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

Monday 19 January 2009 – 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 (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 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 Edexcel Limited copyright policy.
©2009 Edexcel Limited.

Printer's Log. No.
N32680A
W850/R6683/57570 3/3/3

N 3 2 6 8 0 A 0 1 2 4

Turn over

edexcel
advancing learning, changing lives

Leave blank

1. A teacher is monitoring the progress of students using a computer based revision course. The improvement in performance, y marks, is recorded for each student along with the time, x hours, that the student spent using the revision course. The results for a random sample of 10 students are recorded below.

x hours	1.0	3.5	4.0	1.5	1.3	0.5	1.8	2.5	2.3	3.0
y marks	5	30	27	10	–3	–5	7	15	–10	20

[You may use $\sum x = 21.4$, $\sum y = 96$, $\sum x^2 = 57.22$, $\sum xy = 313.7$]

(a) Calculate S_{xx} and S_{xy}.

(3)

(b) Find the equation of the least squares regression line of y on x in the form $y = a + bx$.

(4)

(c) Give an interpretation of the gradient of your regression line.

(1)

Rosemary spends 3.3 hours using the revision course.

(d) Predict her improvement in marks.

(2)

Lee spends 8 hours using the revision course claiming that this should give him an improvement in performance of over 60 marks.

(e) Comment on Lee's claim.

(1)

Leave blank

Question 1 continued

Leave blank

Question 1 continued

Leave blank

Question 1 continued

(Total 11 marks)

Q1

N 3 2 6 8 0 A 0 5 2 4

Turn over

Leave blank

2. A group of office workers were questioned for a health magazine and $\frac{2}{5}$ were found to take regular exercise. When questioned about their eating habits $\frac{2}{3}$ said they always eat breakfast and, of those who always eat breakfast $\frac{9}{25}$ also took regular exercise.

Find the probability that a randomly selected member of the group

(a) always eats breakfast and takes regular exercise,

(2)

(b) does not always eat breakfast and does not take regular exercise.

(4)

(c) Determine, giving your reason, whether or not always eating breakfast and taking regular exercise are statistically independent.

(2)

Leave blank

Question 2 continued

Leave blank

Question 2 continued

Leave blank

Question 2 continued

Q2

(Total 8 marks)

Leave blank

3. When Rohit plays a game, the number of points he receives is given by the discrete random variable X with the following probability distribution.

x	0	1	2	3
$P(X = x)$	0.4	0.3	0.2	0.1

(a) Find $E(X)$.

(2)

(b) Find $F(1.5)$.

(2)

(c) Show that $Var(X) = 1$

(4)

(d) Find $Var(5 - 3X)$.

(2)

Rohit can win a prize if the total number of points he has scored after 5 games is at least 10. After 3 games he has a total of 6 points.
You may assume that games are independent.

(e) Find the probability that Rohit wins the prize.

(6)

Leave blank

Question 3 continued

Leave blank

Question 3 continued

Leave blank

Question 3 continued

(Total 16 marks)

Q3

N 3 2 6 8 0 A 0 1 3 2 4

Leave blank

4. In a study of how students use their mobile telephones, the phone usage of a random sample of 11 students was examined for a particular week.

The total length of calls, y minutes, for the 11 students were

17, 23, 35, 36, 51, 53, 54, 55, 60, 77, 110

(a) Find the median and quartiles for these data.

(3)

A value that is greater than $Q_3 + 1.5 \times (Q_3 - Q_1)$ or smaller than $Q_1 - 1.5 \times (Q_3 - Q_1)$ is defined as an outlier.

(b) Show that 110 is the only outlier.

(2)

(c) Using the graph paper on page 15 draw a box plot for these data indicating clearly the position of the outlier.

(3)

The value of 110 is omitted.

(d) Show that S_{yy} for the remaining 10 students is 2966.9

(3)

These 10 students were each asked how many text messages, x, they sent in the same week.

The values of S_{xx} and S_{xy} for these 10 students are $S_{xx} = 3463.6$ and $S_{xy} = -18.3$.

(e) Calculate the product moment correlation coefficient between the number of text messages sent and the total length of calls for these 10 students.

(2)

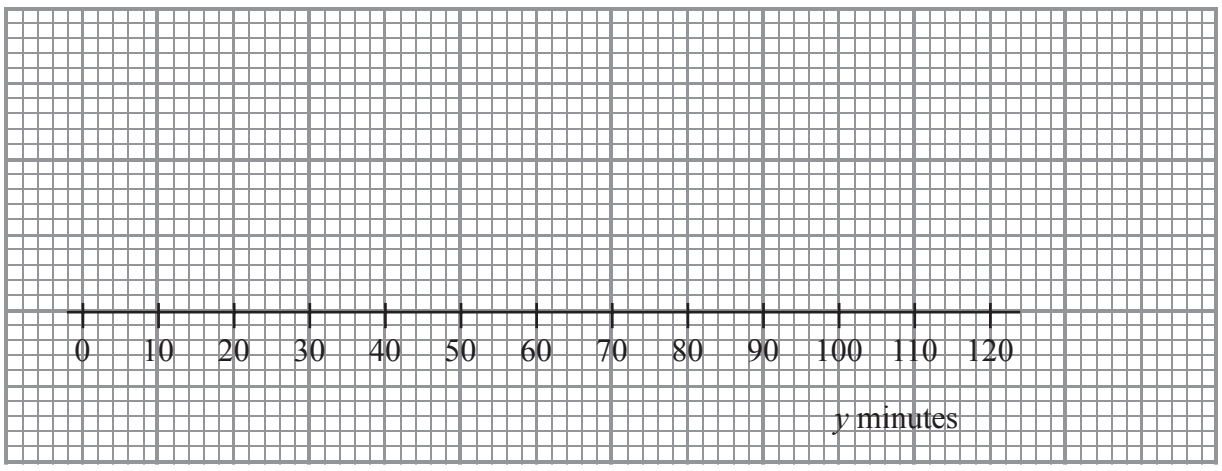
A parent believes that a student who sends a large number of text messages will spend fewer minutes on calls.

(f) Comment on this belief in the light of your calculation in part (e).

(1)

Leave blank

Question 4 continued



N 3 2 6 8 0 A 0 1 5 2 4

Turn over

Leave blank

Question 4 continued

Leave blank

Question 4 continued

(Total 14 marks)

Q4

N 3 2 6 8 0 A 0 1 7 2 4

Turn over

Leave blank

5. In a shopping survey a random sample of 104 teenagers were asked how many hours, to the nearest hour, they spent shopping in the last month. The results are summarised in the table below.

Number of hours	Mid-point	Frequency
0 – 5	2.75	20
6 – 7	6.5	16
8 – 10	9	18
11 – 15	13	25
16 – 25	20.5	15
26 – 50	38	10

A histogram was drawn and the group (8 – 10) hours was represented by a rectangle that was 1.5 cm wide and 3 cm high.

(a) Calculate the width and height of the rectangle representing the group (16 – 25) hours.
(3)

(b) Use linear interpolation to estimate the median and interquartile range.
(5)

(c) Estimate the mean and standard deviation of the number of hours spent shopping.
(4)

(d) State, giving a reason, the skewness of these data.
(2)

(e) State, giving a reason, which average and measure of dispersion you would recommend to use to summarise these data.
(2)

N 3 2 6 8 0 A 0 1 8 2 4

Leave blank

Question 5 continued

Leave blank

Question 5 continued

N 3 2 6 8 0 A 0 2 0 2 4

Leave blank

Question 5 continued

(Total 16 marks)

Q5

Leave blank

6. The random variable X has a normal distribution with mean 30 and standard deviation 5.

(a) Find $P(X < 39)$.

(2)

(b) Find the value of d such that $P(X < d) = 0.1151$

(4)

(c) Find the value of e such that $P(X > e) = 0.1151$

(2)

(d) Find $P(d < X < e)$.

(2)

N 3 2 6 8 0 A 0 2 2 2 4

Leave blank

Question 6 continued

Leave blank

Question 6 continued

Q6

(Total 10 marks)

TOTAL FOR PAPER: 75 MARKS

END

N 3 2 6 8 0 A 0 2 4 2 4