June 2012

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

Thursday 21 June 2012 – 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 (Pink)

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 or symbolic differentiation/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.
Answer ALL the questions.
You must write your answer to 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 Pearson Education Ltd copyright policy.
©2012 Pearson Education Ltd.

Printer's Log. No.
P40472A
W850/R6684/57570 5/5

P 4 0 4 7 2 A 0 1 2 4

Turn over

PEARSON

Leave blank

1. Interviews for a job are carried out by two managers. Candidates are given a score by each manager and the results for a random sample of 8 candidates are shown in the table below.

Candidate	*A*	*B*	*C*	*D*	*E*	*F*	*G*	*H*
Manager *X*	62	56	87	54	65	15	12	10
Manager *Y*	54	47	71	50	49	25	30	44

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

(5)

(b) Test, at the 5% level of significance, whether there is agreement between the rankings awarded by each manager. State your hypotheses clearly.

(5)

Manager *Y* later discovered he had miscopied his score for candidate *D* and it should be 54.

(c) Without carrying out any further calculations, explain how you would calculate Spearman's rank correlation in this case.

(2)

Leave blank

2. A lake contains 3 species of fish. There are estimated to be 1400 trout, 600 bass and 450 pike in the lake. A survey of the health of the fish in the lake is carried out and a sample of 30 fish is chosen.

(a) Give a reason why stratified random sampling cannot be used.
(1)

(b) State an appropriate sampling method for the survey.
(1)

(c) Give one advantage and one disadvantage of this sampling method.
(2)

(d) Explain how this sampling method could be used to select the sample of 30 fish. You must show your working.
(4)

Leave blank

3. (a) Explain what you understand by the Central Limit Theorem.

(2)

A garage services hire cars on behalf of a hire company. The garage knows that the lifetime of the brake pads has a standard deviation of 5000 miles. The garage records the lifetimes, x miles, of the brake pads it has replaced. The garage takes a random sample of 100 brake pads and finds that $\sum x = 1\,740\,000$

(b) Find a 95% confidence interval for the mean lifetime of a brake pad.

(5)

(c) Explain the relevance of the Central Limit Theorem in part (b).

(2)

Brake pads are made to be changed every 20 000 miles on average.
The hire car company complain that the garage is changing the brake pads too soon.

(d) Comment on the hire company's complaint. Give a reason for your answer.

(2)

Leave blank

Question 3 continued

Leave blank

4. Two breeds of chicken are surveyed to measure their egg yield. The results are shown in the table below.

Egg yield / Breed	Low	Medium	High
Leghorn	22	52	26
Cornish	14	32	4

Showing each stage of your working clearly, test, at the 5% significance level, whether or not there is an association between egg yield and breed of chicken. State your hypotheses clearly.

(10)

Leave blank

Question 4 continued

Leave blank

5. Mr Alan and Ms Burns are two Mathematics teachers teaching mixed ability groups of students in a large college. At the end of the college year all students took the same examination. A random sample of 29 of Mr Alan's students and a random sample of 26 of Ms Burns' students are chosen. The results are summarised in the table below.

	Sample Size, n	Mean, $\bar{x}$	Standard Deviation, s
Mr Alan	29	80	10
Ms Burns	26	74	15

(a) Stating your hypotheses clearly, test, at the 10% level of significance whether there is evidence that there is a difference in the mean scores of their students.

(6)

Ms Burns thinks the comparison was unfair as the examination was set by Mr Alan. She looks up a different set of examination results for these students and, although Mr Alan's sample has a higher mean, she calculates the test statistic for this new set of results to be 1.6

However, Mr Alan now claims that the mean marks of his students are higher than the mean marks of Ms Burns' students.

(b) Test Mr Alan's claim, stating the hypotheses and critical values you would use. Use a 10% level of significance.

(3)

Leave blank

Question 5 continued

Leave blank

6. A total of 100 random samples of 6 items are selected from a production line in a factory and the number of defective items in each sample is recorded. The results are summarised in the table below.

Number of defective items	0	1	2	3	4	5	6
Number of samples	6	16	20	23	17	10	8

(a) Show that the mean number of defective items per sample is 2.91

(2)

A factory manager suggests that the data can be modelled by a binomial distribution with $n = 6$. He uses the mean from the sample above and calculates expected frequencies as shown in the table below.

Number of defective items	0	1	2	3	4	5	6
Expected frequency	1.87	10.54	24.82	a	22.01	8.29	b

(b) Calculate the value of a and the value of b giving your answers to 2 decimal places.

(4)

(c) Test, at the 5% level, whether or not the binomial distribution is a suitable model for the number of defective items in samples of 6 items.
State your hypotheses clearly.

(8)

Leave blank

Question 6 continued

Leave blank

7. The heights, in cm, of the male employees in a large company follow a normal distribution with mean 177 and standard deviation 5
The heights, in cm, of the female employees follow a normal distribution with mean 163 and standard deviation 4

A male employee and a female employee are chosen at random.

(a) Find the probability that the male employee is taller than the female employee.
(5)

Six male employees and four female employees are chosen at random.

(b) Find the probability that their total height is less than 17 m.
(6)

Leave blank

Question 7 continued