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

Wednesday 13 January 2010 – 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	
7	
Total	

Materials required for examination
Mathematical Formulae (Pink or Green)

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, 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 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 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.

This publication may be reproduced only in accordance with Edexcel Limited copyright policy.
©2010 Edexcel Limited.

Printer's Log. No.
N35711A
W850/R6683/57570 5/5/5/3

N 3 5 7 1 1 A 0 1 2 4

Turn over

edexcel
advancing learning, changing lives

Leave blank

1. A jar contains 2 red, 1 blue and 1 green bead. Two beads are drawn at random from the jar without replacement.

(a) In the space below, draw a tree diagram to illustrate all the possible outcomes and associated probabilities. State your probabilities clearly.

(3)

(b) Find the probability that a blue bead and a green bead are drawn from the jar.

(2)

Leave blank

Question 1 continued

(Total 5 marks)

Q1

Leave blank

2. The 19 employees of a company take an aptitude test. The scores out of 40 are illustrated in the stem and leaf diagram below.

	2|6 means a score of 26	
0	7	(1)
1	88	(2)
2	4468	(4)
3	2333459	(7)
4	00000	(5)

Find

(a) the median score,

(1)

(b) the interquartile range.

(3)

The company director decides that any employees whose scores are so low that they are outliers will undergo retraining.

An outlier is an observation whose value is less than the lower quartile minus 1.0 times the interquartile range.

(c) Explain why there is only one employee who will undergo retraining.

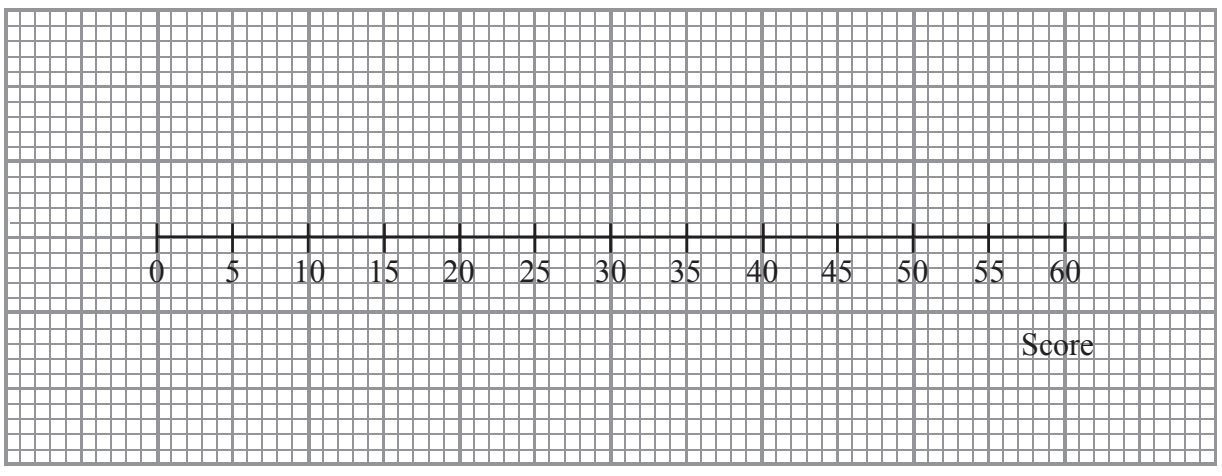
(2)

(d) On the graph paper on page 5, draw a box plot to illustrate the employees' scores.

(3)

Question 2 continued

Leave blank



Q2

(Total 9 marks)

Leave blank

3. The birth weights, in kg, of 1500 babies are summarised in the table below.

Weight (kg)	Midpoint, $x$kg	Frequency, f
0.0 – 1.0	0.50	1
1.0 – 2.0	1.50	6
2.0 – 2.5	2.25	60
2.5 – 3.0		280
3.0 – 3.5	3.25	820
3.5 – 4.0	3.75	320
4.0 – 5.0	4.50	10
5.0 – 6.0		3

[You may use $\sum fx = 4841$ and $\sum fx^2 = 15\,889.5$]

(a) Write down the missing midpoints in the table above.

(2)

(b) Calculate an estimate of the mean birth weight.

(2)

(c) Calculate an estimate of the standard deviation of the birth weight.

(3)

(d) Use interpolation to estimate the median birth weight.

(2)

(e) Describe the skewness of the distribution. Give a reason for your answer.

(2)

Leave blank

Question 3 continued

Leave blank

Question 3 continued

Question 3 continued

Leave blank

(Total 11 marks)

Q3

Leave blank

4. There are 180 students at a college following a general course in computing. Students on this course can choose to take up to three extra options.

112 take systems support,
70 take developing software,
81 take networking,
35 take developing software and systems support,
28 take networking and developing software,
40 take systems support and networking,
4 take all three extra options.

(a) In the space below, draw a Venn diagram to represent this information.

(5)

A student from the course is chosen at random.

Find the probability that this student takes

(b) none of the three extra options,

(1)

(c) networking only.

(1)

Students who want to become technicians take systems support and networking. Given that a randomly chosen student wants to become a technician,

(d) find the probability that this student takes all three extra options.

(2)

Leave blank

Question 4 continued

Leave blank

Question 4 continued

Leave blank

Question 4 continued

Q4

(Total 9 marks)

Leave blank

5. The probability function of a discrete random variable X is given by

$$p(x) = kx^2 \qquad x = 1, 2, 3$$

where k is a positive constant.

(a) Show that $k = \frac{1}{14}$

(2)

Find

(b) $P(X \geqslant 2)$

(2)

(c) $E(X)$

(2)

(d) $\text{Var}(1 - X)$

(4)

N 3 5 7 1 1 A 0 1 4 2 4

Leave blank

Question 5 continued

(Total 10 marks)

Q5

Leave blank

6. The blood pressures, p mmHg, and the ages, t years, of 7 hospital patients are shown in the table below.

Patient	A	B	C	D	E	F	G
t	42	74	48	35	56	26	60
p	98	130	120	88	182	80	135

$$[\sum t = 341, \sum p = 833, \sum t^2 = 18\,181, \sum p^2 = 106\,397, \sum tp = 42\,948]$$

(a) Find S_{pp}, S_{tp} and S_{tt} for these data.

(4)

(b) Calculate the product moment correlation coefficient for these data.

(3)

(c) Interpret the correlation coefficient.

(1)

(d) On the graph paper on page 17, draw the scatter diagram of blood pressure against age for these 7 patients.

(2)

(e) Find the equation of the regression line of p on t.

(4)

(f) Plot your regression line on your scatter diagram.

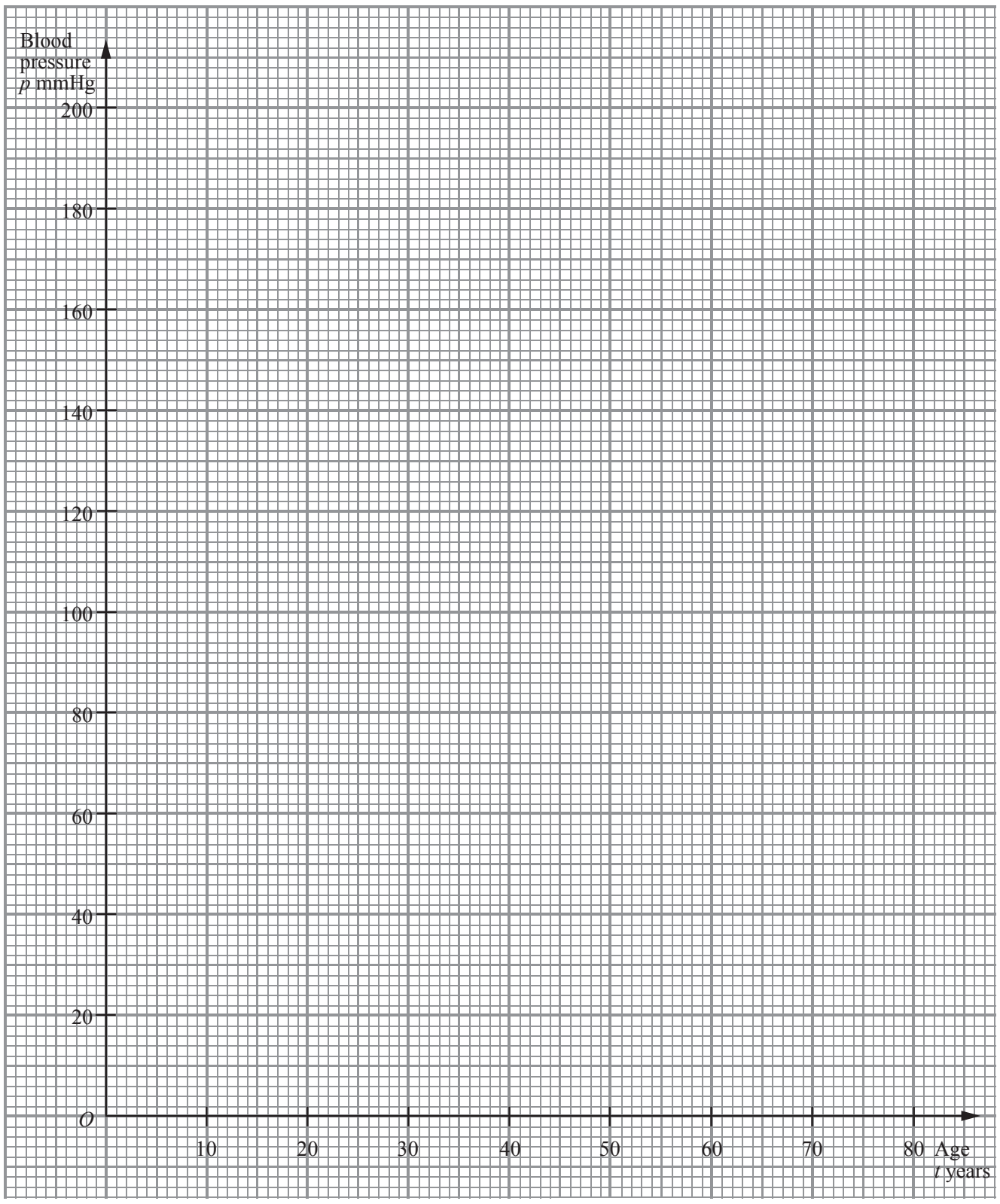
(2)

(g) Use your regression line to estimate the blood pressure of a 40 year old patient.

(2)

N 3 5 7 1 1 A 0 1 6 2 4

Question 6 continued



N 3 5 7 1 1 A 0 1 7 2 4

Leave blank

Question 6 continued

Question 6 continued

Leave blank

(Total 18 marks)

Q6

Leave blank

7. The heights of a population of women are normally distributed with mean μ cm and standard deviation σ cm. It is known that 30% of the women are taller than 172 cm and 5% are shorter than 154 cm.

(a) Sketch a diagram to show the distribution of heights represented by this information.

(3)

(b) Show that $\mu = 154 + 1.6449\sigma$.

(3)

(c) Obtain a second equation and hence find the value of μ and the value of σ.

(4)

A woman is chosen at random from the population.

(d) Find the probability that she is taller than 160 cm.

(3)

Question 7 continued

Leave blank

Leave blank

Question 7 continued

Q7

(Total 13 marks)

TOTAL FOR PAPER: 75 MARKS

END

BLANK PAGE

N35711A02324

BLANK PAGE

N35711A02424