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

Tuesday 15 January 2008 – 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 (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 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 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.
©2008 Edexcel Limited.

Printer's Log. No.
M29283A
W850/R6683/57570 3/3/3/3/3

M 2 9 2 8 3 A 0 1 2 4

Turn over

edexcel
advancing learning, changing lives

https://xtremepape.rs/

Leave blank

1. A personnel manager wants to find out if a test carried out during an employee's interview and a skills assessment at the end of basic training is a guide to performance after working for the company for one year.

The table below shows the results of the interview test of 10 employees and their performance after one year.

Employee	*A*	*B*	*C*	*D*	*E*	*F*	*G*	*H*	*I*	*J*
Interview test, x %.	65	71	79	77	85	78	85	90	81	62
Performance after one year, y %.	65	74	82	64	87	78	61	65	79	69

[You may use $\sum x^2 = 60\ 475, \sum y^2 = 53\ 122, \sum xy = 56\ 076$]

(a) Showing your working clearly, calculate the product moment correlation coefficient between the interview test and the performance after one year.

(5)

The product moment correlation coefficient between the skills assessment and the performance after one year is –0.156 to 3 significant figures.

(b) Use your answer to part (a) to comment on whether or not the interview test and skills assessment are a guide to the performance after one year. Give clear reasons for your answers.

(2)

https://xtremepape.rs/

Leave blank

Question 1 continued

https://xtremepape.rs/

Leave blank

Question 1 continued

https://xtremepape.rs/

Leave blank

Question 1 continued

(Total 7 marks)

Q1

M 2 9 2 8 3 A 0 5 2 4

Turn over

https://xtremepape.rs/

Leave blank

2. Cotinine is a chemical that is made by the body from nicotine which is found in cigarette smoke. A doctor tested the blood of 12 patients, who claimed to smoke a packet of cigarettes a day, for cotinine. The results, in appropriate units, are shown below.

Patient	*A*	*B*	*C*	*D*	*E*	*F*	*G*	*H*	*I*	*J*	*K*	*L*
Cotinine level, x	160	390	169	175	125	420	171	250	210	258	186	243

[You may use $\sum x^2 = 724\ 961$]

(a) Find the mean and standard deviation of the level of cotinine in a patient's blood.
(4)

(b) Find the median, upper and lower quartiles of these data.
(3)

A doctor suspects that some of his patients have been smoking more than a packet of cigarettes per day. He decides to use $Q_3+1.5(Q_3-Q_1)$ to determine if any of the cotinine results are far enough away from the upper quartile to be outliers.

(c) Identify which patient(s) may have been smoking more than a packet of cigarettes a day. Show your working clearly.
(4)

Research suggests that cotinine levels in the blood form a skewed distribution.

One measure of skewness is found using $\dfrac{(Q_1 - 2Q_2 + Q_3)}{(Q_3 - Q_1)}$.

(d) Evaluate this measure and describe the skewness of these data.
(3)

https://xtremepape.rs/

Leave blank

Question 2 continued

https://xtremepape.rs/

Leave blank

Question 2 continued

https://xtremepape.rs/

Leave blank

Question 2 continued

(Total 14 marks)

Q2

M 2 9 2 8 3 A 0 9 2 4

Turn over

https://xtremepape.rs/

Leave blank

3. The histogram in Figure 1 shows the time taken, to the nearest minute, for 140 runners to complete a fun run.

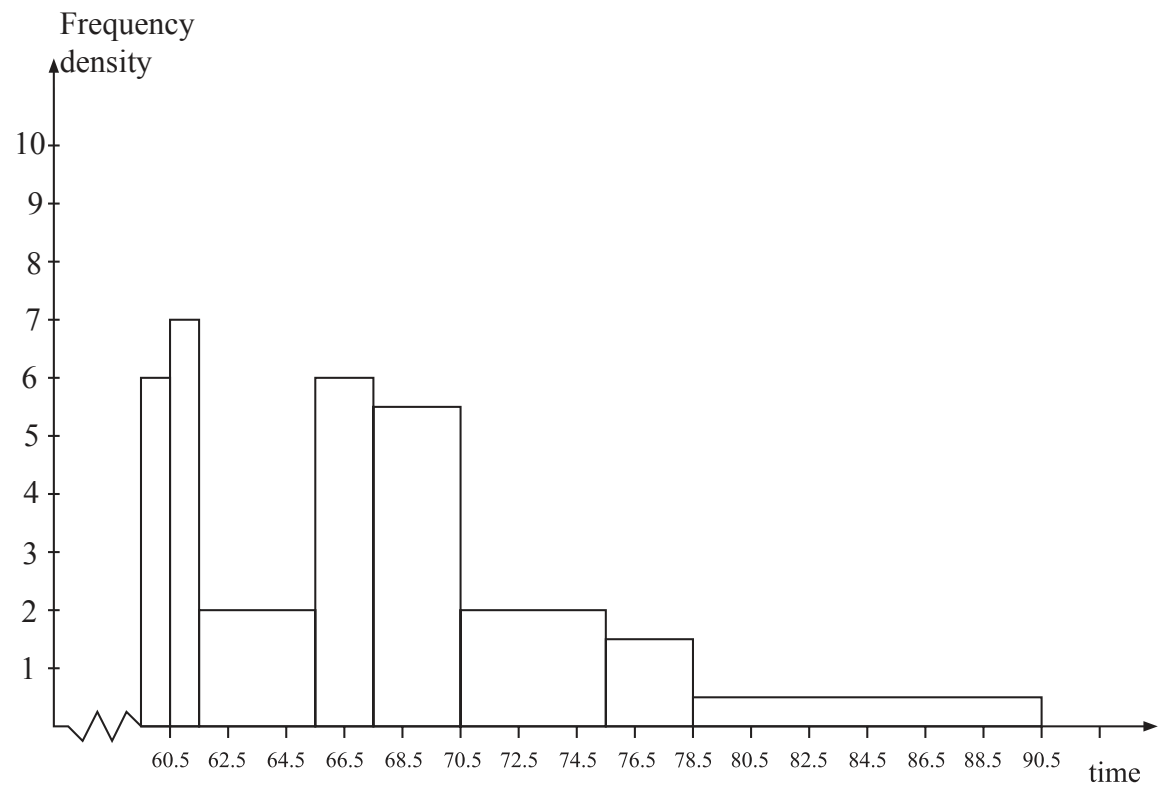

Figure 1

Use the histogram to calculate the number of runners who took between 78.5 and 90.5 minutes to complete the fun run.

(5)

M 2 9 2 8 3 A 0 1 0 2 4

https://xtremepape.rs/

Leave blank

Question 3 continued

Q3

(Total 5 marks)

https://xtremepape.rs/

Leave blank

4. A second hand car dealer has 10 cars for sale. She decides to investigate the link between the age of the cars, x years, and the mileage, y thousand miles. The data collected from the cars are shown in the table below.

Age, x (years)	2	2.5	3	4	4.5	4.5	5	3	6	6.5
Mileage, y (thousands)	22	34	33	37	40	45	49	30	58	58

[You may assume that $\sum x = 41, \sum y = 406, \sum x^2 = 188, \sum xy = 1818.5$]

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

(3)

(b) Find the equation of the least squares regression line in the form $y=a+bx$. Give the values of a and b to 2 decimal places.

(4)

(c) Give a practical interpretation of the slope b.

(1)

(d) Using your answer to part (b), find the mileage predicted by the regression line for a 5 year old car.

(2)

M 2 9 2 8 3 A 0 1 2 2 4

https://xtremepape.rs/

Leave blank

Question 4 continued

https://xtremepape.rs/

Leave blank

Question 4 continued

https://xtremepape.rs/

Leave blank

Question 4 continued

(Total 10 marks)

Q4

M 2 9 2 8 3 A 0 1 5 2 4

Turn over

https://xtremepape.rs/

Leave blank

5. The following shows the results of a wine tasting survey of 100 people.

96 like wine A,
93 like wine B,
96 like wine C,
92 like A and B,
91 like B and C,
93 like A and C,
90 like all three wines.

(a) Draw a Venn Diagram to represent these data.
(6)

Find the probability that a randomly selected person from the survey likes

(b) none of the three wines,
(1)

(c) wine A but not wine B,
(2)

(d) any wine in the survey except wine C,
(2)

(e) exactly two of the three kinds of wine.
(2)

Given that a person from the survey likes wine A,

(f) find the probability that the person likes wine C.
(3)

https://xtremepape.rs/

Leave blank

Question 5 continued

https://xtremepape.rs/

Leave blank

Question 5 continued

https://xtremepape.rs/

Leave blank

Question 5 continued

(Total 16 marks)

Q5

M 2 9 2 8 3 A 0 1 9 2 4

Turn over

https://xtremepape.rs/

Leave blank

6. The weights of bags of popcorn are normally distributed with mean of 200 g and 60% of all bags weighing between 190 g and 210 g.

(a) Write down the median weight of the bags of popcorn.

(1)

(b) Find the standard deviation of the weights of the bags of popcorn.

(5)

A shopkeeper finds that customers will complain if their bag of popcorn weighs less than 180 g.

(c) Find the probability that a customer will complain.

(3)

https://xtremepape.rs/

Leave blank

Question 6 continued

(Total 9 marks)

Q6

M 2 9 2 8 3 A 0 2 1 2 4

Turn over

https://xtremepape.rs/

Leave blank

7. Tetrahedral dice have four faces. Two fair tetrahedral dice, one red and one blue, have faces numbered 0, 1, 2, and 3 respectively. The dice are rolled and the numbers face down on the two dice are recorded. The random variable R is the score on the red die and the random variable B is the score on the blue die.

(a) Find $P(R=3 \text{ and } B=0)$.

(2)

The random variable T is R multiplied by B.

(b) Complete the diagram below to represent the sample space that shows all the possible values of T.

3				
2		2		
1	0			
0				
B / R	**0**	**1**	**2**	**3**

Sample space diagram of *T*

(3)

(c) The table below represents the probability distribution of the random variable T.

t	0	1	2	3	4	6	9
$P(T=t)$	a	b	1/8	1/8	c	1/8	d

Find the values of a, b, c and d.

(3)

Find the values of

(d) $E(T)$,

(2)

(e) $Var(T)$.

(4)

https://xtremepape.rs/

Leave blank

Question 7 continued

https://xtremepape.rs/

Leave blank

Question 7 continued

Q7

(Total 14 marks)

TOTAL FOR PAPER: 75 MARKS

END

M 2 9 2 8 3 A 0 2 4 2 4

https://xtremepape.rs/