OCR
RECOGNISING ACHIEVEMENT

ADVANCED GCE **4768/01**

MATHEMATICS (MEI)

Statistics 3

WEDNESDAY 21 MAY 2008 Afternoon

Time: 1 hour 30 minutes

Additional materials: Answer Booklet (8 pages)
Graph paper
MEI Examination Formulae and Tables (MF2)

INSTRUCTIONS TO CANDIDATES

- Write your name, 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.
- You are permitted to use a graphical calculator in this paper.
- Final answers should be given to a degree of accuracy appropriate to the context.

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 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.

This document consists of **4** printed pages.

1 **(a)** Sarah travels home from work each evening by bus; there is a bus every 20 minutes. The time at which Sarah arrives at the bus stop varies randomly in such a way that the probability density function of X, the length of time in minutes she has to wait for the next bus, is given by

$$f(x) = k(20 - x) \text{ for } 0 \leqslant x \leqslant 20, \text{ where } k \text{ is a constant.}$$

(i) Find k. Sketch the graph of $f(x)$ and use its shape to explain what can be deduced about how long Sarah has to wait. [5]

(ii) Find the cumulative distribution function of X and hence, or otherwise, find the probability that Sarah has to wait more than 10 minutes for the bus. [4]

(iii) Find the median length of time that Sarah has to wait. [3]

(b) **(i)** Define the term 'simple random sample'. [2]

(ii) Explain briefly how to carry out cluster sampling. [3]

(iii) A researcher wishes to investigate the attitudes of secondary school pupils to pollution. Explain why he might prefer to collect his data using a cluster sample rather than a simple random sample. [2]

2 An electronics company purchases two types of resistor from a manufacturer. The resistances of the resistors (in ohms) are known to be Normally distributed. Type A have a mean of 100 ohms and standard deviation of 1.9 ohms. Type B have a mean of 50 ohms and standard deviation of 1.3 ohms.

(i) Find the probability that the resistance of a randomly chosen resistor of type A is less than 103 ohms. [3]

(ii) Three resistors of type A are chosen at random. Find the probability that their total resistance is more than 306 ohms. [3]

(iii) One resistor of type A and one resistor of type B are chosen at random. Find the probability that their total resistance is more than 147 ohms. [3]

(iv) Find the probability that the total resistance of two randomly chosen type B resistors is within 3 ohms of one randomly chosen type A resistor. [5]

(v) The manufacturer now offers type C resistors which are specified as having a mean resistance of 300 ohms. The resistances of a random sample of 100 resistors from the first batch supplied have sample mean 302.3 ohms and sample standard deviation 3.7 ohms. Find a 95% confidence interval for the true mean resistance of the resistors in the batch. Hence explain whether the batch appears to be as specified. [4]

3 **(a)** A tea grower is testing two types of plant for the weight of tea they produce. A trial is set up in which each type of plant is grown at each of 8 sites. The total weight, in grams, of tea leaves harvested from each plant is measured and shown below.

Site	A	B	C	D	E	F	G	H
Type I	225.2	268.9	303.6	244.1	230.6	202.7	242.1	247.5
Type II	215.2	242.1	260.9	241.7	245.5	204.7	225.8	236.0

(i) The grower intends to perform a *t* test to examine whether there is any difference in the mean yield of the two types of plant. State the hypotheses he should use and also any necessary assumption. [3]

(ii) Carry out the test using a 5% significance level. [7]

(b) The tea grower deals with many types of tea and employs tasters to rate them. The tasters do this by giving each tea a score out of 100. The tea grower wishes to compare the scores given by two of the tasters. Their scores for a random selection of 10 teas are as follows.

Tea	Q	R	S	T	U	V	W	X	Y	Z
Taster 1	69	79	85	63	81	65	85	86	89	77
Taster 2	74	75	99	66	75	64	96	94	96	86

Use a Wilcoxon test to examine, at the 5% level of significance, whether it appears that, on the whole, the scores given to teas by these two tasters differ. [8]

4 **(a)** A researcher is investigating the feeding habits of bees. She sets up a feeding station some distance from a beehive and, over a long period of time, records the numbers of bees arriving each minute. For a random sample of 100 one-minute intervals she obtains the following results.

Number of bees	0	1	2	3	4	5	6	7	$\geqslant 8$
Number of intervals	6	16	19	18	17	14	6	4	0

(i) Show that the sample mean is 3.1 and find the sample variance. Do these values support the possibility of a Poisson model for the number of bees arriving each minute? Explain your answer. [3]

(ii) Use the mean in part **(i)** to carry out a test of the goodness of fit of a Poisson model to the data. [10]

(b) The researcher notes the length of time, in minutes, that each bee spends at the feeding station. The times spent are assumed to be Normally distributed. For a random sample of 10 bees, the mean is found to be 1.465 minutes and the standard deviation is 0.3288 minutes. Find a 95% confidence interval for the overall mean time. [4]