OCR
RECOGNISING ACHIEVEMENT

General Certificate of Education
Advanced Supplementary (AS) and Advanced Level
former Oxford and Cambridge modular syllabus

MEI STRUCTURED MATHEMATICS 5516

Statistics 4

Tuesday **19 JUNE 2001** Moming 1 hour 20 minutes

Additional materials:
Answer paper
Graph paper
Students' Handbook

TIME 1 hour 20 minutes

INSTRUCTIONS TO CANDIDATES

Write your name, Centre number and candidate number in the spaces provided on the answer paper/answer booklet.

Answer **three** questions.

Write your answers on the separate answer paper provided.

If you use more than one sheet of paper, fasten the sheets together.

INFORMATION FOR CANDIDATES

The approximate 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 sufficient detail of the working is shown on the answer paper to indicate that a correct method is being used.

This question paper consists of 4 printed pages.

HN/2

[Turn over

1 The lifetimes of a certain kind of electronic component are modelled by the continuous random variable X with probability density function

$$f(x) = \frac{1}{\theta^2} x e^{-x/\theta}$$

for $x > 0$, where $\theta > 0$. Data consisting of a random sample of n of these components' lifetimes are available.

(i) Find the mean of X and deduce that a reasonable estimator of θ is $\hat{\theta} = \frac{1}{2}\overline{X}$ where $\overline{X}$ represents the mean lifetime of the sample. [7]

(ii) Determine whether or not $\hat{\theta}$ is an unbiased estimator of θ and find its mean square error. [13]

2 A construction company operating at many sites uses a computer model to assess the depth of bedrock at each site. Trial borings are also made at some sites to help check the model. Neither the model nor the trial borings can be expected to give completely accurate answers, but it is important that they do not consistently differ from each other.

For a random sample of six sites, the depths (in metres) given by the model and by the trial borings are as follows.

Site	A	B	C	D	E	F
Result from model	9.2	6.5	4.8	8.7	9.6	12.5
Result from trial boring	9.9	6.3	5.1	8.1	9.5	13.0

(a) (i) Use an appropriate t test, at the 5% level of significance, to examine whether the mean difference between the depths given by the model and by the trial borings is zero. State the required distributional assumption. [10]

(ii) Provide a two-sided 90% confidence interval for the mean difference. [4]

(b) Investigate the situation using the Wilcoxon paired sample test, again using a 5% significance level. [6]

3 At a paint factory, a new pigment is being investigated. It is hoped that this will give a greater intensity of colour than the standard pigment. Specimens of paint are prepared using the new pigment and using the standard pigment; each specimen is assessed for intensity of colour.

(a) Initially, the intensities are not directly measured, but a technician ranks the specimens in order of intensity. The results are as follows, where rank 1 indicates the greatest intensity.

Specimen	Pigment	Rank order
1	new	2
2	new	3
3	new	7
4	new	1
5	new	5
6	new	13
7	new	9
8	new	6
9	standard	11
10	standard	14
11	standard	4
12	standard	10
13	standard	8
14	standard	12

The Wilcoxon rank sum test is to be used to examine whether the new pigment gives, on the whole, greater intensity.

(i) State carefully the hypotheses being tested. [4]

(ii) Carry out the test, at the 5% level of significance. [7]

(b) Later, the colour intensities for each specimen are measured photoelectrically. The results are summarised as follows, in a convenient unit.

New pigment: $n_1 = 8$, $\Sigma x = 76.8$, $\Sigma x^2 = 764.26$.
Standard pigment: $n_2 = 6$, $\Sigma y = 49.8$, $\Sigma y^2 = 447.88$.

Using this information, and assuming Normality of the underlying populations, provide a two-sided 95% confidence interval for the difference between the mean colour intensities. What else have you needed to assume? [9]

[Turn over

4 It is thought that the times (in hours) between minor breakdowns on a computer network might be modelled by the exponentially distributed random variable X with probability density function

$$f(x) = \lambda e^{-\lambda x}$$

for $x > 0$, where λ is a parameter ($\lambda > 0$). A random sample of 80 times between minor breakdowns is summarised by the following frequency distribution. In this random sample, $\bar{x} = 20$ hours.

time x (hours)	$0 < x \leqslant 10$	$10 < x \leqslant 20$	$20 < x \leqslant 30$	$30 < x \leqslant 40$	$40 < x \leqslant 50$	$x > 50$
frequency	26	16	9	10	9	10

(i) Show that, for $0 < a < b$,

$$P(a < X \leqslant b) = e^{-\lambda a} - e^{-\lambda b}.$$ [2]

(ii) Using the estimate $\hat{\lambda} = \frac{1}{\bar{x}}$, calculate the expected frequencies corresponding to the $(0, 10)$, $(10, 20)$ and $(20, 30)$ cells of the above table. [5]

(iii) The remaining expected frequencies are as follows.

cell	$30 < x \leqslant 40$	$40 < x \leqslant 50$	$x > 50$
expected frequency	7.02	4.26	6.57

The $(40, 50)$ cell has expected frequency less than 5. Suggest why, despite this, it should perhaps *not* be grouped with another cell or cells when conducting a χ^2 goodness of fit test. [3]

(iv) Carry out a χ^2 goodness of fit test, keeping *all* the cells. Use a 5% significance level. [8]

(v) Discuss briefly your conclusions. [2]

Mark Scheme

Q.1	$\mathrm{f}(x)=\frac{1}{\theta^2}xe^{-\frac{x}{\theta}} \quad (x>0;\ \theta>0)$		
	(i) $\mu=\int_0^\infty \frac{1}{\theta^2}x^2e^{-\frac{x}{\theta}}dx$	M1 by parts M1	
	$=\frac{1}{\theta^2}\left\{\left[x^2.\frac{e^{-\frac{x}{\theta}}}{-\frac{1}{\theta}}\right]_0^\infty+\theta\int_0^\infty e^{-\frac{x}{\theta}}.2x\,dx\right\}$	2, divisible, for algebra	
	$=0+\frac{2}{\theta}\int_0^\infty xe^{-\frac{x}{\theta}}dx$		
	$=\frac{2}{\theta}.\theta^2$ by reference back to pdf; or, integrate by parts again	M1	
	$=2\theta$	1	6
	So we have $\theta=\frac{1}{2}\mu$; reasonable to estimate μ by $\overline{X}$; $\therefore$ reasonable to est θ by $\frac{1}{2}\overline{X}$	E1	1
	(ii) $\mathrm{E}[\hat{\theta}]=\frac{1}{2}\mathrm{E}[\overline{X}]=\frac{1}{2}\mathrm{E}[\overline{X}]$	M1	
	$=\frac{1}{2}.2\theta=\theta$	1	
	$\therefore$ unbiased	1	3
	Being unbiased, $\mathrm{MSE}(\hat{\theta})=\mathrm{var}(\hat{\theta})$	1	
	$=\mathrm{var}(\frac{1}{2}\overline{X})=\frac{1}{4}\mathrm{var}(\overline{X})$	1	
	$=\frac{1}{4}\frac{\mathrm{var}(x)}{n}$	1	3
	So we need $\mathrm{var}(x)=\mathrm{E}[X^2]-(\mathrm{E}[X])^2$	M1	
	$\mathrm{E}[X^2]=\int_0^\infty \frac{1}{\theta^2}x^3e^{-\frac{x}{\theta}}dx$	M1 by parts M1	
	$=\frac{1}{\theta^2}\left\{\left[x^3.\frac{e^{-\frac{x}{\theta}}}{-\frac{1}{\theta}}\right]_0^\infty+\theta.3\int_0^\infty x^2e^{-\frac{x}{\theta}}dx\right\}$		
	$=\theta^2.2\theta$ from (i)	M1 or by parts …	
	$=\frac{1}{\theta^2}\{0+6\theta^4\}=6\theta^2$	1	
	$\therefore \mathrm{var}(X)=6\theta^2-(2\theta)^2=2\theta^2$	1	
	$\therefore \mathrm{var}(\hat{\theta})=\frac{\theta^2}{2n}$	1	7

Q.2	(a) Must be PAIRED COMPARISON t procedure		
	(i) 9.2 6.5 4.8 8.7 9.6 12.5 9.9 6.3 5.1 8.1 9.5 13.0 Differences -0.7 0.2 -0.3 0.6 0.1 -0.5	M1	
	$\bar{d} = -0.1 \quad s_{n-1}^2 = 0.236, \ s_{n-1} = 0.4858$	A1	
	Accept $s_n^2 = 0.19\dot{6}$, $s_n = 0.4435$, but ONLY if correctly used in sequel		
	Test statistic is $\dfrac{-0.1-0}{\frac{0.4858}{\sqrt{6}}}$	M1 M1	
	$= -0.50(42)$	A1	
	Refer to t_5	1	
	May be awarded even if test statistic is wrong, but NO f.t. if wrong		
	Dt 5% pt is 2.571 (NO f.t. if wrong)	1	
	Not significant.	1	
	Seems no overall mean difference between model and trial borings.	1	
	Needs Normality of differences.	1	10
	(ii) CI is given by		
	-0.1	M1	
	± 2.015	B1	
	$\times \dfrac{0.4858}{\sqrt{6}} = -0.1 \pm 0.39963$	M1	
	= (-0.499(63), 0.299(63)) (cao for **BOTH**)	A1	
	Zero out of 4 if not same dist as used for test, except if recovering to t_5. The two M1 marks, only, are to be awarded if t_6 is used here *and* in (i).		4
	(b) Ranks of $\lvert d\rvert$ are (6) 2 (3) 5 1 (4) . ( ) denotes a negative d	M1	
	T = 8 or 13	1 (correct answer from candidate's ds)	
	Refer smaller value to appropriate table.	M1	
	Dt 5% pt for $n = 6$ is ZERO [note for examiner – st 5% pt is 2].	1	
	Result is not significant.	1	
	Seems on the whole model and trial borings give 'the same' results.	1	6

Q.3	(a)(i)	Strictly,		
		Let colour intensity have c.d.f. $F(x)$ for standard pigment	1	
		And c.d.f. $F(x - \Delta)$ [NB same F] for new pigment	1	
		H_0 is $\Delta = 0$	1	
		H_1 is $\Delta < 0$	1	4
		If expressed verbally, Distributions have similar shape (E2) H_0 : location-parameters (allow medians, means) equal (E) H_1 : location-parameters (allow medians, means) different, being greater for new pigment (E1)		
	(ii)	Ranks are NEW 1 2 3 5 6 7 9 13 STANDARD 4 8 10 11 12 14		
		Rank sum 59 (or 46 if larger sample used at this stage) (Mann-Whitney, if directly calculated, is 38)	1	
		Refer to table of Wilcoxon rank sum (or Mann-Whitney) statistic	M1	
		Upper 5% tail is needed	M1	
		WILCOXON FORM | MANN-WHITNEY FORM		
		WILCOXON FORM: Lower 5% value for (6, 8) is 31 (might be obtained as M-W value of $10 + \frac{1}{2}m(m+1)$). MANN-WHITNEY FORM: Lower 5% value for (6, 8) is 10 (might be obtained as Wilcoxon value of $31 - \frac{1}{2}m(m+1)$)	1	
		WILCOXON FORM: Mean is $\frac{1}{2}m(m+n+1) = 45$, So upper 5% value is 59 (or, refer test statistic value of 31 to lower tail). MANN-WHITNEY FORM: Mean is $\frac{1}{2}mn = 24$, so upper 5% value is 38 (or, refer test statistic value of 10 to lower tail) (M-W test statistic might be calculated via Wilcoxon, as $59 - \frac{1}{2}m(m+1) = 38$).	1	
		WILCOXON FORM: 59 is significant at the 5% level MANN-WHITNEY FORM: 38 is significant at the 5% level	1	
		New pigment appears to lead to greater intensity (allowable as ft).	1	7
	(b)	Must assume population variances are the same	1	
		$n_1 = 8 \quad \bar{x} = \frac{76.8}{8} = 9.6 \quad s_1^2 = \frac{1}{7}(26.98) = 3.854;$ $n_2 = 6 \quad \bar{y} = \frac{49.8}{6} = 8.3 \quad s_2^2 = \frac{1}{5}(34.54) = 6.908$		
		Pooled $s^2 = \frac{26.98 + 34.54}{12} = 5.12\dot{6}$	M1 A1	
		$9.6 - 8.3 \pm 2.179\sqrt{5.12\dot{6}}\sqrt{\frac{1}{8}+\frac{1}{6}}$		
		9.6 − 8.3	M1	
		$\sqrt{\ }\sqrt{\ }$	M1	
		2.179	B1	
		$t_{12}(5\%)$	M1	
		$= 1.3 \pm 2.179 \times 2.264 \times 0.540$ = 2.66(45)		
		$= (-1.36(45), 3.96(45))$	A1 A1 cao	9

Q.4	(i) $P(a<x<b)=\int_a^b \lambda e^{-\lambda x}\,dx$	M1	
	$=\left[-e^{-\lambda x}\right]_a^b = e^{-\lambda a} - e^{-\lambda b}$	1	2
	(ii) $\hat{\lambda}=\frac{1}{\bar{x}}=\frac{1}{20}=0.05$		
	$P(0<X<10)=e^{-0}-e^{-0.5}=1-0.6065=0.3935$ $\therefore e=80\times 0.3935$		
	Award M1 **once** for a probability, M1 **once** for $e=80\times$ prob $=31.48$	A1	
	$P(10<X<20)=e^{-0.5}-e^{-1}=0.6065-0.3679=0.2386$ $\therefore e=19.09$	A1	
	$P(20<X<30)=e^{-1}-e^{-1.5}=0.3679-0.2231=0.1447$ $\therefore e=11.58$	A1	5
	(iii) We have o_i 26 16 9 10 9 10 e_i 31.48 19.09 11.58 7.02 4.26 6.57 Discussion about the 'e' of 4.26 SHOULD include $e<5$ is only a rule of thumb, not a hard-and-fast law	E1	
	and MIGHT include points such as • 4.26 is not much less than 5 • some other 'e' values are not much more than 5 – arbitrary and unsatisfactory to treat them differently • this cell might turn out to contain important information – unsatisfactory to sacrifice it • these aren't many cells anyway – unsatisfactory to reduce their number still further	Award E1 E1 for any **two** sensible comments	3
	(iv) X^2 (= 0.95395 + 0.5002 + 0.5748 + 1.2650 + 5.2741 + 1.7907) = 10.36 [10.3587]	M1 A1	
	Refer to χ_4^2 [or ZERO; FT if df wrong, unless ≈ 80]	3	
	Upper 5% pt is 9.488. No f.t. if wrong	1	
	Significant.	1	
	Suggests model does not fit data. ZERO for 'data do not fit model'	1	8
	(v) The main point is that the data are 'heavy in the tail' and 'light near the origin'	E2	2

Examiner’s Report

Statistics 4 (5516)

General Comments

The great majority of candidates were well prepared for this paper and there were many excellent scripts and very few poor scripts. The question on estimation continues to be the least popular question by far, but there were some excellent concise solutions.

Although many candidates were able to carry out most calculations successfully, this was far less the case when any explanation was required. Statements of hypotheses, conclusions at the end of hypothesis tests and required assumptions are all areas where improvements are needed.

Comments on Individual Questions

Question 1

Few candidates attempted this question, but there were some very good solutions. The best solutions were often based on the general result $\int_0^\infty x^n e^{-x} dx = n!$

In part (i) virtually all candidates realised they needed $\int_0^\infty \frac{1}{\theta^2} x^2 e^{-\frac{x}{\theta}} dx$ and usually managed to obtain the correct result. Many candidates had difficulty explaining why $\frac{1}{2}\bar{X}$ is a reasonable estimator of θ.

In part (ii) candidates clearly understood the meaning of an unbiased estimator but often their explanation lacked clarity. Many candidates were much less clear about mean square error, but managed to pick up some marks by calculating the variance of X.

(i) 2θ; (ii) $\frac{\theta^2}{2n}$.

Question 2

In part (a)(i) virtually all candidates realised that a paired comparison t test was required and had a good understanding of the procedure. A small number of candidates became confused between s_n and s_{n-1}, but this was rare. Most candidates carried out the test successfully with the only errors seen more than a few times being the use of t_4 or t_6 or the use of a one-tailed test. Many candidates lost marks were in the final conclusion of the test and also in the required distributional assumption.

Many candidates gave a conclusion which made no reference to the context of the question and this is insufficient. Candidates commonly mentioned underlying Normality, but did not realise that it is the Normality of differences that is required.

Part (ii) was done well by most candidates, with only a handful reverting to z values.

Part (b) was handled very well indeed with most candidates scoring highly. Again, thought, the conclusion was done poorly by many candidates.

(a)(i) Not significant, seems no overall mean difference between model and trial borings;
(a)(ii) (−0.4996, 0.2996);
(b) Not significant, on the whole the model and the trial give 'the same' results.

Question 3

In (a)(i) the question asks for the hypotheses to be stated carefully and there are 4 marks available. In these circumstances it should be clear that

H_0: the two pigments give the same intensity

H_1: the new pigment gives a greater intensity
are not sufficient

What was required is as follows, or something very similar:

H_0: the distributions have a similar shape with location parameters equal

H_1: the distributions have a similar shape with the new pigment having a greater location parameter.

Part (a)(ii) was well done by most candidates who coped well with the fact that the test statistic was equal to the critical value given in the tables. Where errors were made it was usually because of a confusion between the Wilcoxon form and the Mann-Whitney form. A smaller number of candidates compared their test statistic with the critical value from the wrong tail.

Part (b) was not well done and the correct confidence interval was rarely seen.

Common errors were:

a failure to attempt to find a pooled variance;

a confusion between s_n and s_{n-1} in the calculation of the pooled variance;

use of the wrong t distribution, often t_{13} or, more rarely, the N(0, 1) distribution;

use of $\frac{\sigma}{\sqrt{n}}$ in the confidence interval;

use of $\sqrt{\frac{1}{8}+\frac{1}{6}}$ as a divisor in the confidence interval.

Most candidates gave the required assumption that the population variances are the same, although often amongst several other possible assumptions.

(a)(ii) significant, new pigment appears to lead to greater intensity
(b) (–1.3645, 3.9645)

Question 4

Part (i) was done well by most candidates, although the result was often faked, particularly by candidates who tried to obtain the result using the cumulative distribution function. Many of these candidates found the cdf to be $-e^{-\lambda x}$ rather than $1 - e^{-\lambda x}$.

Part (ii) was done extremely well with virtually all candidates correctly obtaining the expected frequencies.

In part (iii) most candidates made a number of sensible comments as to why grouping might not be appropriate. One important point rarely mentioned was that $e < 5$ is only a rule of thumb.

Most candidates demonstrated in (iv) that they could correctly calculate the χ^2 statistic. Many however used χ_5^2, not realising λ had been estimated using the data. χ_3^2 was also seen occasionally.

In part (v) many candidates focused on the expected frequency of 4.26 and a possible recalculation of the χ^2 statistic, but with grouping. The required response was the poor fit in the tails.

(ii) 31.48, 19.09, 11.58; (iv) significant, suggests model does not fit data.