OCR
RECOGNISING ACHIEVEMENT

ADVANCED SUBSIDIARY GCE

MEI STATISTICS

Statistics 2 (Z2)

G242

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]**