Centre No.						Paper Reference							Surname	Initial(s)
Candidate No.						6	6	9	1	/	0	1	Signature	

Paper Reference(s)

6691/01

Edexcel GCE

Statistics S3

Advanced/Advanced Subsidiary

Monday 16 June 2008 – Afternoon

Time: 1 hour 30 minutes

Examiner's use only		

Team Leader's use only		

Question Number	Leave Blank
1	
2	
3	
4	
5	
6	
7	
Total	

Materials required for examination
Mathematical Formulae (Green)

Items included with question papers
Nil

Candidates may use any calculator allowed by the regulations of the Joint Council for Qualifications. Calculators must not have the facility for symbolic algebra manipulation, differentiation and integration, or have retrievable mathematical formulae stored in them.

Instructions to Candidates

In the boxes above, write your centre number, candidate number, your surname, initials and signature.
Check that you have the correct question paper.
You must write your answer for each question in the space following the question.
Values from the statistical tables should be quoted in full. When a calculator is used, the answer should be given to an appropriate degree of accuracy.

Information for Candidates

A booklet 'Mathematical Formulae and Statistical Tables' is provided.
Full marks may be obtained for answers to ALL questions.
The marks for individual questions and the parts of questions are shown in round brackets: e.g. **(2)**.
There are 7 questions in this question paper. The total mark for this paper is 75.
There are 24 pages in this question paper. Any blank pages are indicated.

Advice to Candidates

You must ensure that your answers to parts of questions are clearly labelled.
You should show sufficient working to make your methods clear to the Examiner.
Answers without working may not gain full credit.

This publication may be reproduced only in accordance with Edexcel Limited copyright policy.
©2008 Edexcel Limited.

Printer's Log. No.
N30984A
W850/R6691/57570 3/3/3

N 3 0 9 8 4 A 0 1 2 4

Turn over

edexcel
advancing learning, changing lives

Leave blank

1. Some biologists were studying a large group of wading birds. A random sample of 36 were measured and the wing length, x mm of each wading bird was recorded. The results are summarised as follows

$$\sum x = 6046 \qquad \sum x^2 = 1\,016\,338$$

(a) Calculate unbiased estimates of the mean and the variance of the wing lengths of these birds.

(3)

Given that the standard deviation of the wing lengths of this particular type of bird is actually 5.1 mm,

(b) find a 99% confidence interval for the mean wing length of the birds from this group.

(5)

Leave blank

Question 1 continued

(Total 8 marks)

Q1

Leave blank

2. Students in a mixed sixth form college are classified as taking courses in either Arts, Science or Humanities. A random sample of students from the college gave the following results

		Course		
		Arts	Science	Humanities
Gender	Boy	30	50	35
	Girl	40	20	42

Showing your working clearly, test, at the 1% level of significance, whether or not there is an association between gender and the type of course taken. State your hypotheses clearly.

(11)

Leave blank

Question 2 continued

Leave blank

Question 2 continued

Question 2 continued

Leave blank

(Total 11 marks)

Q2

Leave blank

3. The product moment correlation coefficient is denoted by r and Spearman's rank correlation coefficient is denoted by r_s.

(a) Sketch separate scatter diagrams, with five points on each diagram, to show

(i) $r = 1$,

(ii) $r_s = -1$ but $r > -1$.

(3)

Two judges rank seven collie dogs in a competition. The collie dogs are labelled A to G and the rankings are as follows

Rank	1	2	3	4	5	6	7
Judge 1	*A*	*C*	*D*	*B*	*E*	*F*	*G*
Judge 2	*A*	*B*	*D*	*C*	*E*	*G*	*F*

(b) (i) Calculate Spearman's rank correlation coefficient for these data.

(6)

(ii) Stating your hypotheses clearly, test, at the 5% level of significance, whether or not the judges are generally in agreement.

(5)

Leave blank

Question 3 continued

Leave blank

Question 3 continued

Leave blank

Question 3 continued

(Total 14 marks)

Q3

Leave blank

4. The weights of adult men are normally distributed with a mean of 84 kg and a standard deviation of 11 kg.

(a) Find the probability that the total weight of 4 randomly chosen adult men is less than 350 kg.

(5)

The weights of adult women are normally distributed with a mean of 62 kg and a standard deviation of 10 kg.

(b) Find the probability that the weight of a randomly chosen adult man is less than one and a half times the weight of a randomly chosen adult woman.

(6)

Question 4 continued

Leave blank

Leave blank

Question 4 continued

Leave blank

Question 4 continued

(Total 11 marks)

Q4

Leave blank

5. A researcher is hired by a cleaning company to survey the opinions of employees on a proposed pension scheme. The company employs 55 managers and 495 cleaners.

To collect data the researcher decides to give a questionnaire to the first 50 cleaners to leave at the end of the day.

(a) Give 2 reasons why this method is likely to produce biased results.

(2)

(b) Explain briefly how the researcher could select a sample of 50 employees using

(i) a systematic sample,

(ii) a stratified sample.

(6)

Using the random number tables in the formulae book, and starting with the top left hand corner (8) and working across, 50 random numbers between 1 and 550 inclusive were selected. The first two suitable numbers are 384 and 100.

(c) Find the next two suitable numbers.

(2)

Leave blank

Question 5 continued

(Total 10 marks)

Q5

Leave blank

6. Ten cuttings were taken from each of 100 randomly selected garden plants. The numbers of cuttings that did not grow were recorded.

The results are as follows

No. of cuttings which did not grow	0	1	2	3	4	5	6	7	8, 9 or 10
Frequency	11	21	30	20	12	3	2	1	0

(a) Show that the probability of a randomly selected cutting, from this sample, not growing is 0.223

(2)

A gardener believes that a binomial distribution might provide a good model for the number of cuttings, out of 10, that do not grow.

He uses a binomial distribution, with the probability 0.2 of a cutting not growing. The calculated expected frequencies are as follows

No. of cuttings which did not grow	0	1	2	3	4	5 or more
Expected frequency	r	26.84	s	20.13	8.81	t

(b) Find the values of r, s and t.

(4)

(c) State clearly the hypotheses required to test whether or not this binomial distribution is a suitable model for these data.

(2)

The test statistic for the test is 4.17 and the number of degrees of freedom used is 4.

(d) Explain fully why there are 4 degrees of freedom.

(2)

(e) Stating clearly the critical value used, carry out the test using a 5% level of significance.

(3)

Question 6 continued

Leave blank

Leave blank

Question 6 continued

Leave blank

Question 6 continued

(Total 13 marks)

Q6

Leave blank

7. A sociologist is studying how much junk food teenagers eat. A random sample of 100 female teenagers and an independent random sample of 200 male teenagers were asked to estimate what their weekly expenditure on junk food was. The results are summarised below.

	n	mean	s.d.
Female teenagers	100	£5.48	£3.62
Male teenagers	200	£6.86	£4.51

(a) Using a 5% significance level, test whether or not there is a difference in the mean amounts spent on junk food by male teenagers and female teenagers. State your hypotheses clearly.

(7)

(b) Explain briefly the importance of the central limit theorem in this problem.

(1)

Question 7 continued

Leave blank

Question 7 continued

Leave blank

Q7

(Total 8 marks)

TOTAL FOR PAPER: 75 MARKS

END