RECOGNISING ACHIEVEMENT

Oxford, Cambridge and RSA Examinations

Advanced Subsidiary General Certificate of Education
Advanced General Certificate of Education

MEI STRUCTURED MATHEMATICS 2616

Statistics 4

Wednesday **16 JANUARY 2002** Mornng 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 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 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 5 printed pages and 3 blank pages.

[Turn over

1 The random variables X, Y, Z have independent Normal distributions as follows, where σ^2 is an unknown constant.

$$X \sim \mathrm{N}(0, \sigma^2) \qquad Y \sim \mathrm{N}(0, \sigma^2) \qquad Z \sim \mathrm{N}(0, 2\sigma^2)$$

You are *given* the following distributional results.

$$X^2 \sim \sigma^2\chi_1^2 \qquad Y^2 \sim \sigma^2\chi_1^2 \qquad Z^2 \sim 2\sigma^2\chi_1^2$$

$$\mathrm{E}(\chi_1^2) = 1 \qquad \mathrm{Var}(\chi_1^2) = 2$$

It is proposed to estimate σ^2 by

$$T = \tfrac{1}{3}(X^2 + Y^2 + kZ^2)$$

where k is a constant to be determined.

(i) Find the expected value of T. [3]

(ii) Find the variance of T. [4]

(iii) Find the value of k for which T is an unbiased estimator of σ^2. State the value of the variance of T for this value of k. [3]

(iv) Show that the mean square error of T is

$$\frac{\sigma^4}{9}(12k^2 - 4k + 5). \qquad [3]$$

[You may use without proof the formula for mean square error.]

(v) Find the value of k for which the mean square error of T is a minimum. State the value of the mean square error of T for this value of k. [5]

(vi) Comment on your results in parts **(iii)** and **(v)**. [2]

2 A psychologist is studying the possible effect of hypnosis on dieting and weight loss. Nine people (who may be considered as a random sample from the population under study) volunteer to take part in an experiment. Their weights are measured. Then, under hypnosis, they are told that they will seldom feel hungry and will eat less than usual. After a month, their weights are measured again. The results (in kg) are as follows.

Person	Initial weight	Weight after one month
A	83.7	81.5
B	83.9	80.0
C	68.2	68.8
D	74.9	74.1
E	81.0	82.6
F	72.8	69.2
G	61.3	63.4
H	77.9	74.7
I	69.6	66.2

(i) Use an appropriate *t* test to examine whether, overall, the mean weight has been reduced over the month, at the 5% level of significance. [10]

(ii) Provide an alternative analysis using an appropriate Wilcoxon test, again at the 5% level of significance. [6]

(iii) What distributional assumption is needed in part (i) but not in part (ii)? By considering the data, comment briefly and informally on whether this assumption appears to hold. [You may wish to use a simple diagram.] [4]

3 A factory receives deliveries of an electrical component from two suppliers, A and B. The resistances of these components are critical. Inspectors are checking whether, on average, resistances of components from supplier A are the same as those from supplier B. The resistances are measured for a random sample of nine components from supplier A and for a random sample of eight components from supplier B. The results, in ohms, are as follows.

Supplier A: 18.62 18.44 18.47 18.45 18.29 18.65 18.41 18.50 18.44

Supplier B: 18.53 18.64 18.58 18.62 18.72 18.65 18.45 18.69

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

(ii) Carry out the test, at the 5% significance level. [8]

(iii) Provide a one-sided 95% confidence interval giving a lower bound for $\mu_B - \mu_A$, where μ_B denotes the mean resistance for supplier B's components and μ_A similarly for supplier A. Interpret carefully the meaning of this interval. [5]

(iv) Suppose that, following long experience with the components, the inspectors were prepared to take the true underlying variances ($\sigma_A{}^2$ for the resistances of supplier A's components, $\sigma_B{}^2$ for supplier B's) as known. Desribe briefly [no calculations are required] how the procedures in parts **(ii)** and **(iii)** should be modified. [3]

4 A 2 × 2 contingency table with its observed frequencies may be represented algebraically as

$$\begin{array}{cc|l} a & b & a+b \\ c & d & c+d \\ \hline a+c & b+d & a+b+c+d=N \end{array} \qquad (1)$$

It can be shown that the usual X^2 test statistic (without Yates' correction) for testing independence of the rows and columns may be written as

$$X^2 = \frac{N(ad-bc)^2}{(a+b)(c+d)(a+c)(b+d)}. \qquad (2)$$

(i) Use formula (2) to carry out the test, at the 5% level of significance, for the following data. The data are extracted from the medical records of a random sample of patients of a large general practice, showing for part of a particular year the frequencies of contracting or not contracting influenza for patients who had or had not had influenza inoculations.

		Influenza	
		Yes	No
Inoculated	Yes	8	18
	No	35	17

[8]

(ii) State an advantage and a disadvantage of using the formula (2) when carrying out the test rather than calculating X^2 in the conventional way as $\sum \frac{(o-e)^2}{e}$. [2]

(iii) Write down an algebraic expression for the expected frequency corresponding to a in the contingency table (1), on the usual assumption of independence of rows and columns. Denoting this by e_{11}, show that

$$a - e_{11} = \frac{ad-bc}{N}.$$

Write down similar expressions for $b - e_{12}$, $c - e_{21}$ and $d - e_{22}$, where the e_{ij} denote the respective other expected frequencies. Hence derive formula (2). [10]

Mark Scheme

MEI Statistics 4 (2616)

Marking Scheme – January 2002

Q1. $X \sim N(0, \sigma^2)$ $\qquad Y \sim N(0, \sigma^2)$ $\qquad Z \sim N(0, 2\sigma^2)$

$X^2 \sim \sigma^2\chi_1^2$ $\qquad Y^2 \sim \sigma^2\chi_1^2$ $\qquad Z^2 \sim 2\sigma^2\chi_1^2$

$T = \frac{1}{3}\left(X^2 + Y^2 + kZ^2\right)$, to estimate σ^2

(i) $E(T) = \frac{1}{3}\left(\sigma^2 + \sigma^2 + k.2\sigma^2\right) = \frac{\sigma^2}{3}(2 + 2k)$ [1] **[3]**

[B1] [1]

(ii) $\text{Var}(T) = \frac{1}{9}\left(\sigma^4 2 + \sigma^4 2 + k^2 4\sigma^4 2\right) = \frac{\sigma^4}{9}\left(4 + 8k^2\right)$ [1] **[4]**

[1] [M1] [1]

(iii) For T to be unbiased, we require $E(T) = \sigma^2$ [M1]

So $\frac{\sigma^2}{3}(2 + 2k) = \sigma^2 \rightarrow \underline{\underline{k = \frac{1}{2}}}$ [A1]

For $k = ½$, $\text{Var}(T) = \frac{\sigma^4}{9}(4 + 2) = \underline{\underline{\frac{2}{3}\sigma^4}}$ [A1] **[3]**

(iv) $\text{MSE}(T) = \text{Var}(T) + \{\text{bias}\}^2$

$= \frac{\sigma^4}{9}\left(4 + 8k^2\right) + \{\frac{\sigma^2}{3}(2 + 2k) - \sigma^2\}^2$ [M1]

$= \frac{\sigma^4}{9}\left(4 + 8k^2 + (2 + 2k\text{-}3)^2\right)$

$= \frac{\sigma^4}{9}\left(4 + 8k^2 + 4k^2 - 4k + 1\right)$

$= \frac{\sigma^4}{9}\left(12k^2 - 4k + 5\right)$ [A2] divisible, for this algebra;

beware printed answer **[3]**

(v) For minimum, consider $\frac{d\text{MSE}}{dk} = 0$ [M1]

i.e. $\frac{\sigma^4}{9}(24k - 4) = 0 \rightarrow \underline{\underline{k = \frac{1}{6}}}$

[1] [1]

Check $\underline{\text{min}}$imum, e.g $\frac{d^2\text{MSE}}{dk^2} = \frac{\sigma^4}{9} \cdot 24 \ > 0$ [1]

For $k = \frac{1}{6}$, $\text{MSE}(T) = \frac{\sigma^4}{9}\left(12 \times \frac{1}{36} - \frac{4}{6} + 5\right)$

$= \frac{\sigma^4}{9}\left(\frac{1}{3} - \frac{2}{3} + 5\right) = \underline{\underline{\frac{14\sigma^4}{27}}}$ [A1] **[5]**

(vi) The unbiased estimator in (iii) has more variability than the estimator in (v).
[Might notice that the unbiased estimator gives more weight to the
more variable Z.] [E2] **[2]**

Q2.

83.7	81.5
83.9	80.0
68.2	68.8
74.9	74.1
81.0	82.6
72.8	69.2
61.3	63.4
77.9	74.7
69.6	66.2

(i) MUST be PAIRED COMPARISON t test.

Differences are
2.2 3.9 –0.6 0.8 –1.6 3.6 –2.1 3.2 3.4 [M1]

$\bar{d} = 1\cdot 4\dot{2}$ $s_{n-1} = 2.36(05)$, $s_{n-1}{}^2 = 5\cdot 97(19)$ [A1] Accept $s_n = 2\cdot 22(55)$, $s_n{}^2 = 4\cdot 95(28)$, but ONLY if correctly used in sequence.

Test statistic is

[M1] [M1] [M1] $\dfrac{1.4\dot{2} - 0}{\dfrac{2.36(05)}{\sqrt{9}}} = 1.80(75)$ [A1]

Refer to t_8. [1] May be awarded even if test statistic is wrong. No FT if wrong.

Single tail 5% point is 1.860 [1] No FT.

Not significant [1]

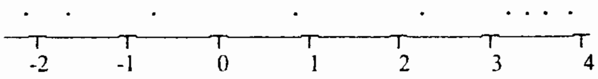
Seems weight unchanged [1] **[10]**

(ii) Must be PAIRED WILCOXON test.
Ranks of |d| are

5 9 $\underset{\text{neg}}{\underset{\uparrow}{1}}$ 2 $\underset{\text{neg}}{\underset{\uparrow}{3}}$ 8 $\underset{\text{neg}}{\underset{\uparrow}{4}}$ 6 7 [M1]

Test statistic = 1 + 3 + 4 = 8 [or 37] [A1]
Refer to paired Wilcoxon table with $n = 9$ [M1]
lower 5% point is 8 [upper is 37] [1]
∴ the observed 8 [or 37] is significant. [1]
seems weight has changed downwards [1] **[6]**

(iii) Normality of differences. [1]
Consider e.g. dotplot of d's :- [M1] Or for any other relevant display/discussion of the data.



Sample is admittedly small, but appears nothing like Normal. [E2]

[4]

Q3.

(i) $H_0 : \mu_A = \mu_B$
$H_1 : \mu_A \neq \mu_B$ [1]

where μ_A, μ_B are the population mean resistances for components from A and B. [1]

Normality of both populations [1], same variance [1]. **[4]**

(ii) $n_1 = 9$ $\bar{x} = 18.47\dot{4}$ $s^2_{n-1} = 0.0117$ $(s_{n-1} = 0.1083)$ [s^2_n=0.0104, s_n=0.1021]
$n_2 = 8$ $\bar{y} = 18.61$ $s^2_{n-1} = 0.0077$ $(s_{n-1} = 0.0878)$ [s^2_n=0.00675, s_n=0.0822]

Pooled $s^2 = \dfrac{8 \times 0.0117 + 7 \times 0.0077}{15} = 0.0098\dot{3}$ [M1]—>Must be CORRECT method
[A1] but FT anything reasonable into the test and CI

Test statistic is

[M1] $\dfrac{18.47\dot{4} - 18.61\,(-0)}{\sqrt{0.0098\dot{3}}\sqrt{\frac{1}{9}+\frac{1}{8}}} = \dfrac{-0.13\dot{5}}{0.0482} = -2.812$ [A1] [beware: very ill-conditioned]

$\uparrow$
=0.09916

Refer to t_{15} [1] May be awarded even if test statistic is wrong. No FT if wrong.
Double tail 5% pt is 2.131 [1] No FT if wrong.
Significant. [1]
Seems there is a diference
(and "B > A"). [1] **[8]**

(iii) CI:-
lower bound given by [M1] if $0.13\dot{5}$ and 0.0482 both correct (FT cand's values)
$+0.13\dot{5} - 1.753 \times 0.0482 = 0.136 - 0.084(5) = 0.051(5)$
[B1]

i.e. interval is (0.051(5), ∞). [A1]

Inerpretation: in repeated sampling, lower bounds obtained in this way fall below the true value of $\mu_B - \mu_A$ on 95% of all occasions. [E2] **[5]**

(iv) SD in test statistic and CI becomes

$$\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$$

[1] for (almost!) any form involving the two separate σ^2s

[1] if correct

and N(0,1) is used as the reference distribution [1] **[3]**

Q4.

a	b	$a+b$
c	d	$c+d$
$a+c$	$b+d$	N

$$X^2 = \frac{N(ad\text{-}bc)^2}{(a+b)(c+d)(a+c)(b+d)}$$

(i)

		Influenza		
		Yes	No	
Inoculated	Yes	8	18	26
	No	35	17	52
		43	35	78

$$X^2 = \frac{78(8\times17-35\times18)^2}{26\cdot52\cdot43\cdot35}$$

[3] for correct identification of <u>all</u> terms in formula;
[-1] each error.

So $X^2 = \dfrac{78(-494)^2}{26\cdot52\cdot43\cdot35} = \dfrac{78\times244036}{2034760} = 9.35(482)$ [A1] c.a.o.

Refer to χ_1^2 [1] No FT if wrong.
Upper 5% point is 3.84 [1] No FT if wrong.
Significant [1]
Seems there is association. [1]

[8]

(ii)

Calculation is easier. [E1]
The e's are not computed – we can't see where any associations might lie [E1]
[Other <u>sensible</u> comments will be rewarded]

[2]

(iii)

$$e_{11} = \frac{(a+b)(a+c)}{N}$$ [1]

$$\therefore a-e_{11} = a-\frac{(a+b)(a+c)}{N}$$ [1]

$$= \frac{a(a+b+c+d)-(a+b)(a+c)}{N} = \frac{ad-bc}{N}$$ [1]

$b-e_{12} = \frac{bc\text{–}ad}{N}$ $c-e_{21} = \frac{bc\text{–}ad}{N}$ $d-e_{22} = \frac{ad\text{–}bc}{N}$ [2]if <u>all</u> correct;
allow [1]if any one is correct.

$$X^2 = \frac{[(ad-bc)/N]^2}{(a+b)(a+c)/N} + \frac{[(bc-ad)/N]^2}{(a+b)(b+d)/N} + \frac{[(bc-ad)/N]^2}{(a+c)(c+d)/N} + \frac{[(ad-bc)/N]^2}{(b+d)(c+d)/N}$$ [1]

$$= \underset{\underset{[1]}{\uparrow}}{\frac{(ad-bc)^2}{N}} \times \underset{\underset{[1]}{\uparrow}}{\frac{(b+d)(c+d)+(a+c)(c+d)+(a+b)(b+d)+(a+b)(a+c)}{(a+b)(c+d)(a+c)(b+d)}}$$

$$= \frac{(ad-bc)^2}{N} \; \overset{\overset{[1]}{\downarrow}}{\frac{(a+b+c+d)^2}{denom.}} \qquad = \frac{N(ad - bc)^2}{denom.} \quad [1]$$

Reward other algebra routes similarly. **[10]**

Examiner’s Report

Statistics 4 (2616)

General Comments

The big feature was that there were only 16 candidates (from 6 centres) – very many fewer, by about an order of magnitude, than ever before for the winter sitting of this unit. Happily, much of the work was of a good standard, indeed sometimes outstandingly good. Some candidates, however, were badly let down by serious lack of even fairly elementary skill in algebra, which is rather disturbing for a unit at this level.

Comments on Individual Questions

Question 1 (estimation of a variance, bias, mean square error)

This question, which was based on some given results about the chi-squared distribution with 1 degree of freedom, was moderately popular. It was pleasing to see some very good, careful and thorough solutions; these candidates were clearly not afraid of the more mathematical aspects of this question. Nor should they be, of course. As has been said before in these reports, there are usually plenty of intermediate results, and it is only necessary to proceed carefully through the question a step at a time. Most attempts got a long way, often right through to the end, by doing exactly that, though there were a few candidates who could only make small progress. The expected value and variance of T gave little difficulty [$2\sigma^2(1 + k)/3$ and $4\sigma^4(1 + 2k^2)/9$ respectively], and candidates readily realised that, for this to be an unbiased estimator of σ^2, the value of k must be ½. Most candidates then correctly used the mean square error formula to derive the result quoted on the question paper, and this was duly differentiated to show that $k = \frac{1}{6}$ gives a minimum – except that a few candidates at this stage neglected to confirm that this is indeed a *min*imum. The discussion at the end was usually good, candidates realising that the unbiased estimator has greater variance than the estimator constructed in part (v), so that the latter is in many ways to be preferred.

Question 2 (paired *t* and Wilcoxon tests)

This question was usually done well. The t test statistic was correctly worked out [1.80(75)] and correctly referred to the t_8 distribution [critical point 1.860] to show that the result was not significant. In the Wilcoxon test, it happened that the value of the test statistic [8] was *exactly* the value given in the tables. Only some of the candidates knew the important point that the tables show values that are *just in* the critical region, so that this test finds the result to be signficant. Candidates knew that the distributional result required for the t test had something to do with Normality, but 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. There were some very good discussions of the data and whether it appeared that the required Normality held. Several candidates used the very useful "dotplot" diagram that is exhibited in the published markscheme; others approached the situation in other ways. Some candidates, however, did not consider the data at all, merely opining in a general way that the situation was likely to be Normal. These candidates missed the point altogether.

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

Another question that was often done well. That said, the opening statements of the hypotheses and assumptions were not always fully correct. This is another area where candidates at this level are expected to be completely secure. For example, the Normality must refer to *both* underlying *populations*, and equality of variance to *population* (not sample) variances. Proceeding onwards, candidates generally knew how to form the pooled estimator of the assumed common variance, how to construct the test statistic [value is around –2.81, but the calculation is very dependent on the accuracy of intermediate working], that it should be referred to t_{15} and that the result is significant (which was duly followed in nearly all cases by the short verbal conclusion in the context of the original problem that is always expected in these questions, and which had also been successfully done in question 2). Candidates also usually had the correct idea of forming the lower confidence bound [0.051(5)], but there were some errors here. Perhaps unsurprisingly, an occasional upper bound appeared, or a two-sided interval. Much more surprisingly, and disturbingly, there were candidates who had successfully carried out the test but then used an incorrect standard error for the confidence bound. Candidates' interpretations of the bound were slightly insecure; candidates appear to have got used to interpreting a two-sided confidence interval, but were not always quite sure what to say in this one-sided case. At the end of the question, candidates knew that, if the true underlying variances were known, they should be used directly in the test statistic which in then referred to N(0,1).

Question 4 (2×2 contingency table, using an alternative formula for the value of the test statistic)

The alternative formula used in this question is quite well known, though it might have been unfamiliar to many candidates. Use of it to obtain the value of the test statistic [9.35(48)] caused no difficulties, and the test was correctly carried out by reference to χ^2_1. In part (ii), all candidates noted that this is a much easier calculation, and most noted the concomitant drawback that the individual expected frequencies are not calculated, meaning that we cannot directly see where any apparent associations appear to lie. Part (iii) was a guided algebraic derivation of the alternative formula. Very pleasingly, there were quite a few careful and completely correct derivations. At the other extreme, some candidates did not try it at all. In the middle, there were some candidates who made an honest attempt but got stuck – and some who were dishonest. These candidates made a slip at some stage and then "corrected" this by some obviously false statement made so that the given answer could be achieved. All candidates should be warned that it is pointless to do this. They will not pull the wool over the examiner's eyes!