OCR
RECOGNISING ACHIEVEMENT

OXFORD CAMBRIDGE AND RSA EXAMINATIONS

Advanced Subsidiary General Certificate of Education
Advanced General Certificate of Education

MEI STRUCTURED MATHEMATICS 2615

Statistics 3

Thursday **12 JANUARY 2006** Afternoon 1 hour 20 minutes

Additional materials:
Answer booklet
Graph paper
MEI Examination Formulae and Tables (MF12)

TIME 1 hour 20 minutes

INSTRUCTIONS TO CANDIDATES

- Write your name, centre number and candidate number in the spaces provided on the answer booklet.
- Answer **all** the questions.
- 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 advised that an answer may receive no marks unless you show sufficient detail of the working to indicate that a correct method is being used.
- Final answers should be given to a degree of accuracy appropriate to the context.
- The total number of marks for this paper is 60.

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

Registered Charity 1066969
[Turn over

1 The random variable X has probability density function $\mathrm{f}(x)$ given by

$$\mathrm{f}(x) = \frac{k}{x^{k+1}}, \qquad x \geqslant 1, \quad k > 1.$$

(i) Sketch the graph of $\mathrm{f}(x)$. [3]

(ii) Obtain the cumulative distribution function of X. [3]

(iii) Find the median of X. [2]

(iv) Obtain the mean of X. [3]

(v) In a certain country, the distribution of incomes in units of \$10 000 is modelled by X with $k = 3$. Find

(*A*) the median income, [1]

(*B*) the mean income, [1]

(*C*) the proportion of incomes greater than the mean, [1]

(*D*) the proportion of incomes greater than \$100 000. [1]

2 Granulated sugar is sold in "1 kg" bags. The bags are filled by a machine which is set so that the weight, in grams, of sugar delivered to a bag is a Normally distributed random variable X with mean 1010 and standard deviation 8, independently for all bags.

(i) Find the probability that a bag of sugar is underweight (i.e. contains less than 1 kg). [2]

(ii) A supermarket sells bags of sugar in "catering packs" consisting of six randomly chosen "1 kg" bags. State the distribution of the weight of sugar in a catering pack. [3]

(iii) A catering pack is underweight if it contains less than 6 kg in total. Find the probability that a catering pack is underweight. Explain why this is less than the probability in part **(i)**. [3]

(iv) The supermarket also sells "2 kg" bags of sugar which are filled by a machine linked to the first machine in such a way that the weight, in grams, of sugar delivered to them is the random variable $1.99X$, independently for all bags. Find the probability that a "2 kg" bag is underweight. [3]

The machine that fills "1 kg" bags undergoes maintenance. On being started again, the weight (in grams) of sugar delivered to a bag is still a Normally distributed random variable with standard deviation 8, independently for all bags, but the mean might have changed.

(v) An inspector now selects 20 "1 kg" bags of sugar at random and finds that the average weight of sugar in them is 1004.2 g. Obtain a two-sided 95% confidence interval for the true mean weight of sugar in the large production run of bags from which the sample was taken. [4]

3 The production process at a certain factory needs the supply of a large volume of water on a continuous basis. The water is received by pipeline direct from a water company. The volume of water supplied varies somewhat from day to day due to amount of rainfall and other demands on the water company's resources. Engineers at the factory have determined that a mean daily volume of 34.8 units is required for efficient operation of the production process. If it is less than this, the process tends to work inefficiently; if it is more, excess water tends to be wasted.

The engineers measure the daily volume supplied on 8 occasions and find it to be as follows.

34.1 35.6 31.3 34.7 34.8 33.9 32.7 33.7

(i) State the appropriate null and alternative hypotheses for the usual t test that the engineers might use. [2]

(ii) State two conditions necessary for the correct use of this test. [2]

(iii) Carry out the test, using a 10% significance level. [7]

(iv) A statistician working for the water company uses the above data and correctly calculates a two-sided 95% confidence interval, based on the t distribution, for the true mean daily volume delivered as $(32.73, 34.97)$. State the conclusion that this would imply for the hypothesis test if it had been carried out at the 5% level. Briefly discuss this result in comparison with your result in part **(iii)**. [4]

4 Geneticists are investigating a situation where there are three types of offspring, T_1, T_2 and T_3. It is believed that an appropriate model for the situation is that the three types occur in the proportions $p^2 : 2pq : q^2$ respectively, where $p + q = 1$.

In a sample of size 80, it is found that there are 5 of type T_1, 38 of type T_2 and 37 of type T_3.

(i) In an initial stage of the modelling, it is assumed that the value of p is $\frac{1}{2}$. Test at the 5% level of significance whether it is reasonable to suppose that the model applies with $p = \frac{1}{2}$. [6]

(ii) The model is then refined by estimating p from the data. The estimate, $\hat{p}$, is given by

$$\hat{p} = \frac{n_1 + \frac{1}{2}n_2}{n_1 + n_2 + n_3}$$

where n_1, n_2 and n_3 are the observed numbers of types T_1, T_2 and T_3 respectively. Test at the 5% level of significance whether it is reasonable to suppose that this refined model applies. [7]

(iii) Discuss your conclusions. [2]

BLANK PAGE

2615 January 2006