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

Friday 22 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	
Total	

Materials required for examination	Items included with question papers
Mathematical Formulae (Pink)	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 6 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.
P40473A
W850/R6684/57570 5/4/5/3

P 4 0 4 7 3 A 0 1 2 4

Turn over

PEARSON

Leave blank

1. A medical student is investigating whether there is a difference in a person's blood pressure when sitting down and after standing up. She takes a random sample of 12 people and measures their blood pressure, in mmHg, when sitting down and after standing up.

The results are shown below.

Person	*A*	*B*	*C*	*D*	*E*	*F*	*G*	*H*	*I*	*J*	*K*	*L*
Sitting down	135	146	138	146	141	158	136	135	146	161	119	151
Standing up	131	147	132	140	138	160	127	136	142	154	130	144

The student decides to carry out a paired *t*-test to investigate whether, on average, the blood pressure of a person when sitting down is more than their blood pressure after standing up.

(a) State clearly the hypotheses that should be used and any necessary assumption that needs to be made.

(2)

(b) Carry out the test at the 1% level of significance.

(7)

Leave blank

Question 1 continued

Q1

(Total 9 marks)

P 4 0 4 7 3 A 0 3 2 4

Turn over

Leave blank

2. A biologist investigating the shell size of turtles takes random samples of adult female and adult male turtles and records the length, x cm, of the shell. The results are summarised below.

	Number in sample	Sample mean $\bar{x}$	$\sum x^2$
Female	6	19.6	2308.01
Male	12	13.7	2262.57

You may assume that the samples come from independent normal distributions with the same variance.

The biologist claims that the mean shell length of adult female turtles is 5 cm longer than the mean shell length of adult male turtles.

(a) Test the biologist's claim at the 5% level of significance.

(10)

(b) Given that the true values for the variance of the population of adult male turtles and adult female turtles are both 0.9 cm^2,

(i) show that when samples of size 6 and 12 are used with a 5% level of significance, the biologist's claim will be accepted if $4.07 < \bar{X}_F - \bar{X}_M < 5.93$ where $\bar{X}_F$ and $\bar{X}_M$ are the mean shell lengths of females and males respectively.

(ii) Hence find the probability of a type II error for this test if in fact the true mean shell length of adult female turtles is 6 cm more than the mean shell length of adult male turtles.

(6)

Leave blank

Question 2 continued

P 4 0 4 7 3 A 0 5 2 4

Turn over

Leave blank

3. The sample variance of the lengths of a random sample of 9 paving slabs sold by a builders' merchant is 36 mm^2. The sample variance of the lengths of a random sample of 11 paving slabs sold by a second builders' merchant is 225 mm^2. Test at the 10 % significance level whether or not there is evidence that the lengths of paving slabs sold by these builders' merchants differ in variability. State your hypotheses clearly.

(You may assume the lengths of paving slabs are normally distributed.)

(5)

Leave blank

4. A newspaper runs a daily Sudoku. A random sample of 10 people took the following times, in minutes, to complete the Sudoku.

5.0 4.5 4.7 5.3 5.2 4.1 5.3 4.8 5.5 4.6

Given that the times to complete the Sudoku follow a normal distribution,

(a) calculate a 95% confidence interval for

(i) the mean,

(ii) the variance,

of the times taken by people to complete the Sudoku.

(13)

The newspaper requires the average time needed to complete the Sudoku to be 5 minutes with a standard deviation of 0.7 minutes.

(b) Comment on whether or not the Sudoku meets this requirement. Give a reason for your answer.

(3)

Leave blank

Question 4 continued

Leave blank

5. Boxes of chocolates manufactured by Philippe have a mean weight of μ grams and a standard deviation of σ grams. A random sample of 25 of these boxes are weighed. Using this sample, the unbiased estimate of μ is 455 and the unbiased estimate of σ^2 is 55.

(a) Test, at the 5% level of significance, whether or not σ is greater than 6. State your hypotheses clearly.
(6)

(b) Test, at the 5% level of significance, whether or not μ is more than 450.
(6)

(c) State an assumption you have made in order to carry out the above tests.
(1)

Leave blank

Question 5 continued

Leave blank

6. When a tree seed is planted the probability of it germinating is p.
A random sample of size n is taken and the number of tree seeds, X, which germinate is recorded.

(a) (i) Show that $\hat{p}_1 = \dfrac{X}{n}$ is an unbiased estimator of p.

(ii) Find the variance of $\hat{p}_1$.

(4)

A second sample of size m is taken and the number of tree seeds, Y, which germinate is recorded.

Given that $\hat{p}_2 = \dfrac{Y}{m}$ and that $\hat{p}_3 = a(3\hat{p}_1 + 2\hat{p}_2)$ is an unbiased estimator of p,

(b) show that

(i) $a = \dfrac{1}{5}$,

(ii) $\mathrm{Var}(\hat{p}_3) = \dfrac{p(1-p)}{25}\left(\dfrac{9}{n} + \dfrac{4}{m}\right)$.

(6)

(c) Find the range of values of $\dfrac{n}{m}$ for which

$$\mathrm{Var}(\hat{p}_3) < \mathrm{Var}(\hat{p}_1) \text{ and } \mathrm{Var}(\hat{p}_3) < \mathrm{Var}(\hat{p}_2)$$

(3)

(d) Given that $n = 20$ and $m = 60$, explain which of $\hat{p}_1$, $\hat{p}_2$ or $\hat{p}_3$ is the best estimator.

(3)

Leave blank

Question 6 continued