Centre No.						Paper Reference							Surname	Initial(s)
Candidate No.						6	6	8	6	/	0	1	Signature	

Paper Reference(s)

6686/01

Edexcel GCE

Statistics S4

Advanced/Advanced Subsidiary

Wednesday 18 June 2008 – 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	**Items included with question papers**
Mathematical Formulae (Green)	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 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.

Printer's Log. No.
H31327A
W850/R6686/57570 3/3/3
H 3 1 3 2 7 A 0 1 2 4

Turn over

edexcel
advancing learning, changing lives

Leave blank

1. A random sample $X_1, X_2, ..., X_{10}$ is taken from a population with mean μ and variance σ^2.

(a) Determine the bias, if any, of each of the following estimators of μ.

$$\theta_1 = \frac{X_3 + X_4 + X_5}{3},$$

$$\theta_2 = \frac{X_{10} - X_1}{3},$$

$$\theta_3 = \frac{3X_1 + 2X_2 + X_{10}}{6}.$$

(4)

(b) Find the variance of each of these estimators.

(5)

(c) State, giving reasons, which of these three estimators for μ is

(i) the best estimator,

(ii) the worst estimator.

(4)

Leave blank

Question 1 continued

H31327A0324

Turn over

Leave blank

2. A large number of students are split into two groups *A* and *B*. The students sit the same test but under different conditions. Group A has music playing in the room during the test, and group B has no music playing during the test. Small samples are then taken from each group and their marks recorded. The marks are normally distributed.

The marks are as follows:

Sample from Group *A*	42	40	35	37	34	43	42	44	49
Sample from Group *B*	40	44	38	47	38	37	33		

(a) Stating your hypotheses clearly, and using a 10% level of significance, test whether or not there is evidence of a difference between the variances of the marks of the two groups.
(8)

(b) State clearly an assumption you have made to enable you to carry out the test in part (a).
(1)

(c) Use a two tailed test, with a 5% level of significance, to determine if the playing of music during the test has made any difference in the mean marks of the two groups. State your hypotheses clearly.
(7)

(d) Write down what you can conclude about the effect of music on a student's performance during the test.
(1)

Leave blank

Question 2 continued

Leave blank

3. The weights, in grams, of mice are normally distributed. A biologist takes a random sample of 10 mice. She weighs each mouse and records its weight.

The ten mice are then fed on a special diet. They are weighed again after two weeks.

Their weights in grams are as follows:

Mouse	*A*	*B*	*C*	*D*	*E*	*F*	*G*	*H*	*I*	*J*
Weight before diet	50.0	48.3	47.5	54.0	38.9	42.7	50.1	46.8	40.3	41.2
Weight after diet	52.1	47.6	50.1	52.3	42.2	44.3	51.8	48.0	41.9	43.6

Stating your hypotheses clearly, and using a 1% level of significance, test whether or not the diet causes an increase in the mean weight of the mice.

(8)

Leave blank

4. A town council is concerned that the mean price of renting two bedroom flats in the town has exceeded £650 per month. A random sample of eight two bedroom flats gave the following results, £x, per month.

705, 640, 560, 680, 800, 620, 580, 760

[You may assume $\sum x = 5345$ $\sum x^2 = 3621025$]

(a) Find a 90% confidence interval for the mean price of renting a two bedroom flat. **(6)**

(b) State an assumption that is required for the validity of your interval in part (a). **(1)**

(c) Comment on whether or not the town council is justified in being concerned. Give a reason for your answer. **(2)**

Leave blank

5. A machine is filling bottles of milk. A random sample of 16 bottles was taken and the volume of milk in each bottle was measured and recorded. The volume of milk in a bottle is normally distributed and the unbiased estimate of the variance, s^2, of the volume of milk in a bottle is 0.003

(a) Find a 95% confidence interval for the variance of the population of volumes of milk from which the sample was taken.

(5)

The machine should fill bottles so that the standard deviation of the volumes is equal to 0.07

(b) Comment on this with reference to your 95% confidence interval.

(3)

Leave blank

Question 5 continued

(Total 8 marks)

Q5

H31327A01524

Turn over

Leave blank

6. A drug is claimed to produce a cure to a certain disease in 35% of people who have the disease. To test this claim a sample of 20 people having this disease is chosen at random and given the drug. If the number of people cured is between 4 and 10 inclusive the claim will be accepted. Otherwise the claim will not be accepted.

(a) Write down suitable hypotheses to carry out this test.
(2)

(b) Find the probability of making a Type I error.
(3)

The table below gives the value of the probability of the Type II error, to 4 decimal places, for different values of p where p is the probability of the drug curing a person with the disease.

P(cure)	0.2	0.3	0.4	0.5
P(Type II error)	0.5880	r	0.8565	s

(c) Calculate the value of r and the value of s.
(3)

(d) Calculate the power of the test for $p = 0.2$ and $p = 0.4$
(2)

(e) Comment, giving your reasons, on the suitability of this test procedure.
(2)

Leave blank

Question 6 continued

Leave blank

7. An engineering firm buys steel rods. The steel rods from its present supplier are known to have a mean tensile strength of 230 N/mm^2.

A new supplier of steel rods offers to supply rods at a cheaper price than the present supplier. A random sample of ten rods from this new supplier gave tensile strengths, x N/mm^2, which are summarised below.

Sample size	Σx	Σx^2
10	2283	524 079

(a) Stating your hypotheses clearly, and using a 5% level of significance, test whether or not the rods from the new supplier have a tensile strength lower than the present supplier. (You may assume that the tensile strength is normally distributed).
(7)

(b) In the light of your conclusion to part (a) write down what you would recommend the engineering firm to do.
(1)

Leave blank

Question 7 continued