June 2007

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

Wednesday 20 June 2007 – 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 formulas 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.
Answer ALL the questions.
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.
©2007 Edexcel Limited.

Printer's Log. No.
N29602A
W850/R6691/57570 3/3/3/3/3

N 2 9 6 0 2 A 0 1 2 4

Turn over

edexcel
advancing learning, changing lives

Leave blank

1. During a village show, two judges, P and Q, had to award a mark out of 30 to some flower displays. The marks they awarded to a random sample of 8 displays were as follows:

Display	A	B	C	D	E	F	G	H
Judge P	25	19	21	23	28	17	16	20
Judge Q	20	9	21	13	17	14	11	15

(a) Calculate Spearman's rank correlation coefficient for the marks awarded by the two judges.

(6)

After the show, one competitor complained about the judges. She claimed that there was no positive correlation between their marks.

(b) Stating your hypotheses clearly, test whether or not this sample provides support for the competitor's claim. Use a 5% level of significance.

(4)

Leave blank

Question 1 continued

Leave blank

2. The Director of Studies at a large college believed that students' grades in Mathematics were independent of their grades in English. She examined the results of a random group of candidates who had studied both subjects and she recorded the number of candidates in each of the 6 categories shown.

	Maths grade A or B	Maths grade C or D	Maths grade E or U
English grade A or B	25	25	10
English grade C to U	15	30	15

(a) Stating your hypotheses clearly, test the Director's belief using a 10% level of significance. You must show each step of your working.

(9)

The Head of English suggested that the Director was losing accuracy by combining the English grades C to U in one row. He suggested that the Director should split the English grades into two rows, grades C or D and grades E or U as for Mathematics.

(b) State why this might lead to problems in performing the test.

(1)

Leave blank

Question 2 continued

Leave blank

3. The time, in minutes, it takes Robert to complete the puzzle in his morning newspaper each day is normally distributed with mean 18 and standard deviation 3. After taking a holiday, Robert records the times taken to complete a random sample of 15 puzzles and he finds that the mean time is 16.5 minutes. You may assume that the holiday has not changed the standard deviation of times taken to complete the puzzle.

Stating your hypotheses clearly test, at the 5% level of significance, whether or not there has been a reduction in the mean time Robert takes to complete the puzzle.

(7)

Leave blank

4. A quality control manager regularly samples 20 items from a production line and records the number of defective items x. The results of 100 such samples are given in Table 1 below.

x	0	1	2	3	4	5	6	7 or more
Frequency	17	31	19	14	9	7	3	0

Table 1

(a) Estimate the proportion of defective items from the production line.

(2)

The manager claimed that the number of defective items in a sample of 20 can be modelled by a binomial distribution. He used the answer in part (a) to calculate the expected frequencies given in Table 2.

x	0	1	2	3	4	5	6	7 or more
Expected frequency	12.2	27.0	r	19.0	s	3.2	0.9	0.2

Table 2

(b) Find the value of r and the value of s giving your answers to 1 decimal place.

(3)

(c) Stating your hypotheses clearly, use a 5% level of significance to test the manager's claim.

(7)

(d) Explain what the analysis in part (c) tells the manager about the occurrence of defective items from this production line.

(1)

Leave blank

Question 4 continued

Leave blank

5. In a trial of diet A a random sample of 80 participants were asked to record their weight loss, x kg, after their first week of using the diet. The results are summarised by

$$\sum x = 361.6 \quad \text{and} \quad \sum x^2 = 1753.95$$

(a) Find unbiased estimates for the mean and variance of weight lost after the first week of using diet A.

(5)

The designers of diet A believe it can achieve a greater mean weight loss after the first week than a standard diet B. A random sample of 60 people used diet B. After the first week they had achieved a mean weight loss of 4.06 kg, with an unbiased estimate of variance of weight loss of 2.50 kg^2.

(b) Test, at the 5% level of significance, whether or not the mean weight loss after the first week using diet A is greater than that using diet B. State your hypotheses clearly.

(7)

(c) Explain the significance of the central limit theorem to the test in part (b).

(1)

(d) State an assumption you have made in carrying out the test in part (b).

(1)

Leave blank

Question 5 continued

Leave blank

6. A random sample of the daily sales (in £s) of a small company is taken and, using tables of the normal distribution, a 99% confidence interval for the mean daily sales is found to be

$$(123.5, 154.7)$$

Find a 95% confidence interval for the mean daily sales of the company.

(6)

Leave blank

7. A set of scaffolding poles come in two sizes, long and short. The length L of a long pole has the normal distribution $N(19.7, 0.5^2)$. The length S of a short pole has the normal distribution $N(4.9, 0.2^2)$. The random variables L and S are independent.

A long pole and a short pole are selected at random.

(a) Find the probability that the length of the long pole is more than 4 times the length of the short pole.

(7)

Four short poles are selected at random and placed end to end in a row. The random variable T represents the length of the row.

(b) Find the distribution of T.

(3)

(c) Find $P(|L - T| < 0.1)$.

(5)

Leave blank

Question 7 continued