OXFORD CAMBRIDGE AND RSA EXAMINATIONS

Advanced Subsidiary General Certificate of Education
Advanced General Certificate of Education

MEI STRUCTURED MATHEMATICS **2616**

Statistics 4

Friday **17 JANUARY 2003** 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 any **three** questions.
- You are permitted to use a graphical calculator in this paper.

INFORMATION FOR CANDIDATES

- The allocation 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 4 printed pages.

1 The random variable X has the rectangular distribution on $(-\theta, \theta)$, so that its probability density function is

$$f(x)=\frac{1}{2\theta} \qquad -\theta \leqslant x \leqslant \theta$$

where $\theta > 0$. Data consisting of a random sample of size n from this distribution are available. The data are arranged in ascending order and are then represented by $U_1, U_2, \ldots, U_n$ where $U_1 \leqslant U_2 \leqslant \ldots \leqslant U_n$. You are *given* that the range $R = U_n - U_1$ has probability density function

$$g(r)=\frac{n(n-1)}{(2\theta)^n} r^{n-2}(2\theta-r) \qquad 0 \leqslant r \leqslant 2\theta.$$

(i) Show that the expected value of R is $\dfrac{n-1}{n+1}(2\theta)$. [6]

(ii) Deduce whether or not $\frac{1}{2}R$ is an unbiased estimator of θ. [3]

(iii) Given that $E(R^2)=\dfrac{n(n-1)}{(n+1)(n+2)}(2\theta)^2$, show that

$$\text{Var}(R)=\frac{8(n-1)\theta^2}{(n+1)^2(n+2)}.$$ [4]

(iv) Hence find the mean square error of $\frac{1}{2}R$ as an estimator of θ. [5]

(v) Use the results of parts **(ii)** and **(iv)** to comment briefly on the use of $\frac{1}{2}R$ as an estimator of θ for cases where the sample is large. [2]

2 A mobile phone company is analysing calls made by customers on different tariffs.

(a) The lengths of calls for a random sample of 60 customers on tariff A have mean 3.62 minutes and standard deviation 1.12 minutes. The lengths of calls for a random sample of 80 customers on tariff B have mean 4.19 minutes and standard deviation 1.85 minutes. [Standard deviations are as defined using divisor $n - 1$.] Test at the 4% level of significance whether there is evidence that the true mean lengths of calls for customers on tariffs A and B differ, stating carefully the null and alternative hypotheses you are testing. Provide a two-sided 99% confidence interval for the true mean difference. [12]

(b) At an initial stage of investigating tariffs C and D, the company records the lengths of random samples of 8 calls by customers on tariff C and 7 calls by customers on tariff D. The lengths, in minutes, are as follows.

Tariff C	0.84	3.92	4.74	0.39	5.18	4.36	1.79	3.94
Tariff D	0.65	1.36	2.83	0.26	0.34	3.65	1.67	

Use an appropriate non-parametric test, at the 10% level of significance, to examine whether there are, on the whole, any differences between lengths of calls on these two tariffs. [8]

3 A company purchases a chemical from a supplier. It is specified that the chemical should contain no more than 7.5% of impurity. To investigate this, the company arranges that a random sample of deliveries are checked both by the supplier and by the company itself. The percentages of impurity as found by the supplier and the company are as follows.

Delivery	A	B	C	D	E	F	G	H	I	J
Supplier's determination	7.7	9.4	6.6	5.5	8.1	4.9	5.9	6.9	9.0	7.4
Company's determination	7.5	9.1	6.8	5.4	8.0	4.7	5.6	6.9	9.3	7.7

(a) Use an appropriate t test to examine, at the 5% level of significance, whether the mean determinations of percentage of impurity by the supplier and the company may be assumed equal, stating carefully your null and alternative hypotheses and the required distributional assumption. [11]

(b) The company decides to rely on the supplier's determinations. Using the data for the supplier in the table above, provide a one-sided 95% confidence interval giving an upper confidence bound for the mean percentage of impurity, stating the required distributional assumption. What can you conclude in respect of the specification that there should be no more than 7.5% of impurity? [9]

4 A courier delivery company is investigating the time intervals, x, between successive deliveries to a particular busy location. The table below gives frequencies for 100 randomly chosen time intervals. It also gives probabilities for the time intervals using the exponential distribution with probability density function $\lambda e^{-\lambda x}$ as a model for the underlying random variable X, where the value of the parameter λ is estimated from the data as 0.48.

Time interval x (hours)	Frequency	Probability
$0 < x \leqslant 1$	34	0.3812
$1 < x \leqslant 1.5$	12	0.1320
$1.5 < x \leqslant 2$	10	0.1039
$2 < x \leqslant 3$	16	0.1460
$3 < x \leqslant 6$	20	0.1808
$x > 6$	8	0.0561

(i) Verify that the entry 0.1320 for $P(1 < X \leqslant 1.5)$ is correct. [3]

(ii) Examine the fit of this model to the data. [8]

Some time later, a further small survey gives the following random sample of nine time intervals, in hours, between successive deliveries.

1.20 3.25 0.23 6.16 0.84 2.50 8.45 4.32 2.86

(iii) The median time interval given by the model used above is 1.444 hours. Test at the 10% level of significance whether this median may be assumed still to apply. [9]

Mark Scheme

Question			Part Marks	Notes	Question Total
1	(i)	$\text{E}(R) = \int_0^{2\theta} r g(r)\,dr$	M1		
		$= \frac{n(n-1)}{(2\theta)^n}\int_0^{2\theta} r.r^{n-2}(2\theta - r)\,dr$	M1		
		$= \frac{n(n-1)}{(2\theta)^n}\int_0^{2\theta}\left(2\theta r^{n-1} - r^n\right)dr$	1		
		$= \frac{n(n-1)}{(2\theta)^n}\left[2\theta\frac{r^n}{n} - \frac{r^{n+1}}{n+1}\right]_0^{2\theta}$	1		
		$= \frac{n(n-1)}{(2\theta)^n}\left\{\frac{(2\theta)^{n+1}}{n} - \frac{(2\theta)^{n+1}}{n+1}\right\}$			
		$= n(n-1)(2\theta)\left\{\frac{1}{n} - \frac{1}{n+1}\right\}$			
		$= \frac{n-1}{n+1}(2\theta)$	2	Divisible, for algebra. Beware printed answer	6
	(ii)	Consider $\text{E}\left(\frac{1}{2}R\right)$	M1		
		$= \frac{1}{2}\text{E}(R) = \frac{n-1}{n+1}\theta$	1		
		$\neq \theta$, $\therefore$ biased	1		3
	(iii)	$\text{Var}(R) = \text{E}(R^2) - \{\text{E}(R)\}^2$	M1		
		$= \frac{n(n-1)}{(n+1)(n+2)}(2\theta)^2 - \left(\frac{n-1}{n+1}\right)^2(2\theta)^2$			
		$= \frac{4\theta^2(n-1)}{n+1}\left\{\frac{n}{n+2} - \frac{n-1}{n+1}\right\}$			
		$= \frac{4\theta^2(n-1)}{n+1}\cdot\frac{n^2+n-n^2-2n+n+2}{(n+1)(n+2)}$			
		$= \frac{8\theta^2(n-1)}{(n+1)^2(n+2)}$	3	Divisible, for algebra. Beware printed answer	4
	(iv)	$\text{MSE}\left(\frac{1}{2}R\right) = \text{Var}\left(\frac{1}{2}R\right) + \left\{\text{bias}\left(\frac{1}{2}R\right)\right\}^2$	M1		
		We have $\text{Var}\left(\frac{1}{2}R\right) = \frac{2\theta^2(n-1)}{(n+1)^2(n+2)}$	1		
		Bias $\left(\frac{1}{2}R\right) = \frac{n-1}{n+1}\theta - \theta$			
		$= \frac{n\theta - \theta - n\theta - \theta}{n+1} = -\frac{2\theta}{n+1}$	1		
		$\therefore \text{MSE} = \frac{2\theta^2}{(n+1)^2}\left(\frac{n-1}{n+2} + 2\right)$			
		$= \frac{2\theta^2}{(n+1)^2}\frac{n-1+2n+4}{n+2} = \frac{6\theta^2}{(n+1)(n+2)}$	2	Divisible	5
	(v)	In large samples, $\frac{1}{2}R$ is 'nearly unbiased' and has 'small' MSE	E1, E1		2

Question		Part Marks	Notes	Question Total
2 (a)	MUST be N(0, 1) test and CI for comparing means.			
	$H_0 : \mu_A = \mu_B$	1	If **both** correct. Do NOT allow $\overline{X}_A = \overline{X}_B$ or similar. Allow verbal statement.	
	$H_1 : \mu_A \neq \mu_B$	1	If μ_A, μ_B are adequately defined in words.	
	Test statistic is $\frac{3.62 - 4.19}{\sqrt{\frac{(1.12)^2}{60} + \frac{(1.85)^2}{80}}}$	M1		
	$= \frac{-0.57}{\sqrt{0.06368(79...)} = 0.2524} = -2.26(-2.2586)$	A1		
	Refer to N(0, 1)	1	No FT if wrong	
	4% critical point is 2.054	1	No FT if wrong	
	Significant	1		
	Seems mean lengths of calls differ	1		
	CI is given by			
	−0.57	M1		
	± 2.576	B1		
	× 0.2524	M1		
	= −0.57 ± 0.65(01…) = (−1.22, 0.08)	A1		12
(b)	MUST be Wilcoxon rank-sum test (or Mann-Whitney form thereof)			
	Use of ranks	M1		
	Ranks are C: 5 11 14 3 15 13 8 12 D: 4 6 9 1 2 10 7	A1		
	Rank sum (for D) is 39 (Mann-Whitney is 11)	1		
	Refer to tables of Wilcoxon (or M-W) statistic	1		
	Lower 5% tail is needed	1		
	Value for (7, 8) is 41 (or 13 if M-W used)	1		
	Result is significant	1		
	Seems on the whole there are differences in call lengths	1		8

Question		Part Marks	Notes	Question Total
3 (a)	$H_0 : \mu_D = 0$ (or $\mu_S = \mu_C$) $H_1 : \mu_D \neq 0$ (or $\mu_S \neq \mu_C$)	1	Do not allow $\bar{X}_S = \bar{X}_C$ or similar	
	Where μ_D is 'mean for S – mean for C'	1	For adequate verbal definition	
	Normality of differences is required	1		
	MUST be paired comparison t test Differences are 0.2 0.3 −0.2 0.1 0.1 0.2 0.3 0 −0.3 −0.3			
	$\bar{d} = 0.04$ $s_{n-1}^2 = 0.053\dot{7}$, $s_{n-1} = 0.2319$	A1	Accept $s_n^2 = 0.0484$, $s_n = 0.22$ ONLY if correctly used in sequel	
	Test statistic is $\frac{0.04-0}{\frac{0.2319}{\sqrt{10}}} = 0.5454(5)$	M1 M1 A1		
	Refer to t_9	1	May be awarded even if test statistic is wrong. No FT if wrong	
	Dt 5% pt is 2.262	1	No FT if wrong	
	Not significant	1		
	Seems mean determinations by supplier and company are the same	1		11
(b)	We now require Normality for the supplier's determinations	1		
	For the supplier $\bar{x} = 7.14$ $s_{n-1}^2 = 2.162\dot{6}$, $s_{n-1} = 1.4706$	A1	Accept $s_n^2 = 1.9464$, $s_n = 1.3951$ ONLY if correctly used in sequel	
	One-sided CI is given by 7.14	M1		
	+	M1		
	1.833	B1	t_9 (upper 5%)	
	$\times \frac{1.4706}{\sqrt{10}}$	M1		
	$= 7.14 + 0.85(24)$ $= 7.99$	A1	Depends on all 4 preceding marks. FT if wrong	
	The upper confidence bound exceeds the specified upper limit of 7.5 - some evidence that the specification is not being met	E2		9

Question		Part Marks	Notes	Question Total
4 (i)	$f(x) = \lambda e^{-\lambda x}$ with $\lambda = 0.48$			
	$P(1 < X \leq 1.5) = \int_1^{1.5} 0.48e^{-0.48x}dx$	M1		
	$= \left[-e^{-0.48x}\right]_1^{1.5} = e^{-0.48} - e^{-0.72}$	1		
	$= 0.6187(8) - 0.4867(5) = 0.1320(3)$	1		3
(ii)	o_i e_i	M1 A1	For e_i (ZERO, but FT, if not to this level of accuracy).	
	34 38.12			
	12 13.20 $X^2 = 0.4453 + 0.1091$			
	10 10.39 $+ 0.0146 + 0.1342 + 0.2039$			
	16 14.60 $+ 1.0182 = 1.92(536)$	M1 A1	For X^2	
	20 18.08			
	8 5.61			
	Refer to χ_4^2	2	Allow 1 for χ_5^2, but no FT	
	Not significant at any sensible significance level	1		
	Seems model fits OK	1		8
(iii)	Data / median 1.444 / difference / rank of |difference|	M1	For differences. ZERO in this section if differences not used	
	1.20 −0.244 1			
	3.25 1.806 6	M1 A1	For ranks. FT if ranks wrong	
	0.23 −1.214 4			
	6.16 4.716 8			
	0.84 −0.604 2			
	2.50 1.056 3			
	8.45 7.006 9			
	4.32 2.876 7			
	2.86 1.416 5			
	$T = 1 + 2 + 4 = 7$ (or $3 + 5 + 6 + 7 + 8 + 9 = 38$)	1		
	Refer to tables of Wilcoxon single sample (/ paired) statistic	1		
	Lower (or upper if 38 used) 5% tail is needed.	1		
	Value for $n = 9$ is 8 (or 37 if 38 used)	1		
	Result is significant	1		
	Seems median is no longer 1.444	1		9

Examiner’s Report

2616 Statistics 4

General Comments

There were 36 candidates, from 9 centres; still small numbers, but a welcome improvement form the very low entry in January 2002. Most of the work was of a reasonable standard; there were some distinctly good scripts; and a few weak ones.

Comments on Individual Questions

Question 1 (estimation, bias, mean square error)

This estimation theory question was based on the range of a random sample from a rectangular (continuous uniform) distribution. As usual, it was the least popular question on the paper. Also as usual, some candidates sailed through easily and efficiently, though others came adrift at various points. The opening expected value, which required only integration of a simple polynomial, was usually correctly achieved. It was surprising that one or two candidates could not immediately relate this to the request concerning unbiasedness in part (ii). In part (iii), most candidates worked carefully through to find the variance, but there were some appalling dishonest fudges of the given answer. Candidates really must be told **not to do this.** There is no disgrace in not being able to obtain a given answer but then going on to use it. The dishonesty, is to get to an incorrect expression that is manifestly nothing like the given answer and then just assume that it equals it. Given answers are there to be *helpful*, both as a check on work and so that candidates who cannot derive them may nevertheless use them in the remainder of the question.

The next part was to obtain a mean square error. A few candidates did not know what to do here. Most, however, used the usual formula with the bias and variance already obtained, generally successfully. A few, in a sense rather pleasingly, worked with the basic definition of MSE as E[(statistic – parameter)^2], which was not a difficult route in this case but often can be. The answer here is $6\theta^2/\{(n+1)(n+2)\}$.

The closing commentary in respect of large samples was somewhat disappointing, though many candidates had some worthwhile ideas. Discussion was expected to the effect that the estimator would have small MSE and would be nearly unbiased.

Question 2 (Normal test for comparing means, and confidence interval; Wilcoxon rank sum test)

This question was often done well.

In part (a), the Normal test was perhaps the first time that the candidates have ever been asked to work to a 4% significance level, but in most cases this did not distract them from proceeding correctly through the analysis. [Value of test statistic is –2.26, critical point is 2.054.] The confidence interval, at the more conventional 99% level [critical point 2.576], was also usually correctly done [answer (–1.22, 0.08)].

A few candidates decided to work this part as a t test and confidence interval, usually with 138 degrees of freedom. Ultimately this is not correct, and this is not really ameliorated by arguments concerning large-sample approximations. These *are* large samples, undoubtedly large enough for the central limit theorem to apply, all of which leads to the *Normal* test and confidence interval, taking the sample standard deviations as population values. Use of t would further require the population variances to be assumed equal, and there is no evidence that this is so (indeed, if an F test were used, there is very strong evidence that they are not. This is, of course, not in the Statistics 4 syllabus, but the point is that *nothing* is said in the question to suggest equality of the population variances).

Statements of the null and alternative hypotheses were often full and correct but not always so. Care must be taken to ensure that *population* parameters are referred to. This comment also applies to part (a) of question 3.

In part (b), the Wilcoxon test was usually done well [value of test statistic is 39, critical point is 41]. Some candidates, despite correctly finding the value of the test statistic and the critical point, erroneously thought that this meant that the result is not significant; this error has been almost completely absent in previous years, and it is disturbing to see it creeping in. Some candidates preferred the Mann-Whitney formulation of the test, which is of course entirely acceptable, either by directly calculating the Mann-Whitney statistic in the first place or by calculating the Wilcoxon form and then adjusting it (which, despite being acceptable, seems bizarre!).

Question 3 (paired *t* test; one-sided confidence interval for the mean based on one of the paired samples)

This question also was often done well.

The pairing in part (a) was usually successfully worked with [value of test statistic is 0.545, refer to t_9, critical point 2.262]. A few candidates, however, left out the zero difference for delivery H; such an approach is appropriate for some non-parametric tests but is a serious error for the *t* procedure. Concerning the required distributional assumption, it is appropriate to quote exactly from last January's report: "Candidates knew that the distributional result required for the *t* test had something to do with Normality, but as usual many were reluctant to state the key points that this referred to the *underlying distribution* of the *differences*. At the level of Statistics 4, candidates are expected to be completely secure about such matters".

In part (b), one of the paired samples was used to construct an upper confidence bound for the underlying population mean. Most candidates were successful here [critical point to be used, again from t_9 but not for the same reason, is 1.833; confidence bound is 7.99]. A few candidates continued to use results from part (a); in some cases this appeared to have been due to carelessness in reading the question, but in others it might have been a deeper misunderstanding. The interpretation at the end was somewhat disappointing. Several candidates proceeded to a general interpretation of a one-sided confidence interval, not noticing that this was not quite what the question asked for. It was an interpretation in the context of meeting the specification that was sought. In fact the upper confidence bound (7.99) exceeded (by some way) the specification limit (7.5), thus giving an interpretation of, in a sense, some evidence of the specification not being met. It was some sort of discussion on these lines that was sought.

Question 4 (chi-squared goodness of fit test based on exponential distribution; Wilcoxon single-sample test for a median)

Parts (i) and (ii) concerned the exponential distribution and the goodness of fit test. Part (i) was a simple exercise in verifying a probability using the pdf (some candidates did it via the cumulative distribution function, which is of course acceptable). Often this was found trivial, but several candidates were not fully convincing – or, sadly, in some cases utterly unconvincing, in their derivation of the given answer. Just stating that it is 0.1320 (the given answer) is scarcely enough! Stating that it is 0.13203 indicates that it has been genuinely worked out, and giving intermediate numerical values for $e^{-0.48}$ and $e^{-0.72}$ is safe. Surprisingly, there were a few candidates who did not know what to do at all in this part.

Nearly all candidates then proceeded correctly to the chi-squared test (though there were a few bizarre other approaches), correctly calculating the value of the test statistic as 1.92. This should be referred to χ^2_4; a substantial minority incorrectly used χ^2_5. The result is not significant at any sensible significance level, so it seems the model fits satisfactorily.

In part (iii), the test for the median based on the ranks of absolute values of differences was usually well done [value of test statistic is 7, critical point is 8, result is significant]. Some candidates, however, tried to use a form of *t* test, which is wholly wrong here.