

Paper Reference(s)

**6683**

**Edexcel GCE**

**Statistics S1**

**Advanced/Advanced Subsidiary**

**Thursday 5 June 2003 – Morning**

**Time: 1 hour 30 minutes**

**Materials required for examination**

Answer Book (AB16)

Graph Paper (ASG2)

Mathematical Formulae (Lilac)

**Items included with question papers**

Nil

**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.**

**Instructions to Candidates**

---

In the boxes on the answer book, write the name of the examining body (Edexcel), your centre number, candidate number, the unit title (Statistics S1), the paper reference (6683), your surname, other name and signature.

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.

This paper has seven questions.

**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.

1. In a particular week, a dentist treats 100 patients. The length of time, to the nearest minute, for each patient's treatment is summarised in the table below.

Time (minutes)	4 – 7	8	9 – 10	11	12 – 16	17 – 20
Number of patients	12	20	18	22	15	13

Draw a histogram to illustrate these data.

**(5)**

---

2. The lifetimes of batteries used for a computer game have a mean of 12 hours and a standard deviation of 3 hours. Battery lifetimes may be assumed to be normally distributed.

Find the lifetime,  $t$  hours, of a battery such that 1 battery in 5 will have a lifetime longer than  $t$ .

**(6)**

---

3. A company owns two petrol stations  $P$  and  $Q$  along a main road. Total daily sales in the same week for  $P$  (£ $p$ ) and for  $Q$  (£ $q$ ) are summarised in the table below.

	$p$	$q$
Monday	4760	5380
Tuesday	5395	4460
Wednesday	5840	4640
Thursday	4650	5450
Friday	5365	4340
Saturday	4990	5550
Sunday	4365	5840

When these data are coded using  $x = \frac{p - 4365}{100}$  and  $y = \frac{q - 4340}{100}$ ,

$$\Sigma x = 48.1, \Sigma y = 52.8, \Sigma x^2 = 486.44, \Sigma y^2 = 613.22 \text{ and } \Sigma xy = 204.95.$$

- (a) Calculate  $S_{xy}$ ,  $S_{xx}$  and  $S_{yy}$ . (4)
- (b) Calculate, to 3 significant figures, the value of the product moment correlation coefficient between  $x$  and  $y$ . (3)
- (c) (i) Write down the value of the product moment correlation coefficient between  $p$  and  $q$ .  
(ii) Give an interpretation of this value. (3)
-

4. The discrete random variable  $X$  has probability function

$$P(X=x) = \begin{cases} k(x^2 - 9), & x = 4, 5, 6 \\ 0, & \text{otherwise,} \end{cases}$$

where  $k$  is a positive constant.

- (a) Show that  $k = \frac{1}{50}$ . **(3)**
- (b) Find  $E(X)$  and  $\text{Var}(X)$ . **(6)**
- (c) Find  $\text{Var}(2X - 3)$ . **(2)**
- 

5. The random variable  $X$  represents the number on the uppermost face when a fair die is thrown.

- (a) Write down the name of the probability distribution of  $X$ . **(1)**
- (b) Calculate the mean and the variance of  $X$ . **(3)**

Three fair dice are thrown and the numbers on the uppermost faces are recorded.

- (c) Find the probability that all three numbers are 6. **(2)**
- (d) Write down all the different ways of scoring a total of 16 when the three numbers are added together. **(4)**
- (e) Find the probability of scoring a total of 16. **(2)**
-

6. The number of bags of potato crisps sold per day in a bar was recorded over a two-week period. The results are shown below.

20, 15, 10, 30, 33, 40, 5, 11, 13, 20, 25, 42, 31, 17

(a) Calculate the mean of these data. **(2)**

(b) Draw a stem and leaf diagram to represent these data. **(3)**

(c) Find the median and the quartiles of these data. **(3)**

An outlier is an observation that falls either  $1.5 \times$  (interquartile range) above the upper quartile or  $1.5 \times$  (interquartile range) below the lower quartile.

(d) Determine whether or not any items of data are outliers. **(3)**

(e) On graph paper draw a box plot to represent these data. Show your scale clearly. **(3)**

(f) Comment on the skewness of the distribution of bags of crisps sold per day. Justify your answer. **(2)**

---

**TURN OVER FOR QUESTION 7**

7. Eight students took tests in mathematics and physics. The marks for each student are given in the table below where  $m$  represents the mathematics mark and  $p$  the physics mark.

		Student							
		<i>A</i>	<i>B</i>	<i>C</i>	<i>D</i>	<i>E</i>	<i>F</i>	<i>G</i>	<i>H</i>
Mark	<i>m</i>	9	14	13	10	7	8	20	17
	<i>p</i>	11	23	21	15	19	10	31	26

A science teacher believes that students' marks in physics depend upon their mathematical ability. The teacher decides to investigate this relationship using the test marks.

- (a) Write down which is the explanatory variable in this investigation. (1)
- (b) Draw a scatter diagram to illustrate these data. (3)
- (c) Showing your working, find the equation of the regression line of  $p$  on  $m$ . (8)
- (d) Draw the regression line on your scatter diagram. (2)

A ninth student was absent for the physics test, but she sat the mathematics test and scored 15.

- (e) Using this model, estimate the mark she would have scored in the physics test. (2)

---

**END**