OXFORD CAMBRIDGE AND RSA EXAMINATIONS

Advanced Subsidiary General Certificate of Education
Advanced General Certificate of Education

MEI STRUCTURED MATHEMATICS 2616

Statistics 4

Monday **19 JANUARY 2004** Morning 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.

[Turn over

1 The random variable X has the Poisson distribution with parameter λ. It is desired to estimate $\theta = \lambda^2$. The estimator $\hat{\theta} = X^2 - X$ is proposed, where X now denotes a single observation from the distribution.

(i) Show that $\hat{\theta}$ is an unbiased estimator of θ. [7]

(ii) You are given that the variance of $\hat{\theta}$ is $2\lambda^2(2\lambda + 1)$ and that, in this situation, a theoretical minimum for the variance of unbiased estimators of θ is $4\theta^{3/2}$. Use this information to find the efficiency of $\hat{\theta}$ by comparing its variance with the theoretical minimum. Show that this efficiency approaches 1 as λ becomes large. [5]

(iii) Show that $\dfrac{(x^2 - x)^2}{x!} = \dfrac{x(x-1)}{(x-2)!}$.

Hence show that

$$\sum_{x=0}^{\infty} (x^2 - x)^2 \frac{e^{-\lambda}\lambda^x}{x!} = \lambda^2 \sum_{x=2}^{\infty} x(x-1) \frac{e^{-\lambda}\lambda^{x-2}}{(x-2)!}$$

and, by substituting $x = y + 2$ in the right-hand side,

$$E(\hat{\theta}^2) = \lambda^2 E\big[(Y+1)(Y+2)\big] \qquad \text{where } Y \sim \text{Poisson}(\lambda).$$

Hence confirm that

$$\text{Var}(\hat{\theta}) = 2\lambda^2(2\lambda + 1).$$ [8]

2 A supermarket has been experiencing difficulty with the oven in its fresh bread department, and managers are checking it. The recipe for a particular type of bread specifies that the oven should operate at 230 °C, though some variation due to environmental factors can be tolerated. The operating temperature is measured at 12 random times and is found to be as follows (in °C).

227 214 218 206 228 231 226 202 236 241 207 223

(i) Use a Wilcoxon test to examine at the 5% significance level whether it is reasonable to assume that the median operating temperature is 230 °C. [9]

(ii) Public health inspectors require a one-sided $97\frac{1}{2}\%$ confidence interval giving an upper confidence bound for the mean operating temperature. Calculate this interval, and interpret it in terms of the specification. [9]

(iii) State the assumption necessary for the analysis in part **(ii)** that is not needed for that in part **(i)**. [2]

3 A company has two factories, A and B. Each factory discharges effluent. Inspections have to be made of the percentage (by volume) of a particular chemical in the effluent.

Determinations of this percentage are made at 10 randomly chosen times at factory A and also at 10 randomly chosen times at factory B. They are as follows.

Factory A:	4.15	4.29	3.87	4.19	4.51	4.33	4.23	4.48	4.27	4.14
Factory B:	4.56	4.46	4.60	4.58	4.39	4.27	4.04	4.53	4.68	4.26

As part of the inspection, it is necessary to carry out a t test to examine whether, on the whole, the percentage at factory A is the same as that at factory B.

(i) State the null and alternative hypotheses and the required assumptions for this t test. [4]

(ii) Carry out the test, at the 5% significance level. [10]

(iii) Provide a two-sided 99% confidence interval for the mean difference between the percentages. Interpret this interval. [6]

Turn over for Question 4

4 An insurance broker is investigating whether the amounts of money paid out in respect of claims for personal accidents are related to the circumstances of the accidents.

Payouts are classified as high, medium or low. Circumstances of accidents are classified as work-related, travel-related, home-related or other. Data for a random sample of 200 claims are shown in the table.

		Circumstances of accident			
		Work-related	Travel-related	Home-related	Other
Payout	High	40	7	26	15
	Medium	22	9	4	17
	Low	14	10	18	18

(i) State the null and alternative hypotheses under examination in the usual χ^2 test applied to this contingency table. [2]

(ii) Carry out the test, at the 5% significance level. [12]

(iii) Discuss your conclusions. [6]

Mark Scheme

Q.1	(i)	Consider $\mathrm{E}(\hat{\theta}) = \mathrm{E}(X^2 - X) = \mathrm{E}(X^2) - \mathrm{E}(X)$	M1	
		We have $\mathrm{E}(X) = \lambda$	1	
		and $\mathrm{E}(X^2) = \mathrm{var}(X) + \{\mathrm{E}(X)\}^2$	M1	
		$= \lambda + \{\lambda\}^2$	1, 1	
		$\therefore \mathrm{E}(\hat{\theta}) = \lambda + \lambda^2 - \lambda = \lambda^2$	1	
		i.e. unbiased	1	7
		ALITER: Via $\mathrm{E}(X) = \sqrt{\theta}$ and $\mathrm{var}(X) = \sqrt{\theta} \to \mathrm{E}(\hat{\theta}) = \sqrt{\theta} + (\sqrt{\theta})^2 - \sqrt{\theta}$		
	(ii)	Efficiency = <u>inverse</u>	M1	
		<u>ratio of variances</u>	M1	
		$\frac{4\theta^{\frac{3}{2}}}{2\lambda^2(2\lambda+1)}$ for any correct expression with θ and/or λ	1	
		$\frac{4\lambda^3}{4\lambda^3+2\lambda^2} = \frac{1}{1+\frac{1}{2\lambda}}$ which → 1 as λ → ∞	1	
		For a form from which behaviour as λ → ∞ can be deduced	1	
		These 2 marks may be awarded if expression is upside-down		5
	(iii)	$\frac{(x^2-x)^2}{x!} = \frac{x(x-1)}{(x-2)!}$	1	
		$\sum_{x=0}^{\infty}(x^2-x)^2\frac{e^{-\lambda}\lambda^x}{x!} = \sum_{x=2}^{\infty}(x^2-x)^2\frac{e^{-\lambda}\lambda^x}{x!}$		
		[for <u>convincingly</u> deleting $x = 0$ and $x = 1$ terms]	1	
		$= \lambda^2\sum_{x=2}^{\infty}x(x-1)\frac{e^{-\lambda}\lambda^{x-2}}{(x-2)!}$		
		for extracting λ^2	1	
		So $\mathrm{E}(\hat{\theta}^2) = \mathrm{E}\{(X^2 - X)^2\}$ = this ↑		
		$= \lambda^2\sum_{y=0}^{\infty}(y+2)(y+1)\frac{e^{-\lambda}\lambda^y}{y!}$ for substitution (including $\sum_0$)	1	
		$= \lambda^2\mathrm{E}[(Y+1)(Y+2)]$ for recognising E	1	
		$= \lambda^2\mathrm{E}[Y^2 + 3Y + 2]$		
		$= \lambda^2\{\lambda + \lambda^2 + 3\lambda + 2\}$	1	
		$= \lambda^4 + 4\lambda^3 + 2\lambda^2$		
		$\therefore \mathrm{var}(\hat{\theta})$ = this ↑ $- \{\mathrm{E}(\hat{\theta})\}^2$	1	
		$= \lambda^4 + 4\lambda^3 + 2\lambda^2 - \{\lambda^2\}^2$		
		$= 4\lambda^3 + 2\lambda^2$		
		$= 2\lambda^2(2\lambda + 1)$ beware printed answer	1	8

Q.2	(i)	Data / Median 230 / Difference / Rank of |difference|		
		227 -3 3	ZERO in 'difference' section if differences not used; M1	
		214 -16 9		
		218 -12 8		
		206 -24 11		
		228 -2 2		
		231 1 1	FT if ranks wrong; M1	
		226 -4 4	A1	
		202 -28 12		
		236 6 5		
		241 11 7		
		207 -23 10		
		223 -7 6		
		T = 1 + 5 + 7 = 13 (or 2 + 3 + 4 + 6 + 8 + 9 + 10 + 11 + 12 = 65)	1	
		Refer to tables of Wilcoxon single sample (/paired) statistic	1	
		Lower (or upper if 65 used) 2½% tail is needed	1	
		Value for n = 12 is 13 (or 65 if 65 used)	1	
		Result is significant	1	
		Some evidence median is not 230	1	9
	(ii)	$\bar{x} = 221.58\dot{3}$	A1	
		s_{n-1} = 12.339(6)	A1	
		Accept s_n = 11.814(2) ONLY if correctly used in sequel		
		One-sided CI is given by		
		$221.58\dot{3}$	M1	
		+	M1	
		$2.201 \times \frac{12.339(6)}{\sqrt{12}}$ [t_{11} (upper 2½%)]	B1 M1	
		$= 221.58\dot{3} + 7.840(2) = 229.42(4)$ c.a.o.	A1	
		The specified mean (230) is (just) beyond the upper confidence bound – some evidence the specification is not being met	E2	9
	(iii)	Normality of underlying population	1, 1	2

Q.3	(i)	$H_0 : \mu_A = \mu_B \quad H_1 : \mu_A \neq \mu_B$ Where μ_A and μ_B are the population mean percentages at A and B Normality of both populations same variance	1 1 1 1	4
	(ii)	$n_1 = 10\ \bar{x} = 4.246\ s_{n-1}^2 = 0.0332\ (s_{n-1} = 0.1823)\ \lfloor s_n^2 = 0.0299, s_n = 0.1730 \rfloor$ $n_2 = 10\ \bar{y} = 4.437\ s_{n-1}^2 = 0.0388\ (s_{n-1} = 0.1970)\ \lfloor s_n^2 = 0.0349, s_n = 0.1869 \rfloor$ if all correct	B1	
		Pooled $s^2 = 0.036$ For any reasonable attempt at pooling (and FT into test and CI) if correct	M1 A1	
		Test statistic is $\frac{4.246 - 4.437(-0)}{\sqrt{0.036}\sqrt{\frac{1}{10}+\frac{1}{10}}}$	M1	
		$= \frac{-0.191}{0.0848(53)} = -2.251$	A1	
		Refer to t_{18} mark may be awarded even if test statistic is wrong No FT if wrong	1	
		Dt 5% point: t_{15} is 2.131, t_{20} is 2.086 — For any reasonable attempt at interpolation If essentially correct (2.101) (nearer to t_{20})	M1 A1	
		OR award **BOTH** marks for convincing argument based on $\|2.251\| > \|2.131\|$ Significant Seems there is a different (and "B > A")	 1 1	10
	(iii)	CI given by $-0.191 \pm (2.878) \times 0.0848(53)$	M1 M1	
		t_{15} is 2.947, t_{20} is 2.845 — as in test above	M1 A1	
		$= -0.191 \pm 0.244$ $= (-0.435, 0.053)$ In repeated sampling, 99% of such intervals would 'capture' the time mean difference.	 A1 E1	6

Q.4	(i)	H_0 : no association H_1 : association	1 1	2
	(ii)	6 degrees of freedom [or zero]	2	

o_i	WR	TR	HR	Other	e_i			
High	40	7	26	15	33.44	11.44	21.12	22(.00)
Med	22	9	4	17	19.76	6.76	12.48	13(.00)
Low	14	10	18	18	22.8(0)	7.8(0)	14.4(0)	15(.00)

Deduct 1 per error. Must be to this level of accuracy. A4

Contributions to χ^2

1.28689	1.72322	1.12758	2.22727
0.25393	0.74225	5.76205	1.23077
3.39649	0.62051	0.9	0.6

$\chi^2 = 19.87$ M1

awrt 19.9 A2

[give A1 if $\in(19.5, 20.2)$]

Refer to χ_6^2 [see above; FT if wrong, unless ≈ 200]

Upper 5% point is 12.59 1

Significant 1

Seems there is association ZERO if $H_0 \leftrightarrow H_1$ 1 | 12

(iii) There are very many fewer home-related payouts than expected – home-related payouts tend to be either high or low. There are many fewer low work-related payouts than expected – these tend to be high. Travel-related payouts have some tendency not to be high, otherwise 'other' payouts. E6 | 6

Examiner’s Report

2616 Statistics 4

General Comments

There were 36 candidates from 8 centres, which is about par for the course for this unit nowadays. The work showed the usual spread of ability, but many candidates were able to score quite high marks freely.

One general point to be made concerns the statement of conclusion in context after carrying out a hypothesis test (or, sometimes, after finding a confidence interval). Some candidates are too assertive. For example, rejection of a null hypothesis by a statistical procedure cannot imply that the hypothesis is *definitely* wrong. An element of doubt always remains, and this should be reflected in the statement of conclusion in context by some form of words such as "it appears that..." or " it is reasonable to assume that..." or "there is some evidence that...", or some such phrase.

Comments on Individual Questions

Question 1 (Expectation and estimation theory)

This question was based on estimation of the square of the parameter of a Poisson distribution. As usual, it was the least popular question on the paper, but there were still a decent number of answers, many of which were quite good. The first part was to show that a given estimator was unbiased; candidates usually did this readily enough, using the standard results for the mean and variance of the Poisson distribution. The next part involved finding the efficiency of this estimator, using a given result. Again, this was usually done successfully, except that several candidates used the quotient of variances the wrong way up (i.e. the reciprocal of the correct quantity), resulting in a slight loss of marks. Candidates ought to know the correct result to use in this situation, and further ought to have recognised a problem when their "efficiency" in this case came out to be greater than 1. The third part was mostly an expectation exercise leading to the variance of the estimator that had been worked with. Candidates generally met with some success in working through this, but there was insufficient care in showing that the first two terms in the summation (for $x = 0$ and $x = 1$) are zero; this must be done with the original left-hand side of the "Hence show that" equation in the question, before cancellations involving x and $x - 1$ are undertaken. On the whole, however, it is pleasing to report that the work in this question was quite good.

Question 2 (Single-sample Wilcoxon test and one-sided *t* confidence interval)

The single-sample Wilcoxon test was usually done well, except by a small number of candidates who appeared to have no idea what to do at all. There were also a few candidates who, sadly, used the sign test; these candidates earned no marks as they had not answered the question, which *explicitly* required the Wilcoxon test. [Value of test statistic is 13, critical value for $n = 12$ is 13, so the result is significant.]

The one sided t confidence interval was also usually correctly found though, as ever, there were a few candidates who used an incorrect distribution [should be t_{11}, with upper 2½% point 2.201] and occasionally an incorrect standard error. The upper confidence bound should have come out as 229.42, so the specification value (230) is (just) beyond the confidence bound, giving some evidence that the specification is not being met.

Nearly all candidates knew that underlying Normality is required for the t-test but not for the Wilcoxon procedure.

Question 3 (Unpaired *t* test and confidence interval)

The question opened with a requirement to state the null and alternative hypotheses. These statements were often full and correct but not always so. Some candidates were not careful enough in ensuring that *population* parameters were referred to, or that there was a statement of what "μ" is. Candidates who preferred to make verbal statements sometimes did not ensure that they referred to differences in the *mean* between the two factories.

Likewise, statements of the required assumptions, though often fully correct, were sometimes not sufficiently sound. Normality of *both* underlying populations is required, and the population variances (not the sample variances) are required to be equal.

The test itself, in part (ii) of the question, was done correctly by most candidates, but there were some errors (including, occasionally, the drastic one of working it as a paired procedure). Some candidates made mistakes in pooling the estimates of variance. In some cases this was due to lamentable confusion between sample variances divided by *n* and by *n* – 1, but other mistakes occurred too. Some candidates should have known that they had made errors when they achieved a pooled estimate which was not somewhere between the two separate sample variances! In forming the test statistics, some candidates did not evaluate the standard error correctly with a factor of √{(1/10) + (1/10)}; errors such as √{/1/20} occurred. The null distribution here was t_{18}, and some form of interpolation between the tabulated values for t_{15} and t_{20} was expected (or a convincing argument that, as the calculated value of the test statistic [–2.25] exceeded both (in absolute value), it must be significant); pleasingly, most candidates had a reasonable idea what to do here.

The confidence interval in the last part was usually correct [answer is (–0.435, 0.053)]. But there was the usual crop of mistakes – even by candidates who had done the test correctly – including having an interval not centred at –0.191 and/or having an incorrect standard error (again, even though this had been right in the test), and the bizarre but recurring error of using a different distribution for the confidence interval from that which had been used in the test.

Question 4 (Chi-squared contingency table)

The hypotheses here were usually stated correctly (though one candidate said $\mu = 0$ for the null and $\mu \neq 0$ for the alternative – what on earth could that mean in this context?), and the arithmetic of the test was also usually correct [value of test statistic is 19.87, 6 degrees of freedom, critical point 12.59, significant]. Occasionally the number of degrees of freedom was incorrect, and there were occasional strange errors in calculating the expected frequencies. As usual, it was the discussion of results that was sometimes a bit insecure, though nearly all candidates were able to say something sensible and some of the discussions were intelligent and complete.