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 17 June 2009 – 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	
8	
Total	

Materials required for examination
Mathematical Formulae (Orange or 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.
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 8 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.

This publication may be reproduced only in accordance with Edexcel Limited copyright policy.
©2009 Edexcel Limited.

Printer's Log. No.
M34278A
W850/R6691/57570 8/4/4/3

M 3 4 2 7 8 A 0 1 2 0

Turn over

edexcel
advancing learning, changing lives

Leave blank

1. A telephone directory contains 50 000 names. A researcher wishes to select a systematic sample of 100 names from the directory.

(a) Explain in detail how the researcher should obtain such a sample.

(2)

(b) Give one advantage and one disadvantage of

(i) quota sampling,

(ii) systematic sampling.

(4)

Leave blank

Question 1 continued

Q1

(Total 6 marks)

M 3 4 2 7 8 A 0 3 2 0

Turn over

Leave blank

2. The heights of a random sample of 10 imported orchids are measured. The mean height of the sample is found to be 20.1 cm. The heights of the orchids are normally distributed.

Given that the population standard deviation is 0.5 cm,

(a) estimate limits between which 95% of the heights of the orchids lie,

(3)

(b) find a 98% confidence interval for the mean height of the orchids.

(4)

A grower claims that the mean height of this type of orchid is 19.5 cm.

(c) Comment on the grower's claim. Give a reason for your answer.

(2)

Leave blank

Question 2 continued

Q2

(Total 9 marks)

M 3 4 2 7 8 A 0 5 2 0

Turn over

Leave blank

3. A doctor is interested in the relationship between a person's Body Mass Index (BMI) and their level of fitness. She believes that a lower BMI leads to a greater level of fitness. She randomly selects 10 female 18 year-olds and calculates each individual's BMI. The females then run a race and the doctor records their finishing positions. The results are shown in the table.

Individual	*A*	*B*	*C*	*D*	*E*	*F*	*G*	*H*	*I*	*J*
BMI	17.4	21.4	18.9	24.4	19.4	20.1	22.6	18.4	25.8	28.1
Finishing position	3	5	1	9	6	4	10	2	7	8

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

(5)

(b) Stating your hypotheses clearly and using a one tailed test with a 5% level of significance, interpret your rank correlation coefficient.

(5)

(c) Give a reason to support the use of the rank correlation coefficient rather than the product moment correlation coefficient with these data.

(1)

Leave blank

Question 3 continued

Q3

(Total 11 marks)

Leave blank

4. A sample of size 8 is to be taken from a population that is normally distributed with mean 55 and standard deviation 3. Find the probability that the sample mean will be greater than 57.

(5)

Leave blank

Question 4 continued

Q4

(Total 5 marks)

M 3 4 2 7 8 A 0 9 2 0

Turn over

Leave blank

5. The number of goals scored by a football team is recorded for 100 games. The results are summarised in Table 1 below.

Number of goals	Frequency
0	40
1	33
2	14
3	8
4	5

Table 1

(a) Calculate the mean number of goals scored per game.

(2)

The manager claimed that the number of goals scored per match follows a Poisson distribution. He used the answer in part (a) to calculate the expected frequencies given in Table 2.

Number of goals	Expected Frequency
0	34.994
1	r
2	s
3	6.752
$\geqslant 4$	2.221

Table 2

(b) Find the value of r and the value of s giving your answers to 3 decimal places.

(3)

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

(7)

M 3 4 2 7 8 A 0 1 0 2 0

Leave blank

Question 5 continued

Leave blank

Question 5 continued

Leave blank

Question 5 continued

Q5

(Total 12 marks)

Leave blank

6. The lengths of a random sample of 120 limpets taken from the upper shore of a beach had a mean of 4.97 cm and a standard deviation of 0.42 cm. The lengths of a second random sample of 150 limpets taken from the lower shore of the same beach had a mean of 5.05 cm and a standard deviation of 0.67 cm.

(a) Test, using a 5% level of significance, whether or not the mean length of limpets from the upper shore is less than the mean length of limpets from the lower shore. State your hypotheses clearly.

(8)

(b) State two assumptions you made in carrying out the test in part (a).

(2)

Leave blank

Question 6 continued

Q6

(Total 10 marks)

Leave blank

7. A company produces climbing ropes. The lengths of the climbing ropes are normally distributed. A random sample of 5 ropes is taken and the length, in metres, of each rope is measured. The results are given below.

120.3 120.1 120.4 120.2 119.9

(a) Calculate unbiased estimates for the mean and the variance of the lengths of the climbing ropes produced by the company.

(5)

The lengths of climbing rope are known to have a standard deviation of 0.2 m. The company wants to make sure that there is a probability of at least 0.90 that the estimate of the population mean, based on a random sample size of n, lies within 0.05 m of its true value.

(b) Find the minimum sample size required.

(6)

Leave blank

Question 7 continued

Q7

(Total 11 marks)

Leave blank

8. The random variable A is defined as

$$A = 4X - 3Y$$

where $X \sim \mathrm{N}(30, 3^2)$, $Y \sim \mathrm{N}(20, 2^2)$ and X and Y are independent.

Find

(a) $\mathrm{E}(A)$,

(2)

(b) $\mathrm{Var}(A)$.

(3)

The random variables Y_1, Y_2, Y_3 and Y_4 are independent and each has the same distribution as Y. The random variable B is defined as

$$B = \sum_{i=1}^{4} Y_i$$

(c) Find $\mathrm{P}(B > A)$.

(6)

M 3 4 2 7 8 A 0 1 8 2 0

Leave blank

Question 8 continued

Leave blank

Question 8 continued

Q8

(Total 11 marks)

TOTAL FOR PAPER: 75 MARKS

END

M 3 4 2 7 8 A 0 2 0 2 0