure that you know what you have to do before starting your answer.
* Answer **all** the questions.
* Do **not** write in the bar codes.
* 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.
* **You are reminded of the need for clear presentation in your answers.**
* The total number of marks for this paper is **72**.
* This document consists of **4** pages. Any blank pages are indicated.

**Turn over**

**1** The values of 5 independent observations from a population can be summarised by

$$\Sigma x = 75.8, \quad \Sigma x^2 = 1154.58.$$

Find unbiased estimates of the population mean and variance. **[4]**

**2** A college has 400 students. A journalist wants to carry out a survey about food preferences and she obtains a sample of 30 pupils from the college by the following method.

- Obtain a list of all the students.

- Number the students, with numbers running sequentially from 0 to 399.

- Select 30 random integers in the range 000 to 999 inclusive. If a random integer is in the range 0 to 399, then the student with that number is selected. If the number is greater than 399, then 400 is subtracted from the number (if necessary more than once) until an answer in the range 0 to 399 is selected, and the student with that number is selected.

  **(i)** Explain why this method is unsatisfactory. **[2]**

  **(ii)** Explain how it could be improved. **[1]**

**3** In a large town, 35% of the inhabitants have access to television channel $C$. A random sample of 60 inhabitants is obtained. Use a suitable approximation to find the probability that 18 or fewer inhabitants in the sample have access to channel $C$. **[6]**

**4** 80 randomly chosen people are asked to estimate a time interval of 60 seconds without using a watch or clock. The mean of the 80 estimates is 58.9 seconds. Previous evidence shows that the population standard deviation of such estimates is 5.0 seconds. Test, at the 5% significance level, whether there is evidence that people tend to underestimate the time interval. **[7]**

**5** The number of customers arriving at a store between 8.50 am and 9 am on Saturday mornings is a random variable which can be modelled by the distribution Po(11.0). Following a series of price cuts, on one particular Saturday morning 19 customers arrive between 8.50 am and 9 am. The store's management claims, first, that the mean number of customers has increased, and second, that this is due to the price cuts.

  **(i)** Test the first part of the claim, at the 5% significance level. **[7]**

  **(ii)** Comment on the second part of the claim. **[1]**

**6** The continuous random variable $X$ has the distribution $N(\mu, \sigma^2)$.

  **(i)** Each of the three following sets of probabilities is impossible. Give a reason in each case why the probabilities cannot both be correct. (You should not attempt to find $\mu$ or $\sigma$.)

    **(a)** $P(X > 50) = 0.7$ and $P(X < 50) = 0.2$ **[1]**

    **(b)** $P(X > 50) = 0.7$ and $P(X > 70) = 0.8$ **[1]**

    **(c)** $P(X > 50) = 0.3$ and $P(X < 70) = 0.3$ **[1]**

  **(ii)** Given that $P(X > 50) = 0.7$ and $P(X < 70) = 0.7$, find the values of $\mu$ and $\sigma$. **[4]**

**7** The continuous random variable $T$ is equally likely to take any value from 5.0 to 11.0 inclusive.

   **(i)** Sketch the graph of the probability density function of $T$. **[2]**

   **(ii)** Write down the value of $E(T)$ and find by integration the value of $Var(T)$. **[5]**

   **(iii)** A random sample of 48 observations of $T$ is obtained. Find the approximate probability that the mean of the sample is greater than 8.3, and explain why the answer is an approximation. **[6]**

**8** The random variable $R$ has the distribution $B(10, p)$. The null hypothesis $H_0: p = 0.7$ is to be tested against the alternative hypothesis $H_1: p < 0.7$, at a significance level of 5%.

   **(i)** Find the critical region for the test and the probability of making a Type I error. **[3]**

   **(ii)** Given that $p = 0.4$, find the probability that the test results in a Type II error. **[3]**

   **(iii)** Given that $p$ is equally likely to take the values 0.4 and 0.7, find the probability that the test results in a Type II error. **[2]**

**9** Buttercups in a meadow are distributed independently of one another and at a constant average incidence of 3 buttercups per square metre.

   **(i)** Find the probability that in 1 square metre there are more than 7 buttercups. **[2]**

   **(ii)** Find the probability that in 4 square metres there are either 13 or 14 buttercups. **[3]**

   **(iii)** Use a suitable approximation to find the probability that there are no more than 69 buttercups in 20 square metres. **[5]**

   **(iv) (a)** Without using an approximation, find an expression for the probability that in $m$ square metres there are at least 2 buttercups. **[2]**

       **(b)** It is given that the probability that there are at least 2 buttercups in $m$ square metres is 0.9. Using your answer to part **(a)**, show numerically that $m$ lies between 1.29 and 1.3. **[4]**