Centre No.						Paper Reference							Surname	Initial(s)
Candidate No.						6	6	8	3	/	0	1	Signature	

Paper Reference(s)

6683/01

Edexcel GCE

Statistics S1

Advanced/Advanced Subsidiary

Friday 14 January 2005 – 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 (Lilac or Green)

Items included with question papers
Nil

Candidates may use any calculator EXCEPT those with the facility for symbolic algebra, differentiation and/or integration. Thus candidates may NOT use calculators such as the Texas Instruments TI 89, TI 92, Casio CFX 9970G, Hewlett Packard HP 48G.

Instructions to Candidates

In the boxes above, write your centre number, candidate number, your surname, initial(s) and signature.
You must write your answer to each question in the space following the question.
If you need more space to complete your answer to any question use additional answer sheets.
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.
This paper has seven questions. Pages 7 and 20 are blank.
The total mark for this paper is 75.

Advice to Candidates

You must ensure that your answers to parts of questions are clearly labelled.
You must show sufficient working to make your methods clear to the examiner. Answers without working may gain no credit.

This publication may only be reproduced in accordance with London Qualifications Limited copyright policy.
©2005 London Qualifications Limited.

Printer's Log. No.
N16741A
W850/R6683/57570 4/3/7/5/4/2/2

N 1 6 7 4 1 A

Turn over

edexcel

Leave blank

1. A company assembles drills using components from two sources. Goodbuy supplies 85% of the components and Amart supplies the rest. It is known that 3% of the components supplied by Goodbuy are faulty and 6% of those supplied by Amart are faulty.

(a) Represent this information on a tree diagram.

(3)

An assembled drill is selected at random.

(b) Find the probability that it is not faulty.

(3)

Leave blank

Question 1 continued

Q1

(Total 6 marks)

Leave blank

2. The number of caravans on Seaview caravan site on each night in August last year is summarised in the following stem and leaf diagram.

Caravans	1 | 0 means 10	Totals
1	0 5	(2)
2	1 2 4 8	(4)
3	0 3 3 3 4 7 8 8	(8)
4	1 1 3 5 8 8 8 9 9	(9)
5	2 3 6 6 7	(5)
6	2 3 4	(3)

(a) Find the three quartiles of these data.

(3)

During the same month, the least number of caravans on Northcliffe caravan site was 31. The maximum number of caravans on this site on any night that month was 72. The three quartiles for this site were 38, 45 and 52 respectively.

(b) On the graph paper on page 5 and using the same scale, draw box plots to represent the data for both caravan sites. You may assume that there are no outliers.

(6)

(c) Compare and contrast these two box plots.

(3)

(d) Give an interpretation to the upper quartiles of these two distributions.

(2)

Leave blank

Leave blank

Question 2 continued

Q2

(Total 14 marks)

BLANK PAGE

Leave blank

3. The following table shows the height x, to the nearest cm, and the weight y, to the nearest kg, of a random sample of 12 students.

x	148	164	156	172	147	184	162	155	182	165	175	152
y	39	59	56	77	44	77	65	49	80	72	70	52

(a) On the graph paper on page 9, draw a scatter diagram to represent these data.

(3)

(b) Write down, with a reason, whether the correlation coefficient between x and y is positive or negative.

(2)

The data in the table can be summarised as follows.

$$\Sigma x = 1962, \quad \Sigma y = 740, \quad \Sigma y^2 = 47\,746, \quad \Sigma xy = 122\,783, \quad S_{xx} = 1745.$$

(c) Find S_{xy}.

(2)

The equation of the regression line of y on x is $y = -106.331 + bx$.

(d) Find, to 3 decimal places, the value of b.

(2)

(e) Find, to 3 significant figures, the mean $\bar{y}$ and the standard deviation s of the weights of this sample of students.

(3)

(f) Find the values of $\bar{y} \pm 1.96s$.

(2)

(g) Comment on whether or not you think that the weights of these students could be modelled by a normal distribution.

(1)

Leave blank

Leave blank

Question 3 continued

Leave blank

Question 3 continued

Q3

(Total 15 marks)

Leave blank

4. The random variable X has probability function

$$P(X = x) = kx, \qquad x = 1, 2, ..., 5.$$

(a) Show that $k = \frac{1}{15}$.

(2)

Find

(b) $P(X < 4)$,

(2)

(c) $E(X)$,

(2)

(d) $E(3X - 4)$.

(2)

Leave blank

Question 4 continued

Q4

(Total 8 marks)

Leave blank

5. Articles made on a lathe are subject to three kinds of defect, *A*, *B* or *C*. A sample of 1000 articles was inspected and the following results were obtained.

31 had a type *A* defect
37 had a type *B* defect
42 had a type *C* defect
11 had both type *A* and type *B* defects
13 had both type *B* and type *C* defects
10 had both type *A* and type *C* defects
6 had all three types of defect.

(a) In the space below, draw a Venn diagram to represent these data.

(6)

Find the probability that a randomly selected article from this sample had

(b) no defects,

(1)

(c) no more than one of these defects.

(2)

An article selected at random from this sample had only one defect.

(d) Find the probability that it was a type *B* defect.

(2)

Two different articles were selected at random from this sample.

(e) Find the probability that both had type *B* defects.

(2)

Leave blank

Question 5 continued

Q5

(Total 13 marks)

Leave blank

6. A discrete random variable is such that each of its values is assumed to be equally likely.

(a) Write down the name of the distribution that could be used to model this random variable.

(1)

(b) Give an example of such a distribution.

(1)

(c) Comment on the assumption that each value is equally likely.

(2)

(d) Suggest how you might refine the model in part (a).

(2)

Leave blank

Question 6 continued

Q6

(Total 6 marks)

Leave blank

7. The random variable X is normally distributed with mean 79 and variance 144.

Find

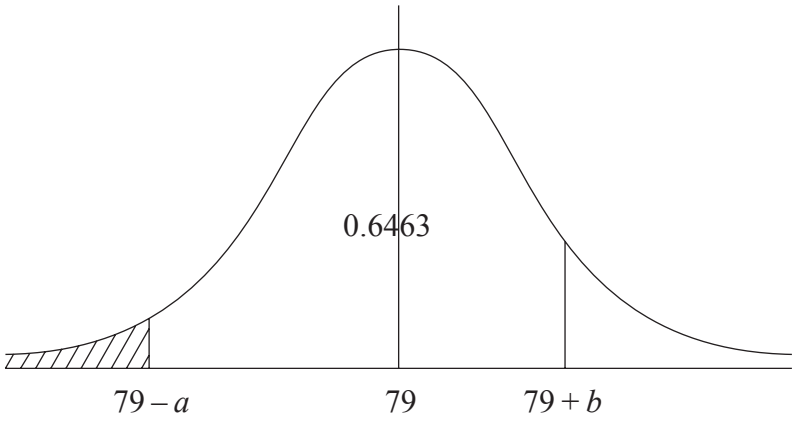
(a) $P(X < 70)$,

(3)

(b) $P(64 < X < 96)$.

(3)

It is known that $P(79 - a \leqslant X \leqslant 79 + b) = 0.6463$. This information is shown in the figure below.



Given that $P(X \geqslant 79 + b) = 2P(X \leqslant 79 - a)$,

(c) show that the area of the shaded region is 0.1179.

(3)

(d) Find the value of b.

(4)

Leave blank

Question 7 continued

Q7

(Total 13 marks)

TOTAL FOR PAPER: 75 MARKS

END

BLANK PAGE