OCR
RECOGNISING ACHIEVEMENT

ADVANCED SUBSIDIARY GCE

MEI STATISTICS

G242

Statistics 2 (Z2)

Candidates answer on the Answer Booklet

OCR Supplied Materials:
- 8 page Answer Booklet
- Graph paper
- MEI Examination Formulae and Tables (MF2)

Other Materials Required:
- Scientific or graphical calculator

Wednesday 9 June 2010
Afternoon

Duration: 1 hour 30 minutes

INSTRUCTIONS TO CANDIDATES

- Write your name clearly in capital letters, your Centre Number and Candidate Number in the spaces provided on the Answer Booklet.
- Use black ink. Pencil may be used for graphs and diagrams only.
- Read each question carefully and make sure that you know what you have to do before starting your answer.
- Answer **all** the questions.
- Do **not** write in the bar codes.
- You are permitted to use a graphical calculator in this paper.
- Final answers should be given to a degree of accuracy appropriate to the context.

INFORMATION FOR CANDIDATES

- The number 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.
- The total number of marks for this paper is **72**.
- This document consists of **4** pages. Any blank pages are indicated.

1 A birdwatcher has learned to recognise different species of birds by their song. He notices that three particular types of warbler regularly found in his 'patch' prefer to sing in trees. He decides to investigate whether there is any association between the type of warbler and the type of tree from which they are heard singing. 200 warblers, regarded as a random sample, are selected and the numbers of warblers in each category are summarised in the table below.

		Tree		
		Willow	Birch	Oak
Warbler	Chiffchaff	10	13	20
	Willow Warbler	39	43	12
	Whitethroat	24	20	19

(i) A test is to be carried out to examine whether these data provide any evidence of an association between these classification factors. State clearly the null and alternative hypotheses. The following tables show some of the expected frequencies and contributions to the test statistic. Calculate the remaining expected frequencies and contributions. Carry out the test at the 5% level of significance. **[11]**

Expected frequencies		Tree		
		Willow	Birch	Oak
Warbler	Chiffchaff	15.695	16.340	10.965
	Willow Warbler	34.310		
	Whitethroat	22.995		

Contributions to the test statistic		Tree		
		Willow	Birch	Oak
Warbler	Chiffchaff	2.0665	0.6827	7.4447
	Willow Warbler	0.6411		
	Whitethroat	0.0439		

(ii) For each type of warbler, comment briefly on how its distribution compares with what would be expected if there were no association. **[3]**

(iii) While out for a walk, the birdwatcher hears the song of a whitethroat. Use the given data to estimate the probability that it is singing from a birch tree. **[2]**

2 A doctor working in a hospital in a poor area of a large city is concerned about the low average birth weight of babies born in the hospital. For babies born in this hospital, the mean birth weight is 2800 grams, which is well below the ideal birth weight. The doctor introduces an extensive prenatal care programme in an attempt to increase the mean birth weight. Following the introduction of the programme, the doctor measures the birth weight of each of a random sample of 12 babies born in the hospital, with results in grams as follows.

2430 2720 2910 3000 3230 2840 2660 3350 3210 2870 2820 3540

(i) Explain why, in this situation, it would not be appropriate to carry out a hypothesis test for a population mean using the Normal distribution. State the assumption necessary for a test based on the t distribution to be valid. **[3]**

(ii) Use these data to estimate the population mean and the population standard deviation. **[3]**

(iii) Use a t test to examine at the 5% significance level whether this sample provides evidence that the prenatal care programme has been successful in increasing the mean birth weight of babies born in this hospital. State your null and alternative hypotheses clearly. **[10]**

3 A regional highway authority is concerned about the high numbers of accidents involving cyclists at roundabouts. A random sample of 150 roundabouts is selected, and the number of accidents involving cyclists at each of these roundabouts over a four-week period is recorded. The results are shown in the following frequency table.

Number of accidents, x	0	1	2	3	4	5	6	$\geqslant 7$
Observed frequency, f	21	36	26	24	23	12	8	0

(i) The sample standard deviation is 1.734, correct to 3 decimal places.

(A) Verify that the sample mean number of accidents is 2.4. **[2]**

(B) Do these statistics give you any reason to doubt the belief that the number of accidents may be modelled using a Poisson distribution? Justify your answer. **[2]**

(ii) The highway authority wishes to carry out a test of the goodness of fit of the Poisson model. The sample mean of 2.4 is used as an estimate of the mean of the underlying population. The following tables show some of the expected frequencies and corresponding contributions to the test statistic. Use the appropriate cumulative probability tables to find the remaining expected frequencies, and calculate the remaining contributions. Carry out the test at the 5% level of significance. **[10]**

Expected frequencies

Number of accidents, x	0	1	2	3	4	5	$\geqslant 6$
Expected frequency	13.605		39.195	31.350	18.810	9.030	

Contributions to the test statistic

Number of accidents, x	0	1	2	3	4	5	$\geqslant 6$
Contribution	4.0196	0.3426			0.9333	0.9768	1.3064

4 As part of a research project involving a particular colony of common seals, a biologist is investigating the length of time that seals spend under water each time they dive. The dive durations, in seconds, for a random sample of 10 adolescent seals are as follows.

243 251 218 227 205 232 198 224 187 264

Over a period of time, the biologist has found that, for this particular seal colony, the median length of dive is 210 seconds. Use a Wilcoxon test to examine, at the 5% significance level, whether the sample provides evidence of a difference between the median dive duration of these adolescent seals and that of the seal colony as a whole. State your null and alternative hypotheses clearly. **[13]**

5 A large brewery supplies beer in bottles labelled as containing 500 ml. The bottles are filled by machine. The random variable X represents the volume of beer, in ml, delivered to each bottle. X is Normally distributed with mean μ and standard deviation 1.29. The value of μ can be adjusted by a machine operator.

(i) Given that $\mu = 502$, find $P(X < 500)$. **[3]**

(ii) Find the value of μ needed to ensure that 1% of bottles filled by the machine contain less than 500 ml. **[3]**

The brewery also sells beer in casks labelled as containing 9 gallons. During one month, a random sample of 40 casks is selected. The sample mean volume of beer is 9.05 gallons and the sample standard deviation is 0.06 gallons.

(iii) Find a two-sided 95% confidence interval for the mean volume of beer per cask. **[5]**

(iv) The brewery aims to avoid the mean volume being less than the advertised 9 gallons. Comment on this, using the confidence interval found in part **(iii)** to support your answer. **[2]**

GCE

Statistics (MEI)

Advanced Subsidiary GCE **G242**

Statistics 2 (Z2)

Mark Scheme for June 2010

This mark scheme is published as an aid to teachers and students, to indicate the requirements of the examination. It shows the basis on which marks were awarded by Examiners. It does not indicate the details of the discussions which took place at an Examiners' meeting before marking commenced.

All Examiners are instructed that alternative correct answers and unexpected approaches in candidates' scripts must be given marks that fairly reflect the relevant knowledge and skills demonstrated.

Mark schemes should be read in conjunction with the published question papers and the Report on the Examination.

OCR will not enter into any discussion or correspondence in connection with this mark scheme.

Any enquiries about publications should be addressed to:

OCR Publications
PO Box 5050
Annesley
NOTTINGHAM
NG15 0DL

Telephone: 0870 770 6622
Facsimile: 01223 552610
E-mail: publications@ocr.org.uk

Q1			
(i)	H_0: there is no association between warbler and tree H_1: there is an association between warbler and tree	B1	
	Expected frequencies <table><tr><td></td><td>Willow</td><td>Birch</td><td>Oak</td></tr><tr><td>Chiffchaff</td><td>15.695</td><td>16.340</td><td>10.965</td></tr><tr><td>Willow Warbler</td><td>34.310</td><td>**35.720**</td><td>**23.970**</td></tr><tr><td>Whitethroat</td><td>22.995</td><td>**23.940**</td><td>**16.065**</td></tr></table>	M1 A1	
	Contributions to X^2 <table><tr><td></td><td>Willow</td><td>Birch</td><td>Oak</td></tr><tr><td>Chiffchaff</td><td>2.0665</td><td>0.6827</td><td>7.4447</td></tr><tr><td>Willow Warbler</td><td>0.6411</td><td>**1.4837**</td><td>**5.9775**</td></tr><tr><td>Whitethroat</td><td>0.0439</td><td>**0.6484**</td><td>**0.5362**</td></tr></table>	M1 A1	
	$X^2 = 19.525$	A1	
	4 degrees of freedom Critical value for 5% significance level is 9.488 As 19.525 > 9.488 the result is significant	B1 B1 M1 A1	
	There is evidence of an association between the warbler and tree.	A1	**11**
(ii)	Chiffchaffs occurred more frequently than expected in Oak trees. Willow Warblers occurred less frequently than expected in Oak trees. Whitethroat occurred more or less as expected.	E1 E1 E1	**3**
(iii)	P(Birch|Whitethroat) = 20/63	M1 A1	**2**
			16

Q2			
(i)	This is a small sample The variance is unknown We must assume birth weights are Normally distributed	B1 B1 B1	**3**
(ii)	Estimate for population mean = 2965 g Estimate for population standard deviation $= \sqrt{\dfrac{106593000 - \dfrac{35580^2}{12}}{11}}$ = 315.983… = 316 to 3 sf	B1 M1 A1 CAO	**3**
(iii)	$H_0 : \mu = 2800$ & $H_1 : \mu > 2800$ Where μ represents the population mean birth weight of babies born after the introduction of the prenatal care programme. $t = \dfrac{2965 - 2800}{SD/\sqrt{12}}$ = 1.809 (using SD = 316) 11 degrees of freedom At 5% level, critical value of t is 1.796 1.809 > 1.796 so the result is significant. Evidence suggests the mean birth weight has increased.	B1 B1 B1 M1 A1 CAO B1 B1 M1A1 A1	**10**
			16

Q3			
(i)*A*	$\sum fx \div \sum f = 360 \div 150$ (= 2.4 A.G.)	M1 A1	**2**
B	Variance = $1.734^2 = 3.0067\ldots$, which seems close to the mean value of 2.4. A Poisson model may be appropriate.	B1 E1(compare mean with variance – allow arguments either way, with relevant conclusion)	**2**
(ii)	H_0: The Poisson model is suitable $P(X = 1) = 0.2177$ & $P(X \geq 6) = 0.0357$ Missing expected frequencies are 32.655 ($x = 1$), and 5.355 ($x \geq 6$) Missing contributions are 4.4421 ($x = 2$) and 1.7232 ($x = 3$) $X^2 = 13.7441$ There are 7 – 1 – 1 = 5 degrees of freedom. At the 5% significance level the critical value is 11.07 The result is significant Evidence suggests that the Poisson model is inappropriate.	B1 (both probabilities) M1 A1 (expected freq) M1 A1 A1 B1 B1 B1 B1	**3** **3** **4**
			14

Q4			
	H_0: population median = 210 H_1: population median ≠ 210	B1 B1	**2**
	Actual differences 33 41 8 17 -5 22 -12 14 -23 54 Associated ranks 8 9 2 5 1 6 3 4 7 10 $T^- = 1 + 3 + 7 = 11$ $T^+ = 8 + 9 + 2 + 5 + 6 + 4 + 10 = 44$ $\therefore T = 11$	B1 M1 A1 B1 B1 B1	**6**
	From $n = 10$ tables – at the 5% level of significance in a two-tailed Wilcoxon single sample test, the critical value of T is 8 $11 > 8$ $\therefore$ the result is not significant The evidence does not suggest that there is a difference between the median dive duration of adolescent seals and the seal population as a whole.	M1 (use of $n = 10$ in tables) A1 M1 A1 E1	**5**
			13

Q5			
(i)	$P(X < 500) = P(Z < \frac{500-502}{1.29}) = P(Z < -1.550)$ $1 - \Phi(1.550) = 1 - 0.9394 = 0.0606$(awrt 0.061)	M1 standardising M1 correct tail A1	**3**
(ii)	From tables $\Phi^{-1}(0.99) = 2.326$ $\frac{500-\mu}{1.29} = -2.326$ $\mu = 500 + 2.326 \times 1.29 = 503$	B1 for 2.326 seen M1 for equation in μ and negative z-value A1	**3**
(iii)	$9.05 \pm 1.96 \times \frac{0.06}{\sqrt{40}}$ (9.03, 9.07)	B1 centred on 9.05 B1 for 1.96 M1 structure A1 A1	**5**
(iv)	As the lower limit of the interval in part (iii) is more than 9 gallons, this does not suggest that the mean volume is below 9 gallons for this month. Allow sensible alternatives	E1 E1	**2**
			13

OCR (Oxford Cambridge and RSA Examinations)
1 Hills Road
Cambridge
CB1 2EU

OCR Customer Contact Centre

14 – 19 Qualifications (General)
Telephone: 01223 553998
Facsimile: 01223 552627
Email: general.qualifications@ocr.org.uk

www.ocr.org.uk

For staff training purposes and as part of our quality assurance programme your call may be recorded or monitored

Chief Examiner's Report

The Principal Examiners' reports that follow discuss the candidates' performances on the individual modules. There is one matter that should be discussed in a general way as it applies to all the statistics modules. This is in respect of arithmetical accuracy in intermediate working and in quotation of final answers.

Most candidates are sensible in their arithmetical work, but there is some unease as to exactly what level of accuracy the examiners are expecting. There is no general answer to this! The standard rubric for all the papers sums the situation up by including "final answers should be given to a degree of accuracy appropriate to the context". Three significant figures may often be the norm for this, but this always needs to be considered in the context of the problem in hand. For example, in quoting from Normal tables, *some* evidence of interpolation is generally expected and so quotation to four decimal places will often be appropriate. But even this does not always apply – quotations of the standard critical points for significance tests such as 1.96, 1.645, 2.576 (maybe even 2.58 – but not 2.57) will commonly suffice.

Talking now in general terms, the examiners always exercise sensible discretion in cases of small variations in the degree of accuracy to which an answer is given. For example, if 3 significant figures are expected (either because of an explicit instruction or because the general context of a problem demands it) but only 2 are given, a candidate is likely to lose an Accuracy mark; but if 4 significant figures are given, there would normally be no penalty. Likewise, answers which are slightly deviant from what is expected in a very minor manner are not penalised (for example, a Normal probability given, after an attempt at interpolation, as 0.6418 whereas 0.6417 was expected). However, there are increasing numbers of cases where candidates give answers which are *grossly* over- or under-specified, such as insistence that the value of a test statistic is (say) 2.128888446667 merely because that is the value that happens to come off the candidate's calculator. **Such gross over-specification indicates a lack of appreciation of the nature of statistical work and, with effect from the January 2011 examinations, will be penalised by withholding of associated Accuracy marks.**

Candidates must however always be aware of the dangers of premature rounding if there are several steps in a calculation. If, say, a final answer is desired that is correct to 3 decimal places, this can in no way be guaranteed if only 3 decimal places are used in intermediate steps; indeed, it may not be safe to carry out the intermediate work even to 4 decimal places. The issue of over-specification may arise for the final answer but not for intermediate stages of the working.

It is worth repeating that most candidates act sensibly in all these respects, but it is hoped that this note may help those who are perhaps a little less confident in how to proceed.

G242 Statistics 2

General comments

This year saw another small entry, similar in size to last year. The majority of this year's candidates were very well prepared and many high marks were produced.

Overall, the candidates demonstrated very good understanding of the statistical methods required and communicated their responses using appropriate statistical terms and in sufficient detail. The parts of questions requiring candidates to interpret information, explain or comment were not as well answered as the parts involving calculation. Some candidates lost marks through incorrect use of their calculator; there were several cases where a correct method was seen but the final answer did not match what was written. Problems identifying the correct number of degrees of freedom were again common.

Comments on individual questions

1) *(Chi-squared test for Association)*

(i) Some candidates mixed up the hypotheses, leading to contradictory conclusions and loss of marks. Some candidates did not include context in either their hypotheses or in their concluding remarks. A few slips with degrees of freedom were seen and incorrect critical values were fairly common. It is expected that candidates should state the number of degrees of freedom used – some did not and were penalised.

(ii) This part was poorly understood. Few candidates showed an understanding of the link between the size of the contribution to the test statistic and the level of association. For willow warblers and chiffchaffs, candidates were expected to identify the cells containing relatively large contributions and comment whether this provided evidence that the warblers were seen more frequently or less frequently than expected in the corresponding tree. For whitethroats, the candidates were expected to comment that the small contributions indicated that they occurred in numbers that would be expected if there were no association between warbler and type of tree.

(iii) This too was poorly answered. Generally, candidates interpreted the question incorrectly, not realising the importance of the condition that the bird heard was a whitethroat. A small number reversed the question, finding the probability that the bird was a whitethroat given that it was singing from a birch tree.

2) *(Hypothesis test using the t distribution)*

(i) This required an understanding of the differences between the situations leading to hypothesis tests based on the Normal distribution and the t distribution. In general, this was not well answered. Several candidates did not comment on the assumption necessary for a t test to be valid. Confusion between population and sample was evident.

(ii) This required candidates to provide estimates for population mean and population standard deviation. This led to full marks in most cases.

(iii) This part was well answered. Occasional marks were lost for failing to define μ as the population mean or for writing hypotheses in terms of some other variable (e.g. x). Candidates were expected to give hypotheses in terms of μ rather than in words. Several candidates stated a correct calculation for the test statistic but did not calculate it correctly; a final value of 0.15075 was seen more than once. Some candidates were unsure of the value to use for the number of degrees of freedom. Some stated a 2-tail critical value despite intending to use a 1-tailed test. Conclusions were stated in appropriate terms, were not too assertive and were given in the context of the question.

3) *(Chi-squared test for goodness of fit)*

(i) (A) was well answered. Most candidates successfully verified the sample mean as 2.4 using the given frequency distribution.
(B) was poorly answered with several candidates making comments about results being random and/or independent, rather than comparing the mean and variance. Several compared mean and standard deviation and earned no credit.

(ii) In general, candidates could find $P(X = 1)$ but many struggled with $P(X \geq 6)$.
The remainder of the question was well answered – some lost marks by using a critical value from the t distribution and some associated 'significant' with not rejecting the null hypothesis.

4) (*Wilcoxon test)*

This was generally well answered. Candidates were required to provide values in their hypotheses and make it clear that the values referred to the population median. Many lost marks here. In the remainder of the question, marks were lost for providing an incorrect critical value. Some contradictory conclusions were seen (as in Q3 (ii)).

5) *(Use of Normal distribution (confidence interval))*

(i) This was well done.

(ii) This was less well handled, with +2.326 leading to an answer of 497 seen on several occasions. Candidates should be encouraged to sketch diagrams to help ensure sensible answers are found.

(iii) This was well answered although some used 1.645 in place of 1.96.

(iv) In this part, many were not convincing in their explanations. Candidates were expected to point out that the value of 9 (gallons) was 'below' the confidence interval – stating that it was 'not contained in' the confidence interval was deemed not to support the statement that the mean value was less than 9 gallons.