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

6683/01

Edexcel GCE

Statistics S1

Advanced/Advanced Subsidiary

Tuesday 16 January 2007 – Morning

Time: 1 hour 30 minutes

Materials	required	for	examination

Mathematical Formulae (Green)

Items included with question papers

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.

Question

Examiner's use only

Team Leader's use only

2	
3	
4	
5	
6	
7	

Instructions to Candidates

ND003924834

In the boxes above, write your centre number, candidate number, your surname, initial(s) and signature.

Check that you have the correct question paper.

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 parts of questions are shown in round brackets: e.g. (2).

There are 7 questions in this question paper. The total for this question paper is 75.

There are 20 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 must show sufficient working to make your methods clear to the examiner. Answers without working may gain no credit.

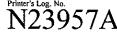
This publication may be reproduced only in accordance with Edexcel Limited copyright policy.



Turn over

Total





W850/R6683/57570 3/3/3/3/3/3/17,400

1. As part of a statistics project, Gill collected data relating to the length of time, to the nearest minute, spent by shoppers in a supermarket and the amount of money they spent. Her data for a random sample of 10 shoppers are summarised in the table below, where t represents time and £m the amount spent over £20.

t (minutes)	£m
15	-3
23	17
5	-19
16	4
30	12
- 6	-9
32	27
23	6
35	20
27	6

(a) Write down the actual amount spent by the shopper who was in the supermarket for 15 minutes.

(1)

(b) Calculate S_{tt} , S_{mm} and S_{tm} .

(You may use
$$\Sigma t^2 = 5478$$
 $\Sigma m^2 = 2101$ $\Sigma tm = 2485$)

(c) Calculate the value of the product moment correlation coefficient between t and m.

(3)

(d) Write down the value of the product moment correlation coefficient between t and the actual amount spent. Give a reason to justify your value.

(2)

On another day Gill collected similar data. For these data the product moment correlation coefficient was 0.178

(e) Give an interpretation to both of these coefficients.

(2)

(f) Suggest a practical reason why these two values are so different.

(1)



Question 1 continued	
	The second of th
<u> </u>	



	t opposition and a set
	14 of the second same on

· · · · · · · · · · · · · · · · · · ·	



5

- 2. In a factory, machines A, B and C are all producing metal rods of the same length. Machine A produces 35% of the rods, machine B produces 25% and the rest are produced by machine C. Of their production of rods, machines A, B and C produce 3%, 6% and 5% defective rods respectively.
 - (a) Draw a tree diagram to represent this information.

(3)

- (b) Find the probability that a randomly selected rod is
 - (i) produced by machine A and is defective,
 - (ii) is defective.

(5)

(c) Given that a randomly selected rod is defective, find the probability that it was produced by machine C.

(3)



(Total 11 marks)

3. The random variable X has probability function

$$P(X=x) = \frac{(2x-1)}{36}$$
 $x = 1, 2, 3, 4, 5, 6.$

(a) Construct a table giving the probability distribution of X.

(3)

Find

(b) $P(2 \le X \le 5)$,

(2)

(c) the exact value of E(X).

(2)

(d) Show that Var(X) = 1.97 to 3 significant figures.

(4)

(e) Find Var(2-3X).

(2)

	· mate sur!
	-
<u></u>	



(Total 13 marks)

4. Summarised below are the distances, to the nearest mile, travelled to work by a random sample of 120 commuters.

Distance (to the nearest mile)	Number of commuters
0–9	10
10-19	19
20–29	43
30–39	25
40–49	8
50-59	6
60–69	5
70–79	3
80-89	1

For this distribution,

(a) describe its shape,

(1)

(b) use linear interpolation to estimate its median.

(2)

The mid-point of each class was represented by x and its corresponding frequency by f giving

$$\Sigma fx = 3550$$
 and $\Sigma fx^2 = 138020$

(c) Estimate the mean and the standard deviation of this distribution.

(3)

One coefficient of skewness is given by

$$\frac{3(\text{mean} - \text{median})}{\text{standard deviation}}$$

(d) Evaluate this coefficient for this distribution.

(3)

(e) State whether or not the value of your coefficient is consistent with your description in part (a). Justify your answer.

(2)





Question 4 continued



Question 4 continued
en e
(Total 14 marks)



13

Leave	
blank	

5.	A teacher recorded, to the nearest hour, the time spent watching television during a particular week by each child in a random sample. The times were summarised in a grouped frequency table and represented by a histogram.				
	One of the classes in the grouped frequency distribution was 20–29 and its associated frequency was 9. On the histogram the height of the rectangle representing that class was 3.6 cm and the width was 2 cm.				
	(a) Give a reason to support the use of a histogram to represent these data.	(1)			
	(b) Write down the underlying feature associated with each of the bars in a histogram	n. (1)			
	(c) Show that on this histogram each child was represented by 0.8 cm ² .	(3)			
	The total area under the histogram was 24 cm ² .				
	(d) Find the total number of children in the group.	(2)			
	<u></u>				
700 a. a. mar haba		Accele celling			



	Leav blan
Question 5 continued	
	l Í
	ı
en de la composition	
	25

(Total 7 marks)

6.	(a) Give two reasons to justify the use of statistical models. (2)
	It has been suggested that there are 7 stages involved in creating a statistical model. They are summarised below, with stages 3, 4 and 7 missing.
	Stage 1. The recognition of a real-world problem.
	Stage 2. A statistical model is devised.
	Stage 3.
	Stage 4.
	Stage 5. Comparisons are made against the devised model.
	Stage 6. Statistical concepts are used to test how well the model describes the real-world problem.
	Stage 7.
	(b) Write down the missing stages. (3)
and an about	



	T
	Leave blank
Question 6 continued	Diank
Question o continueu	
	-
A STATE OF THE PARTY OF THE PAR	
and an other to be a final and the second se	
	<u>'</u>
	l
	l
	1
	1
	l
A second of the	
and the second of the second o	
The state of the s	
The second secon	
	ĺ
]
The state of the s	
lo lo	6
ranger in the second of the	<u>"</u>



(Total 5 marks)

7.	The measure of intelligence, IQ, of a group of students is assumed to be Norma distributed with mean 100 and standard deviation 15.		
	(a) Find the probability that a student selected at random has an IQ less than 91.		
	The probability that a randomly selected student has an IQ of at least $100 + k$ is 0.2090 .		
	(b) Find, to the nearest integer, the value of k.		
	(6) I had, to the hearest integer, the value of h.		
,			



	* A - Care and common - C - 1 - 2 - 2 - 2 - 2 - 2 - 2 - 2 - 2 - 2
en de la composition de la composition La composition de la	
	-
	Pod more annual to the commence of the com-
and the second of the second	
en de la companya de	
	-
	the second second
(Total 1	0 marks)
TOTAL FOR PAPER: 75	



BLANK PAGE