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 5 June 2015 – 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	
Total	

Materials required for examination
Mathematical Formulae (Pink)

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 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.

Printer's Log. No.
P44845A
W850/R6683/57570 5/1/1/1/1/

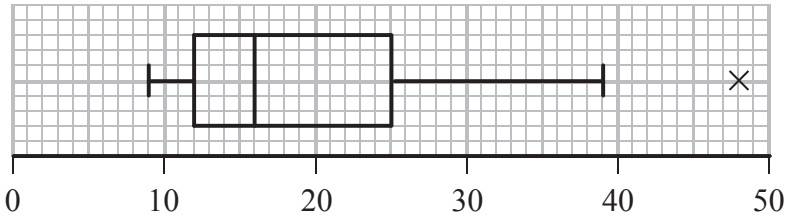
P 4 4 8 4 5 A 0 1 2 4

Turn over

PEARSON

Leave blank

1. Each of 60 students was asked to draw a 20° angle without using a protractor. The size of each angle drawn was measured. The results are summarised in the box plot below.



(a) Find the range for these data.

(1)

(b) Find the interquartile range for these data.

(1)

The students were then asked to draw a 70° angle.
The results are summarised in the table below.

Angle, a, (degrees)	Number of students
$55 \leqslant a < 60$	6
$60 \leqslant a < 65$	15
$65 \leqslant a < 70$	13
$70 \leqslant a < 75$	11
$75 \leqslant a < 80$	8
$80 \leqslant a < 85$	7

(c) Use linear interpolation to estimate the size of the median angle drawn. Give your answer to 1 decimal place.

(2)

(d) Show that the lower quartile is 63°

(2)

For these data, the upper quartile is 75°, the minimum is 55° and the maximum is 84°

An outlier is an observation that falls either
more than 1.5 × (interquartile range) above the upper quartile or
more than 1.5 × (interquartile range) below the lower quartile.

(e) (i) Show that there are no outliers for these data.

(ii) Draw a box plot for these data on the grid on page 3.

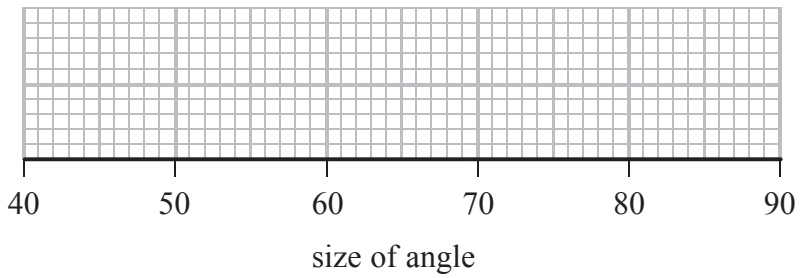
(5)

(f) State which angle the students were more accurate at drawing. Give reasons for your answer.

(3)

Leave blank

Question 1 continued



Leave blank

2. An estate agent recorded the price per square metre, p £/m^2, for 7 two-bedroom houses.

He then coded the data using the coding $q = \dfrac{p - a}{b}$, where a and b are positive constants.

His results are shown in the table below.

p	1840	1848	1830	1824	1819	1834	1850
q	4.0	4.8	3.0	2.4	1.9	3.4	5.0

(a) Find the value of a and the value of b

(2)

The estate agent also recorded the distance, d km, of each house from the nearest train station. The results are summarised below.

$$S_{dd} = 1.02 \quad S_{qq} = 8.22 \quad S_{dq} = -2.17$$

(b) Calculate the product moment correlation coefficient between d and q

(2)

(c) Write down the value of the product moment correlation coefficient between d and p

(1)

The estate agent records the price and size of 2 additional two-bedroom houses, H and J.

House	Price (£)	Size (m^2)
H	156 400	85
J	172 900	95

(d) Suggest which house is most likely to be closer to a train station. Justify your answer.

(3)

Leave blank

Question 2 continued

Leave blank

3. A college has 80 students in Year 12.

20 students study Biology
28 students study Chemistry
30 students study Physics
7 students study both Biology and Chemistry
11 students study both Chemistry and Physics
5 students study both Physics and Biology
3 students study all 3 of these subjects

(a) Draw a Venn diagram to represent this information.

(5)

A Year 12 student at the college is selected at random.

(b) Find the probability that the student studies Chemistry but not Biology or Physics.

(1)

(c) Find the probability that the student studies Chemistry or Physics or both.

(2)

Given that the student studies Chemistry or Physics or both,

(d) find the probability that the student does not study Biology.

(2)

(e) Determine whether studying Biology and studying Chemistry are statistically independent.

(3)

Leave blank

Question 3 continued

Leave blank

4. Statistical models can provide a cheap and quick way to describe a real world situation.

(a) Give two other reasons why statistical models are used.

(2)

A scientist wants to develop a model to describe the relationship between the average daily temperature, x °C, and her household's daily energy consumption, y kWh, in winter.

A random sample of the average daily temperature and her household's daily energy consumption are taken from 10 winter days and shown in the table.

x	–0.4	–0.2	0.3	0.8	1.1	1.4	1.8	2.1	2.5	2.6
y	28	30	26	25	26	27	26	24	22	21

[You may use $\sum x^2 = 24.76$ $\sum y = 255$ $\sum xy = 283.8$ $S_{xx} = 10.36$]

(b) Find S_{xy} for these data.

(3)

(c) Find the equation of the regression line of y on x in the form $y = a + bx$

Give the value of a and the value of b to 3 significant figures.

(4)

(d) Give an interpretation of the value of a

(1)

(e) Estimate her household's daily energy consumption when the average daily temperature is 2 °C

(2)

The scientist wants to use the linear regression model to predict her household's energy consumption in the summer.

(f) Discuss the reliability of using this model to predict her household's energy consumption in the summer.

(2)

Leave blank

Question 4 continued

Leave blank

5. In a quiz, a team gains 10 points for every question it answers correctly and loses 5 points for every question it does not answer correctly. The probability of answering a question correctly is 0.6 for each question. One round of the quiz consists of 3 questions.

The discrete random variable X represents the total number of points scored in one round. The table shows the incomplete probability distribution of X

x	30	15	0	–15
$P(X = x)$	0.216			0.064

(a) Show that the probability of scoring 15 points in a round is 0.432

(2)

(b) Find the probability of scoring 0 points in a round.

(1)

(c) Find the probability of scoring a total of 30 points in 2 rounds.

(3)

(d) Find $E(X)$

(2)

(e) Find $Var(X)$

(3)

In a bonus round of 3 questions, a team gains 20 points for every question it answers correctly and loses 5 points for every question it does not answer correctly.

(f) Find the expected number of points scored in the bonus round.

(3)

Leave blank

Question 5 continued

Leave blank

6. The random variable $Z \sim N(0, 1)$

A is the event $Z > 1.1$
B is the event $Z > -1.9$
C is the event $-1.5 < Z < 1.5$

(a) Find

(i) P(A)

(ii) P(B)

(iii) P(C)

(iv) P($A \cup C$)

(6)

The random variable X has a normal distribution with mean 21 and standard deviation 5

(b) Find the value of w such that $P(X > w \mid X > 28) = 0.625$

(6)

Leave blank

Question 6 continued