OCR
RECOGNISING ACHIEVEMENT

OXFORD CAMBRIDGE AND RSA EXAMINATIONS

Advanced Subsidiary General Certificate of Education
Advanced General Certificate of Education

MEI STRUCTURED MATHEMATICS 2616

Statistics 4

Thursday **30 MAY 2002** 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 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.

1 (i) The random variable X has the exponential distribution with parameter λ ($\lambda > 0$), so that its probability density function is

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

for $x \geqslant 0$. Derive the mean and variance of X. [9]

(ii) The random variable Y is defined by

$$Y = X + \theta$$

where θ is a positive constant. State the mean and variance of Y. [2]

An experimental device for recording the output of particles from a radioactive source is being investigated. The output of a particle is defined as an 'event'. The times between successive events are exponentially distributed with parameter λ. The device is known to have a 'dead time' θ, so that it can only detect an event if it occurs at least time θ after the previous event. Time intervals between events as recorded by the device are assumed to be given by the random variable Y in part **(ii)**. A random sample of n such intervals is recorded. The sample mean is represented by $\bar{Y}$ and the sample variance by S^2.

(iii) Use the mean and variance of Y to show that plausible estimators of λ and θ are

$$\hat{\lambda} = \frac{1}{S} \quad \text{and} \quad \hat{\theta} = \bar{Y} - S.$$ [3]

(iv) Now suppose that the device is used to record events for a standard source for which λ is *known* to be 3. Use the mean of Y to show that in these circumstances θ may plausibly be estimated by

$$\tilde{\theta} = \bar{Y} - \tfrac{1}{3}.$$

State the mean and variance of $\tilde{\theta}$, explaining why the results indicate desirable properties of $\tilde{\theta}$ as an estimator of θ. [6]

2 In a steelworks, several skilled technicians are testing two machines that can be used to cut rods of steel to, approximately, the required lengths. The machines always cut to slightly above the specified length, so that the rod may later be ground down to the required length. However, this causes a waste of time and material. The purpose of the test is to find which machine, if either, is better on the whole at cutting very near to the specified length and thus minimising wastage.

Each technician uses each machine to cut rods to a particular specified length.The excess length is then carefully measured, and the results, in cm, are as follows.

Technician	Machine A	Machine B
1	2.9	1.9
2	1.8	1.4
3	4.7	3.4
4	2.7	3.3
5	2.9	2.0
6	2.4	2.4
7	5.2	3.2
8	2.9	2.1

(i) Use an appropriate t test to examine, at the 5% level of significance, whether either machine is better, stating carefully your null and alternative hypotheses and the required distributional assumption. [12]

(ii) Provide a two-sided 99% confidence interval for the true mean difference in excess lengths between the machines. [4]

(iii) In the design of the experiment, every technician uses both machines. Explain what this is intended to achieve. Discuss also another feature that should be incorporated in the design. [4]

[Turn over

3 A food manufacturer is investigating a new process for preparing the blend of tea for teabags.

(a) Samples of teabags are prepared using the existing and new processes, and cups of tea are made from each in a carefully controlled way. A member of a tasting panel is asked to rank the cups of tea in order of flavour. The results are as follows (rank 1 indicates the best).

Cup	Process	Rank order
1	Existing	5
2	Existing	6
3	Existing	3
4	Existing	12
5	Existing	10
6	Existing	11
7	New	2
8	New	7
9	New	8
10	New	1
11	New	9
12	New	4

Use the Wilcoxon rank sum test, at the 10% level of significance, to examine whether there are, on the whole, any differences between the flavours produced by the two processes. State carefully the hypotheses being tested.

Discuss briefly *two* precautions which should be used in setting up the tasting experiment. [11]

(b) The strengths of the blends can be measured directly. Data for further random samples from the two processes are summarised as follows.

Existing process: $n_1 = 10$, sample mean = 17.3, sample variance (divisor $n - 1$) = 4.62.

New process: $n_2 = 10$, sample mean = 19.2, sample variance (divisor $n - 1$) = 6.48.

Using these data, and stating two required assumptions, test at the 5% level of significance whether the mean strength for the new process is greater than that for the old. [9]

4 **(i)** A 2×2 contingency table has the following observed frequencies.

2	8
13	2

Carry out the usual chi-squared test for independence of the rows and columns, both with and without use of Yates' correction. [You should find that one of the expected frequencies is less than 5. You should *not* group this expected frequency with any other in the calculations.] [7]

(ii) You are given the result $P(Y > y) = 2P(Z > \sqrt{y})$ where $Y \sim \chi_1^2$ and $Z \sim N(0, 1)$. Use this to find the approximate level of significance of the data as given by each of the calculations in part **(i)**. [3]

(iii) Consider the following general notation for an observed 2×2 table.

n_{11}	n_{12}	r_1
n_{21}	n_{22}	r_2
c_1	c_2	n

Regarding the row and column totals (r_1, r_2, c_1, c_2) as fixed, it can be shown that the probability of obtaining this table if the rows and columns are independent is

$$\frac{\binom{r_1}{n_{11}}\binom{r_2}{n_{21}}}{\binom{n}{c_1}}.$$

Use this result to show that the probability of obtaining the table in part **(i)** if the rows and columns are independent is 0.0014455, truncated at the 7th decimal place.

[Note: $\binom{25}{15} = 3\,268\,760$.] [2]

(iv) Explain why the following observed table would be more extreme than the table in part **(i)** if the rows and columns are independent.

1	9
14	1

[2]

(v) You are given that there are also two more observed tables that would be more extreme than the table in part **(i)** if the rows and columns are independent. These are as follows.

0	10
15	0

10	0
5	10

Use the result in part **(iii)** to find the exact level of significance of the data in part **(i)**. [6]

Mark Scheme

(Qu 1) $f(x) = \lambda e^{-\lambda x} \quad x \geq 0 \ (\lambda > 0)$

(i) $E(X) = \int_0^\infty \lambda x e^{-\lambda x}\,dx$ (M1)

$= \lambda \left\{ \left[x \frac{e^{-\lambda x}}{-\lambda} \right]_0^\infty + \int_0^\infty \frac{e^{-\lambda x}}{\lambda}\,dx \right\}$ By parts (M1) (A1)

$= 0 + \left[\frac{e^{-\lambda x}}{-\lambda} \right]_0^\infty = \frac{1}{\lambda}$ (1)

$E(X^2) = \int_0^\infty \lambda x^2 e^{-\lambda x}\,dx$

$= \lambda \left\{ \left[x^2 \frac{e^{-\lambda x}}{-\lambda} \right]_0^\infty + \int_0^\infty \frac{2x e^{-\lambda x}}{\lambda}\,dx \right\}$ (M1) (A1)

$= 0 + 2 \cdot \frac{1}{\lambda} E(X)$ (M1) (or by parts again)

$= \frac{2}{\lambda^2}$ (1)

$\therefore \text{Var}(X) = E(X^2) - (E(X))^2 = \frac{2}{\lambda^2} - \frac{1}{\lambda^2} = \frac{1}{\lambda^2}$ (1) **[9]**

(ii) $Y = X + \theta$

$E(Y) = \frac{1}{\lambda} + \theta$ (1) $\quad \text{Var}(Y) = \frac{1}{\lambda^2}$ (1) **[2]**

(iii) From the expressions for $E(Y)$ and $\text{Var}(Y)$, we take

$\left. \begin{aligned} \bar{Y} &= \frac{1}{\hat{\lambda}} + \hat{\theta} \\ S^2 &= \frac{1}{\hat{\lambda}^2} \end{aligned} \right\}$ (M1)

$\rightarrow \hat{\lambda} = \frac{1}{S}$ (1) , $\hat{\theta} = \bar{Y} - S$ (1) **[3]**

(iv) We now have $\lambda = 3$, so $E[Y] = \frac{1}{3} + \theta$.

So we take $\bar{Y} = \frac{1}{3} + \tilde{\theta}$ (M1) $\rightarrow \tilde{\theta} = \bar{Y} - \frac{1}{3}$ (1)

$E(\tilde{\theta}) = E(\bar{Y}) - \frac{1}{3} = \frac{1}{3} + \theta - \frac{1}{3} = \theta$ (1) ie unbiased (1)

$\text{Var}(\tilde{\theta}) = \text{Var}(\bar{Y}) = \frac{1}{9n}$ (1) which is "small" and gets smaller as n increases (1) **[6]**

(Qu 2)

(i) $H_0: \mu_D = 0 \quad (\text{or } \mu_A = \mu_B)$

$H_1: \mu_D \neq 0 \quad (\text{or } \mu_A \neq \mu_B)$ } (1)

where μ_D is "mean for A − mean for B" (1) for verbal definition of "μ"

Normality of differences. (1)

The test procedure, and the CI in (ii) MUST be PAIRED COMPARISON t test.

Differences are 1·0, 0·4, 1·3, −0·6, 0·9, 0, 2·0, 0·8

$\bar{d} = 0{\cdot}725 \qquad s_{n-1}^2 = 0{\cdot}6364(3),\ s_{n-1} = 0{\cdot}7977(6)$ (A1) Accept $s_n^2 = 0{\cdot}5568(8)$, $s_n = 0{\cdot}7462(4)$, but ONLY if correctly used in sequel.

Test statistic is

$$\frac{0{\cdot}725 - 0}{\frac{0{\cdot}7977}{\sqrt{8}}} = 2{\cdot}57(04)$$ (M1) (M1) (M1) (A1)

Refer to t_7. (1) May be awarded even if test statistic is wrong. No FT if wrong.

Dt 5% point is 2·365. (1) No FT.

Significant. (1)

Seems means differ (and B is better). (1)

[12]

(ii) CI is given by

$$0{\cdot}725 \pm 3{\cdot}499 \frac{0{\cdot}7977}{\sqrt{8}} = 0{\cdot}725 \pm 0{\cdot}987 = (-0{\cdot}262, 1{\cdot}712)$$

(M1) (B1) (M1) (A1)

ZERO out of 4 if not same dist as for test. → **[4]**

(iii) The "pairing" will eliminate differences between technicians. (E2)

To (help to) eliminate any possible "learning" effect, there should be random assignment of the machine used first. (E2)

[4]

Qu 3

(a) Strictly,

let flavour have c.d.f. $F(x)$ for existing process and $F(x-\Delta)$ for new. (1)

H_0 is $\Delta = 0$ (1) H_1 is $\Delta \neq 0$ (1)

If expressed verbally,

distributions have similar shape (1)

H_0 : location-parameters (allow medians, means) equal (1)

H_1 : " " " " " not equal (1) **[3]**

Ranks are

EXISTING	3	5	6	10	11	12
NEW	1	2	4	7	8	9

Rank sum 31 (using "new"; 47 using "existing".
Mann-Whitney, if directly calculated, is 10 or 26) (1)

Refer to tables of Wilcoxon rank-sum (or Mann-Whitney) statistic. (M1)

Lower 5% tail is needed (upper 5% if "existing" used; value for W is 50, for M-W is 29) (1)

Lower 5% value for (6,6) is 28 [7 for M-W]. (1)

31 [or 10 for M-W] is not significant. (1)

Seems on the whole there are no differences between the flavours. (1) **[6]**

"Design" discussion: 1 mark each for any 2 relevant points, eg

- present cups in random order
- adequate time between cups to avoid "carry-over" effect
- ensure appearances do not detract from taste comparisons

(E2) **[2]**

(b) Need

Normality of each population (1)

Same variance (1)

Pooled $s^2 = \dfrac{9 \times 4.62 + 9 \times 6.48}{18} = 5.55$ [or as simple average] (1)

Test statistic is $\dfrac{1.9 - 0}{\sqrt{5.55}\sqrt{\frac{1}{10}+\frac{1}{10}}} = \dfrac{1.9}{1.053(57)} = 1.80(34)$ (M1 A1)

Refer to t_{18} (1) No FT.

Upper 5% pt ≈ 1.73 (1) No FT. ← Must show evidence of interpolation between 1.753 (t_{15}) and 1.725 (t_{20}), and nearer the latter – or a convincing argument that 1.80 is > both.

Significant (1) – seems strength of "new" is greater (1). **[9]**

Qu 4

(i) Observed:

2	8	10
13	2	15
15	10	25

Expected:

6	4
9	6

A1

Without Yates: $X^2 = 11.1$ (1)

With Yates: $X^2 = 8.507$ (1)

Refer to χ^2_1. (1)

Very (1) strong (1) evidence against (1) independence. **7**

(ii) $P(\chi^2_1 > 11.1) = 2P(N(0,1) > \sqrt{11.1} = 3.33) = 2 \times 0.0004 = 0.0008$

$P(\chi^2_1 > 8.507) = 2P(N(0,1) > \sqrt{8.507} = 2.92) = 2 \times 0.0018 = 0.0036$

M1 for general idea, A1, A1 **3**

(iii) We want $\dfrac{\binom{10}{2}\binom{15}{13}}{\binom{25}{15}} = \dfrac{45 \times 105}{3268760} = 0.0014455$ M1 A1 **2**

(iv) Explanation to effect that $\begin{matrix} 1 & 9 \\ 14 & 1 \end{matrix}$ is "further away from expectation" than $\begin{matrix} 2 & 8 \\ 13 & 2 \end{matrix}$, if the rows and columns are independent. E2 **2**

(v) We need to consider all tables which are as extreme or more extreme than the observed table, i.e. M1

$\begin{matrix} 2 & 8 \\ 13 & 2 \end{matrix} \quad \begin{matrix} 1 & 9 \\ 14 & 1 \end{matrix} \quad \begin{matrix} 0 & 10 \\ 15 & 0 \end{matrix} \quad \begin{matrix} 10 & 0 \\ 5 & 10 \end{matrix}$.

Their probabilities under H_0 are

$\binom{10}{2}\binom{15}{13} \Big/ \binom{25}{15} = \dfrac{45 \times 105}{3268760} = 0.001\,445\,5$ M1 award once

$\binom{10}{1}\binom{15}{14} \Big/ \binom{25}{15} = \dfrac{10 \times 15}{3268760} = 0.000\,045\,8$ A1

$\binom{10}{0}\binom{15}{15} \Big/ \binom{25}{15} = \dfrac{1}{3268760} = 0.000\,000\,3$ A1

$\binom{10}{10}\binom{15}{5} \Big/ \binom{25}{15} = \dfrac{1 \times 3003}{3268760} = 0.000\,9186$ A1

∴ level of significance of data = sum of these = 0.0024(107) (1) **6**

Examiner’s Report

2616 Statistics 4

General Comments

There were 132 candidates from 29 centres, a slight decline from last year. The quality of the work was again high with only a handful of candidates being clearly unprepared for the paper. Candidates appeared to have few problems completing the paper in the allotted time, though very few attempted all 4 questions other than in desperation.

Many candidates did not take enough care in stating their hypotheses, particularly in the case of the Wilcoxon rank sum test. These same candidates, in general, were also unaware of the assumptions required to carry out the various teats.

One sign of real improvement was where candidates were asked to make comments on the design of experiments. This was thoughtfully done by the great majority of candidates.

Comments on Individual Questions

Q.1 This, as usual, proved to be the least popular question and was attempted by 48 candidates. For most candidates who attempted this question it was clearly familiar territory and frequently full marks were gained for part (i). A significant number of candidates clearly knew that the required mean and variance were $\frac{1}{\lambda}$ and $\frac{1}{\lambda^2}$ and were determined to obtain these results even though their integration was not quite up to the task.

Virtually all candidates answered part (ii) correctly.

In part (iii) most candidates were not quite sure how to demonstrate that an estimator was plausible, but clearly had an intuitive grasp of the meaning. Explanations here were almost invariably weak.

In part (iv) most candidates knew some of the desirable properties of $\tilde{\theta}$ - unbiased, small for all n, decreasing as n increases, but few mentioned all 3.

(i) mean = $\frac{1}{\lambda}$, variance = $\frac{1}{\lambda^2}$, (ii) mean = $\frac{1}{\lambda} + \theta$ variance = $\frac{1}{\lambda^2}$, (iv) mean = θ variance = $\frac{1}{9n}$.

Q.2 All but one candidate attempted this question and for most it proved a good source of marks. Thankfully very few students did not realise that a paired test was required in part (i). Although most candidates carried out the correct test, many stated their hypotheses either incorrectly or incompletely. Hypotheses need to be in terms of population parameters (rather than sample statistics) and the terms used need to be defined – they are not given in the question. Most realised that the distributional assumption concerned normality, but many did not realise that it was the normality of the differences that was required.

Part (ii) was generally done very well and those candidates who had been largely successful in part (i) had few problems here.

In part (iii) most candidates were able to explain that a paired test enabled the removal of the variability of technicians, although some felt that the ability to use a paired t-test was reason enough in itself. Many went on to give a second feature which should be incorporated – usually that technicians should be assigned at random to the machine that they would use first, but other sound ideas were also put forward.

(i) $t_7 = 2.57$, significant, (ii) (-0.262, 1.712).

Q.3 This was another very popular question, being answered by all but 3 candidates. The hypotheses were not stated clearly by the majority of candidates who merely stated that the two processes were the same, on the whole, for the null hypothesis and that they were not, for the alternative hypothesis. What is required is that the distributions have the same shape and that for the null hypothesis the location parameters are equal with the location parameters not equal for the alternative hypothesis.

The calculation of the test statistic was usually well done, with the majority preferring to use the Mann-Whitney formulation. Again most candidates found the critical value correctly and were able to make the correct decision that the result was not significant. Some candidates then failed to relate this decision in the context of the question.

Most candidates were able to make at least one sensible comment about the design of the experiment and many gave two.

In part (b) a lack of knowledge about the required assumptions was apparent for many candidates. In general candidates knew how to calculate the pooled variance, but a significant number decided to square the given values of the sample variances in their calculation and there was the inevitable confusion between 9 and 10 for some candidates.

A pleasing majority calculated the test statistic correctly (for their figures) and realised that t_{18} was appropriate. Most were then able either to interpolate between t_{15} and t_{20} or give some other convincing argument that the result was significant.

(a) rank sum =31, critical value = 28, not significant; (b) t_{18}= 1.80, significant

Q.4 This question was answered by 89 candidates and proved a good source of marks for most. However, few candidates scored very highly in this question.

In part (i) all candidates were able to calculate the expected values correctly and the vast majority able to calculate the χ^2 statistic correctly without Yates' correction. However the introduction of Yates' correction led to many errors, usually associated with 0.5 being subtracted incorrectly in almost every conceivable way. Most candidates also realised that there was one degree of freedom and that the result was significant at all tabulated levels of significance. Only a handful of candidates thought that this implied independence.

Only the best candidates were able to cope with part (ii) and this part was often left blank.

Parts (iii) and (iv) however were both done successfully by the majority of candidates.

Part (v) again sorted out the best candidates. Only a minority realised that they needed to add together the extreme values to give a significance level and of those it was rare to see $\begin{array}{cc|} 2 & 8 \\ 13 & 2 \end{array}$ included.

(i) χ^2 = 8.51 with Yates and 11.11 without Yates, significant in both cases,
(ii) significance level = 0.0036 with Yates and 0.0008 without Yates, (v) level of significance = 0.0024.