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 13 June 2013 – Morning

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 20 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.

Printer's Log. No.
P41817A
W850/R6691/57570 5/5/5/5/

P 4 1 8 1 7 A 0 1 2 0

Turn over

PEARSON

Leave blank

1. A doctor takes a random sample of 100 patients and measures their intake of saturated fats in their food and the level of cholesterol in their blood. The results are summarised in the table below.

Cholesterol level / Intake of saturated fats	High	Low
High	12	8
Low	26	54

Using a 5% level of significance, test whether or not there is an association between cholesterol level and intake of saturated fats. State your hypotheses and show your working clearly.

(10)

Leave blank

2. The table below shows the number of students per member of staff and the student satisfaction scores for 7 universities.

University	*A*	*B*	*C*	*D*	*E*	*F*	*G*
Number of students per member of staff	14.2	13.1	13.3	11.7	10.5	15.9	10.8
Student satisfaction score	4.1	4.2	3.8	4.0	3.9	4.3	3.7

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

(5)

(b) Stating your hypotheses clearly test, at the 5% level of significance, whether or not there is evidence of a correlation between the number of students per member of staff and the student satisfaction score.

(3)

Leave blank

3. A college manager wants to survey students' opinions of enrichment activities. She decides to survey the students on the courses summarised in the table below.

Course	Number of students enrolled
Leisure and Sport	420
Information Technology	337
Health and Social Care	200
Media Studies	43

Each student takes only one course.

The manager has access to the college's information system that holds full details of each of the enrolled students including name, address, telephone number and their course of study. She wants to compare the opinions of students on each course and has a generous budget to pay for the cost of the survey.

(a) Give one advantage and one disadvantage of carrying out this survey using

(i) quota sampling,

(ii) stratified sampling.

(2)

The manager decides to take a stratified sample of 100 students.

(b) Calculate the number of students to be sampled from each course.

(3)

(c) Describe how to choose students for the stratified sample.

(2)

Leave blank

4. Customers at a post office are timed to see how long they wait until being served at the counter. A random sample of 50 customers is chosen and their waiting times, x minutes, are summarised in Table 1.

Waiting time in minutes (x)	Frequency
0–3	8
3–5	12
5–6	13
6–8	9
8–12	8

Table 1

(a) Show that an estimate of $\bar{x} = 5.49$ and an estimate of $s_x^2 = 6.88$

(3)

The post office manager believes that the customers' waiting times can be modelled by a normal distribution.
Assuming the data is normally distributed, she calculates the expected frequencies for these data and some of these frequencies are shown in Table 2.

Waiting Time	$x < 3$	3–5	5–6	6–8	$x > 8$
Expected Frequency	8.56	12.73	7.56	a	b

Table 2

(b) Find the value of a and the value of b.

(3)

(c) Test, at the 5% level of significance, the manager's belief. State your hypotheses clearly.

(8)

Leave blank

Question 4 continued

Leave blank

5. Blumen is a perfume sold in bottles. The amount of perfume in each bottle is normally distributed. The amount of perfume in a large bottle has mean 50 ml and standard deviation 5 ml. The amount of perfume in a small bottle has mean 15 ml and standard deviation 3 ml.

One large and 3 small bottles of Blumen are chosen at random.

(a) Find the probability that the amount in the large bottle is less than the total amount in the 3 small bottles.

(6)

A large bottle and a small bottle of Blumen are chosen at random.

(b) Find the probability that the large bottle contains more than 3 times the amount in the small bottle.

(6)

Leave blank

Question 5 continued

Leave blank

6. Fruit-n-Veg4U Market Gardens grow tomatoes. They want to improve their yield of tomatoes by at least 1 kg per plant by buying a new variety. The variance of the yield of the old variety of plant is 0.5 kg^2 and the variance of the yield for the new variety of plant is 0.75 kg^2. A random sample of 60 plants of the old variety has a mean yield of 5.5 kg. A random sample of 70 of the new variety has a mean yield of 7 kg.

(a) Stating your hypotheses clearly test, at the 5% level of significance, whether or not there is evidence that the mean yield of the new variety is more than 1 kg greater than the mean yield of the old variety.

(9)

(b) Explain the relevance of the Central Limit Theorem to the test in part (a).

(2)

Leave blank

Question 6 continued

Q6

(Total 11 marks)

Leave blank

7. Lambs are born in a shed on Mill Farm. The birth weights, x kg, of a random sample of 8 newborn lambs are given below.

4.12 5.12 4.84 4.65 3.55 3.65 3.96 3.40

(a) Calculate unbiased estimates of the mean and variance of the birth weight of lambs born on Mill Farm.

(3)

A further random sample of 32 lambs is chosen and the unbiased estimates of the mean and variance of the birth weight of lambs from this sample are 4.55 and 0.25 respectively.

(b) Treating the combined sample of 40 lambs as a single sample, estimate the standard error of the mean.

(7)

The owner of Mill Farm researches the breed of lamb and discovers that the population of birth weights is normally distributed with standard deviation 0.67 kg.

(c) Calculate a 95% confidence interval for the mean birth weight of this breed of lamb using your combined sample mean.

(3)

Leave blank

Question 7 continued