OCR
RECOGNISING ACHIEVEMENT

General Certificate of Education
Advanced Supplementary (AS) and Advanced Level

former Oxford and Cambridge modular syllabus

MEI STRUCTURED MATHEMATICS 5517

Statistics 5

Tuesday **12 JUNE 2001** Afternoon 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 probability density function of the random variable X having the χ_n^2 distribution is

$$f(x) = Kx^{\frac{1}{2}(n-2)}e^{-\frac{1}{2}x}$$

for $x \geqslant 0$, where n is any positive integer and K is a constant (dependent on the value of n).

(i) Show that the moment generating function of X is

$$M(\theta) = K\int_0^\infty x^{\frac{1}{2}(n-2)}e^{-x\left(\frac{1}{2}-\theta\right)}\,dx.$$ [1]

(ii) By making the substitution $x\left(\frac{1}{2}-\theta\right) = \frac{1}{2}u$ and reconsidering the form of $f(x)$, show that

$$M(\theta) = \frac{1}{(1-2\theta)^{\frac{1}{2}n}}.$$

Explain why $M(\theta)$ is only valid for $\theta < \frac{1}{2}$. [7]

(iii) Using results about moment generating functions, which should be carefully stated,

(*A*) show that the sum of q independent random variables each having the χ_1^2 distribution is the random variable having the χ_q^2 distribution,

(*B*) show that the sum of two independent random variables, one having the χ_m^2 distribution and the other the χ_n^2 distribution, is the random variable having the χ_{m+n}^2 distribution. [12]

2 A certain type of domestic security system is specified as operating correctly for at least 2000 hours with probability at least 0.8. Trading standards officers inspect a random sample of 50 of these systems and find that 15 of them have ceased to operate correctly after less than 2000 hours.

Examine, at the 10% significance level, whether the specification is being met, stating clearly your null and alternative hypotheses. [13]

Provide a two-sided 95% confidence interval for the true proportion of the systems that operate correctly for at least 2000 hours. [7]

3 (i) $X_1, X_2, \ldots, X_n$ are independent random variables, all with the distribution $N(\mu_1, \sigma_1^2)$. The random variables $\overline{X}$ and S_1^2 are defined by

$$\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i,$$

$$S_1^2 = \frac{1}{n-1}\sum_{i=1}^{n}\left(X_i - \overline{X}\right)^2.$$

Similarly, $Y_1, Y_2, \ldots, Y_m$ are independent random variables, all with the distribution $N(\mu_2, \sigma_2^2)$, and $\overline{Y}$ and S_2^2 are defined by

$$\overline{Y} = \frac{1}{m}\sum_{j=1}^{m} Y_j,$$

$$S_2^2 = \frac{1}{m-1}\sum_{j=1}^{m}\left(Y_j - \overline{Y}\right)^2.$$

(*A*) State the distribution of S_1^2. [2]

(*B*) State the distribution of $\dfrac{S_1^2}{S_2^2}$ for the case $\sigma_1^2 = \sigma_2^2$. [2]

(ii) It is claimed that the variance of the diameter of a machined engine part is no more than 0.0004 cm^2. Inspectors check a random sample of 13 of these parts and find that their diameters (in cm) are as follows.

2.00 2.01 2.03 1.96 1.97 2.00 1.98 1.96 2.03 2.04 2.00 2.02 1.97

Stating carefully your null and alternative hypotheses, carry out an appropriate test at the 5% level of significance. Provide also a two-sided 95% confidence interval for the population variance. Underlying Normality may be assumed. [11]

(iii) A competing company makes similar parts. A random sample of 15 of these gives a value of S^2, defined as in part **(i)**, of 0.00028 cm^2. Again assuming underlying Normality, use this and the data in part **(ii)** to test at the 5% level whether the population variance for this competing company is smaller than that for the original company. [5]

4 The random variable X is Normally distributed with mean μ and variance 9, i.e.

$$X \sim N(\mu, 9).$$

A random sample of size $n = 16$ is available.

(a) The null hypothesis $H_0 : \mu = 0$ is to be tested against the alternative hypothesis $H_1 : \mu = 2$.

The probability of a type I error (i.e. of rejecting H_0 when in fact it is true) is denoted by α. Similarly, the probability of a type II error is denoted by β.

(i) Show that the usual one-sided 5% significance test based on the sample mean $\bar{X}$ states that H_0 is to be rejected if the value of $\bar{X}$ is greater than 1.234 (to 3 decimal places). Write down the value of α for this test and show that the value of β is 0.154. [6]

(ii) An alternative test is proposed, in which H_0 is to be rejected if the value of $\bar{X}$ is greater than 1. Find the values of α and β for this test. [4]

(iii) Use the values of α and β to comment briefly on these two tests. [2]

(b) The null hypothesis $H_0 : \mu = 0$ is to be tested against the alternative hypothesis $H_1 : \mu > 0$.

(i) The usual one-sided 5% significance test again rejects H_0 if the value of $\bar{X}$ is greater than 1.234. Show that an expression for the operating characteristic of this test is

$$P\left(Z \leqslant \tfrac{4}{3}(1.234 - \mu)\right)$$

where $Z \sim N(0, 1)$. [2]

(ii) Find a corresponding expression for the operating characteristic of the test which rejects H_0 if the value of $\bar{X}$ is greater than 1. [2]

(iii) You are *given* the following values of the operating characteristics for a selection of values of μ; *no further calculations are required.*

μ	Value of operating characteristic	
	Test in (b) (i)	Test in (b) (ii)
−1	0.9986	0.9962
0	0.95	0.9087
1	0.6225	0.5
2	0.1536	0.0913
3	0.0093	0.0038

Use the values of the operating characteristics to comment briefly on these two tests. [4]

Mark Scheme

June 2001 — 5517 Statistics 5

Q.1	$\chi_n^2 : \mathrm{f}(x) = Kx^{\frac{n-2}{2}} e^{-\frac{x}{2}}$		
	(i) $M(\theta) = \mathrm{E}\left[e^{\theta x}\right] = \int_0^\infty e^{\theta x} . Kx^{\frac{n-2}{2}} e^{-\frac{x}{2}} dx$	1	1
	(ii) $\ldots = K \int_0^\infty x^{\frac{n-2}{2}} e^{-x\left(\frac{1}{2}-\theta\right)} dx$ $x\left(\frac{1}{2}-\theta\right) = \frac{1}{2}u$		
	$= K \int_0^\infty \left(\frac{\frac{1}{2}u}{\frac{1}{2}-\theta}\right)^{\frac{n-2}{2}} e^{-\frac{1}{2}u} \frac{\frac{1}{2}du}{\frac{1}{2}-\theta}$ For correct substitution (may be subdivided if partially correct)	2	
	$= \frac{\left(\frac{1}{2}\right)^{\frac{n}{2}}}{\left(\frac{1}{2}-\theta\right)^{\frac{n}{2}}} \int_0^\infty Ku^{\frac{n-2}{2}} e^{-\frac{u}{2}} du$ For re-writing in this form	1	
	(integral) $= 1$	1	
	$= \frac{1}{(1-2\theta)^{\frac{n}{2}}}$ BEWARE PRINTED ANSWER	1	
	If $\theta = \frac{1}{2}$, integral becomes $\int_0^\infty x^{\frac{n-2}{2}} dx$	E1	
	[allow 'substitution collapses' or 'would divide by zero']		
	If $\theta > \frac{1}{2}$, the integrand would have e raised to a positive power (becoming infinite at the upper limit).		
	[allow comment that the substitution would lead to $\int_0^{-\infty} du$]	E1	7
	(iii) Required results are		
	• (convolution theorem) mgf of sum of indep random variables is product of their sep mgfs May be subdivided if statements are not fully correct	2	
	• uniqueness of mgf $\leftrightarrow$ distribution relationship	2	
	A Consider mgf of $\chi_1^2 + \chi_1^2 \ldots + \chi_1^2$	M1	
	$= \frac{1}{(1-2\theta)^{\frac{1}{2}}} \times \ldots \times \frac{1}{(1-2\theta)^{\frac{1}{2}}}$	1	
	$= \frac{1}{(1-2\theta)^{\frac{q}{2}}}$	1	
	mgf of χ_q^2	1	
	Hence result		
	B Consider mgf of $\chi_m^2 + \chi_n^2$	M1	
	$= \frac{1}{(1-2\theta)^{\frac{m}{2}}} \times \frac{1}{(1-2\theta)^{\frac{n}{2}}} = \frac{1}{(1-2\theta)^{\frac{m+n}{2}}}$	1	
	= mgF of χ_{m+n}^2	1	
	Hence result		
	Award (1) ONCE in part A or part B	1	12

Q.2	$H_0 : p = 0.2$ OR $H_0 : p = 0.8$	1	
	$H_1 : p > 0.2$ OR $H_1 : p < 0.8$	1	
	Where p = P (system not OK for at least 2000 hours) if 0.2 used Where p = P (system OK for at least 2000 hours) if 0.8 used	2	4
	15 out of 50 fail before 2000 hours		
	Test statistic is		
	$\frac{14\frac{1}{2} - 10}{\sqrt{50 \times 0.2 \times 0.8}} = \frac{4.5}{\sqrt{8}} = 1.591$ or $\frac{0.29 - 0.2}{\sqrt{\frac{0.2 \times 0.8}{50}}} = \frac{0.9}{\sqrt{0.0032}} = 1.591$ If $p = 0.2$ used If **all** correct	M4 A1	
	OR		
	$\frac{35\frac{1}{2} - 40}{\sqrt{50 \times 0.8 \times 0.2}} = \frac{-4.5}{\sqrt{8}} = -1.591$ or $\frac{0.71 - 0.8}{\sqrt{\frac{0.8 \times 0.2}{50}}} = \frac{-0.09}{\sqrt{0.0032}} = -1.591$ If $p = 0.8$ used		
	Accept 15 (or 0.3) [OR 35 (or 0.7)] for M2 and FT (value = ± 1.7678). Do NOT accept $15\frac{1}{2}$ [or $34\frac{1}{2}$]. Do NOT accept $10 - 14\frac{1}{2}$ [or $40 - 35\frac{1}{2}$], unless CLEARLY explained in sequel.		
	Compare N(0, 1) No FT	1	
	Critical point is 1.282 if $p = 0.2$ used or −1.282 if $p = 0.8$ used No F.T., SIGN must be CORRECT	1	
	Significant	1	
	Seems specification is not being met	1	
	SPECIAL CASES: allow 1 of last 3 marks if work is correct after ▪ Quoting critical points as ±1.282 ▪ -1.591 calculated but then +1.591 referred to +1.282 (or the same with all signs reversed) ▪ stating that critical region is given by $\lvert z \rvert > 1.282$		9
	CI is given by		
	0.7	M2	
	± 1.96	B1	
	$\sqrt{\frac{0.3 \times 0.7}{50}}$	M2	
	$= 0.7 \pm 1.96\sqrt{0.0042} = 0.7 \pm 1.96 \times 0.0648 = 0.7 \pm 0.127 = (0.573, 0.827)$	A1 A1 cao	7

Q.3	(i)	A $S_1^2 \sim \frac{\sigma_1^2}{n-1}\chi^2_{n-1}$ Four key points: if all correct [Allow 1 if any three correct]	2	2
		B $\frac{S_1^2}{S_2^2} \sim F_{n-1,m-1}$ [1 or F, 1 for correct d.f.]	2	2
	(ii)	$n = 13$ $s^2 = 0.0007692$ [if used, $s_n^2 = 0.00071$]		
		$H_0 : s^2 = 0.0004$	1	
		$H_1 : s^2 > 0.0004$	1	
		Test statistic is $\frac{12 \times 0.0007692}{0.0004} = 23.08$	M1 A1	
		Refer to χ^2_{12}	1	
		Upper 5% pt is 21.03	1	
		Significant and suggests variance is too great	1	
		CI is given by [M1 is for general form of CI] 4.404	M1 B1	
		$< \frac{12 \times 0.0007692}{\sigma^2} < 23.34$	B1	
		$\rightarrow 0.0003954 < \sigma^2 < 0.0020959$	A1	11
	(iii)	New s^2 is 0.00028 from 15 observations		
		Consider $\frac{0.0007692}{0.00028} = 2.747$	1	
		And refer to upper tail of $F_{12,\,14}$	1	
		ONE-sided test needed here – upper 5% point is 2.53	1	
		Significant	1	
		Competitor's variance does appear smaller	1	5

Q.4	$X \sim N(\mu, 9)$ Sample of size $n = 16$.		
	(a) $H_0 : \mu = 0$ $H_1 : \mu = 2$		
	(i) one-sided 5% test rejects H_0 if $\frac{\bar{x} - 0}{\frac{3}{4}} > 1.645$	1	
	i.e. if $\bar{x} > \frac{3}{4} \times 1.645 = 1.234$ [1.23375] Beware printed answer	1	
	$\alpha = 0.05$	1	
	$\beta = P(\text{acc } H_0 \mid H_1 \text{ true})$	M1	
	$= P(\bar{X} < 1.234 \mid \mu = 2$ i.e. $\bar{X} \sim N(2, \frac{9}{16}))$	M1	
	$= P(N(0, 1) < \frac{1.234 - 2}{\frac{3}{4}} = -1.021\dot{3}) = 1 - 0.8464 = 0.1536$		
	(printed answer is 0.154)	1	6
	(ii) Reject if $\bar{x} > 1$		
	$\alpha = P(\bar{X} > 1 \mid \bar{X} \sim N(0, \frac{9}{16}))$	M1	
	$= P\left(N(0,1) > \frac{1-0}{\frac{3}{4}} = 1.333\right) = 0.0913$	1	
	$\beta = P(\bar{X} < 1 \mid \bar{X} \sim N(2, \frac{9}{16}))$	M1	
	$= P(N(0, 1) < \frac{1-2}{\frac{3}{4}} = -1.333 = 0.0913$	1	
	[accept symmetry argument for β only if *absolutely* convincing]		4
	(iii) Discussion – standard test has smaller probability of wrongly rejecting H_0 but larger probability of wrongly accepting it.	E2	2
	(b) $H_0 : \mu = 0$ $H_1 : \mu > 0$		
	(i) OC is $P(\text{acc } H_0 \mid \mu)$ i.e. $P(\bar{X} < 1.234 \mid \bar{X} \sim N(\mu, \frac{9}{16}))$	M1	
	$= P(N(0, 1) < \frac{1.234 - \mu}{\frac{3}{4}})$ Beware printed answer	1	2

μ	z	OC
-1	2.9787	0.9986
0	1.645	0.95
1	0.312	0.8225
2	$-1.021\dot{3}$	0.1536
3	$-2.354\dot{6}$	0.0093

	(ii) OC is $P(\bar{X} < 1 \mid \bar{X} \sim N(\mu, \frac{9}{16}))$	M1	
	$= P(N(0, 1) < \frac{1 - \mu}{\frac{3}{4}})$	1	

μ	z	OC
-1	2.667	0.9962
0	1.333	0.9087
1	0	0.5
2	-1.333	0.0913
3	-2.667	0.0038

			2
	(iii) Discussion – standard test has higher OC for all calculated values of μ (indeed, for all values of μ), i.e. a higher probability of accepting the null hypothesis – which is fine if the null hypothesis is actually true, but poor if it is false.	E4	4

Examiner’s Report

Statistics 5 (5517)

General Comments

There were 53 candidates from 18 centres – fewer centres but more candidates than last year.

There was a lot of good work. It is pleasing to see that there are able candidates in the system at this level. Inevitably there was also some work of a lower standard, indeed rather poor in places. But this should not be allowed to obscure the confident message of the good work.

Questions 2 and 3 on the paper were very popular. Questions 1 and 4 were then about equally popular some way behind. A very few candidates attempted all four questions; as usual, all their work was marked and the best three questions were counted.

Comments on Individual Questions

Question 1 (generating functions)

This question used moment generating functions to explore some properties of the chi-squared distribution. Parts (i) and (ii) obtained the moment generating function of the general chi-squared distribution, the answer being given and the candidates in effect being told the method for getting there. Some candidates were well up to this task, but rather a lot could not manage it; the substitution (given in the question) was sometimes poorly handled, and some candidates embarked on utterly fruitless attempts to integrate by parts. Candidates then had to discuss why the moment generating function is valid only for $\theta < ½$. Comments here were disappointing. Many candidates said something about the necessity for the function to be positive, which is true but not really the point; hardly anyone saw that the integration would have broken down, and the answer thus not even reached, had θ been greater than ½ (the power in the exponential function would then be positive, but it is taken to an upper limit of ∞). Also, few candidates noted that, if θ were actually equal to ½, the integration would also have broken down due to division by zero.

Part (iii) of the question invited candidates to prove additive properties of chi-squared distributions using the moment generating function. A few candidates floundered without trace here, but most knew near-enough what to do. However, candidates were sometimes less than secure in stating the results on which their work was based. Indeed, the explicit request in the question to use results "which should be carefully stated" was honoured rather more in the breach than in the observance.

Question 2 (test and confidence interval for a binomial proportion)

This question was often done well, though rarely perfectly. Candidates had first to state the hypotheses that were being tested. Rigour was expected here, not least in defining the "p" parameter that would appear. It was necessary that this clearly referred to a population; the word "probability" obviously carried that implication, but where "proportion" was used it had to be qualified by "population" to earn full marks. Proceeding to the test, some candidates based it on testing $p = 0.2$ and others (the large majority) on testing $p = 0.8$; either approach was entirely acceptable. It was disappointingly rare to find a continuity correction. An occasional candidate averred that the large sample rendered this unnecessary – most candidates said nothing at all – but a glance at the values of the test statistic with and without it [± 1.591 and ± 1.768, the sign depending on whether 0.2 or 0.8 is being tested] should convince otherwise. In referring the value to N(0, 1), candidates were expected to be *careful* in respect of sign, as this is a one-sided test. Thus, if 0.8 is being tested, the critical point for comparison is -1.282 and not $+1.282$ (or just 1.282). Finally, the confidence interval was often done very well, and it is pleasing to see that difficulties in some previous years with this work seem to have largely vanished. A modest number of candidates, however, still wrongly used the null hypothesis value of p (0.2 or 0.8) in the standard deviation rather than the observed sample proportion (0.3 or 0.7). A few candidates were also careless in not seeing that the interval *must* be centred on 0.7 and not 0.3. The t_{50} distribution appeared once, but only once. It might be a record for this error to have occurred so few times, but let us strive for a new record on the next occasion! [The interval is (0.573, 0.827).]

Question 3 (test and confidence interval for variance using chi-squared; F test for comparing variances)

This question was often *very* well done; many solutions were essentially perfect from beginning to end. The most common error was a strange one in part (iii), where an F test was used but based on the *sample* variance for the competing company taken with the hypothesised *population* variance for the original company. This seemed all the stranger as the hypothesised value had just been rejected!

Sadly, there remained some candidates who tied themselves into terrible knots by insisting on working with the "n" version of the sample variance, despite all the definitions given in the question. In nearly all cases, these candidates did not "correct" it properly (or at all) when forming the test statistics; even when they did, their work was often unnecessarily convoluted. It is quite upsetting for the examiner to see candidates getting themselves into such difficulties that are wholly unnecessary.

[Part (ii): value of test statistic is 23.08, 12 degrees of freedom, critical point 21.03; end-points of confidence interval are 0.00039(5) and 0.00209(6). Part (iii): value of test statistic is 2.747, degrees of freedom are 12 and 14, critical point is 2.53.]

Question 4 (type I error, type II error, operating characteristic)

Candidates seemed rather more secure with this work than in previous years. Intermediate answers were quite liberally given in the question, but this time the work that led to them was generally more robust. There remained, however, some candidates who can only be said to have got there by hook or by crook! Still, it was good to see that the values of α and β in part (a)(ii), which were not given in the question, were usually obtained correctly [0.0913 for each of them]. The consequent discussion of the merits of the tests in (a)(i) and (a)(ii) often produced comments that the equality of α and β for the (a)(ii) test was in itself meritorious, in some sense minimising the total risk. While there is perhaps something in this argument, it obscures the comparison that the (a)(ii) test is better than the (a)(i) test in respect of Type II error but worse for Type I error.

In part (b), candidates moved on to considering operating characteristics. This was usually done quite well, maybe aided by the given answer for (b)(i). The answer for (b)(ii) was often given as a direct write-down [$P(Z) \leq (4/3)(1 - \mu)$], presumably by direct comparison with the answer to (b)(i). The discussions of the operating characteristics varied in their quality. Most candidates saw that the operating characteristic of the (b)(ii) test is everywhere less than that of the (b)(i) test, but not everyone could interpret this, or at least not fully, in respect of their relative merits.