OCR
RECOGNISING ACHIEVEMENT

OXFORD CAMBRIDGE AND RSA EXAMINATIONS

Advanced Subsidiary General Certificate of Education
Advanced General Certificate of Education

MEI STRUCTURED MATHEMATICS 4768

Statistics 3

Thursday **12 JANUARY 2006** Afternoon 1 hour 30 minutes

Additional materials:
8 page answer booklet
Graph paper
MEI Examination Formulae and Tables (MF2)

TIME 1 hour 30 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.
- 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.
- 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.
- The total number of marks for this paper is 72.

This question paper consists of 5 printed pages and 3 blank pages.

HN/4
© OCR 2006 [R/102/2658] Registered Charity 1066969

1 A railway company is investigating operations at a junction where delays often occur. Delays (in minutes) are modelled by the random variable T with the following cumulative distribution function.

$$F(t) = \begin{cases} 0 & t \leqslant 0 \\ 1 - e^{-\frac{1}{3}t} & t > 0 \end{cases}$$

(i) Find the median delay and the 90th percentile delay. [5]

(ii) Derive the probability density function of T. Hence use calculus to find the mean delay. [5]

(iii) Find the probability that a delay lasts longer than the mean delay. [2]

You are given that the variance of T is 9.

(iv) Let $\overline{T}$ denote the mean of a random sample of 30 delays. Write down an approximation to the distribution of $\overline{T}$. [3]

(v) A random sample of 30 delays is found to have mean 4.2 minutes. Does this cast any doubt on the modelling? [3]

2 Geoffrey is a university lecturer. He has to prepare five questions for an examination. He knows by experience that it takes about 3 hours to prepare a question, and he models the time (in minutes) taken to prepare one by the Normally distributed random variable X with mean 180 and standard deviation 12, independently for all questions.

(i) One morning, Geoffrey has a gap of 2 hours 50 minutes (170 minutes) between other activities. Find the probability that he can prepare a question in this time. [3]

(ii) One weekend, Geoffrey can devote 14 hours to preparing the complete examination paper. Find the probability that he can prepare all five questions in this time. [3]

A colleague, Helen, has to check the questions.

(iii) She models the time (in minutes) to check a question by the Normally distributed random variable Y with mean 50 and standard deviation 6, independently for all questions and independently of X. Find the probability that the total time for Geoffrey to prepare a question and Helen to check it exceeds 4 hours. [3]

(iv) When working under pressure of deadlines, Helen models the time to check a question in a different way. She uses the Normally distributed random variable $\frac{1}{4}X$, where X is as above. Find the length of time, as given by this model, which Helen needs to ensure that, with probability 0.9, she has time to check a question. [4]

Ian, an educational researcher, suggests that a better model for the time taken to prepare a question would be a constant k representing "thinking time" plus a random variable T representing the time required to write the question itself, independently for all questions.

(v) Taking k as 45 and T as Normally distributed with mean 120 and standard deviation 10 (all units are minutes), find the probability according to Ian's model that a question can be prepared in less than 2 hours 30 minutes. [2]

Juliet, an administrator, proposes that the examination should be reduced in time and shorter questions should be used.

(vi) Juliet suggests that Ian's model should be used for the time taken to prepare such shorter questions but with $k = 30$ and T replaced by $\frac{3}{5}T$. Find the probability as given by this model that a question can be prepared in less than $1\frac{3}{4}$ hours. [3]

[Turn over

3 A production line has two machines, A and B, for delivering liquid soap into bottles. Each machine is set to deliver a nominal amount of 250 ml, but it is not expected that they will work to a high level of accuracy. In particular, it is known that the ambient temperature affects the rate of flow of the liquid and leads to variation in the amounts delivered.

The operators think that machine B tends to deliver a somewhat greater amount than machine A, no matter what the ambient temperature. This is being investigated by an experiment. A random sample of 10 results from the experiment is shown below. Each column of data is for a different ambient temperature.

Ambient temperature	T_1	T_2	T_3	T_4	T_5	T_6	T_7	T_8	T_9	T_{10}
Amount delivered by machine A	246.2	251.6	252.0	246.6	258.4	251.0	247.5	247.1	248.1	253.4
Amount delivered by machine B	248.3	252.6	252.8	247.2	258.8	250.0	247.2	247.9	249.0	254.5

(a) Use an appropriate t test to examine, at the 5% level of significance, whether the mean amount delivered by machine B may be taken as being greater than that delivered by machine A, stating carefully your null and alternative hypotheses and the required distributional assumption. [11]

(b) Using the data for machine A in the table above, provide a two-sided 95% confidence interval for the mean amount delivered by this machine, stating the required distributional assumption. Explain whether you would conclude that the machine appears to be working correctly in terms of the nominal amount as set. [7]

4 Quality control inspectors in a factory are investigating the lengths of glass tubes that will be used to make laboratory equipment.

(i) Data on the observed lengths of a random sample of 200 glass tubes from one batch are available in the form of a frequency distribution as follows.

Length x (mm)	Observed frequency
$x \leqslant 298$	1
$298 < x \leqslant 300$	30
$300 < x \leqslant 301$	62
$301 < x \leqslant 302$	70
$302 < x \leqslant 304$	34
$x > 304$	3

The sample mean and standard deviation are 301.08 and 1.2655 respectively.

The corresponding expected frequencies for the Normal distribution with parameters estimated by the sample statistics are

Length x (mm)	Expected frequency
$x \leqslant 298$	1.49
$298 < x \leqslant 300$	37.85
$300 < x \leqslant 301$	55.62
$301 < x \leqslant 302$	58.32
$302 < x \leqslant 304$	44.62
$x > 304$	2.10

Examine the goodness of fit of a Normal distribution, using a 5% significance level. [7]

(ii) It is thought that the lengths of tubes in another batch have an underlying distribution similar to that for the batch in part **(i)** but possibly with different location and dispersion parameters. A random sample of 10 tubes from this batch gives the following lengths (in mm).

301.3 301.4 299.6 302.2 300.3 303.2 302.6 301.8 300.9 300.8

(*A*) Discuss briefly whether it would be appropriate to use a t test to examine a hypothesis about the population mean length for this batch. [2]

(*B*) Use a Wilcoxon test to examine at the 10% significance level whether the population median length for this batch is 301 mm. [9]