

Paper Reference(s)

6683

Edexcel GCE

Statistics S1

Advanced/Advanced Subsidiary

Tuesday 2 November 2004 – 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. As part of their job, taxi drivers record the number of miles they travel each day. A random sample of the mileages recorded by taxi drivers Keith and Asif are summarised in the back-to-back stem and leaf diagram below.

Totals	Keith												Asif							Totals	
(9)			8	7	7	4	3	2	1	1	0	18	4	4	5	7	(4)				
(11)	9	9	8	7	6	5	4	3	3	1	1	19	5	7	8	9	9	(5)			
(6)						8	7	4	2	2	0	20	0	2	2	4	4	8	(6)		
(6)						9	4	3	1	0	0	21	2	3	5	6	6	7	9	(7)	
(4)							6	4	1	1		22	1	1	2	4	5	5	8	(7)	
(2)										2	0	23	1	1	3	4	6	6	7	8	(8)
(2)										7	1	24	2	4	8	9	(4)				
(1)											9	25	4	(1)							
(2)										9	3	26		(0)							

Key: 0 | 18 | 4 means 180 for Keith and 184 for Asif

The quartiles for these two distributions are summarised in the table below.

	Keith	Asif
Lower quartile	191	<i>a</i>
Median	<i>b</i>	218
Upper quartile	221	<i>c</i>

- (a) Find the values of *a*, *b* and *c*. **(3)**

Outliers are values that lie outside the limits

$$Q_1 - 1.5(Q_3 - Q_1) \quad \text{and} \quad Q_3 + 1.5(Q_3 - Q_1).$$

- (b) On graph paper, and showing your scale clearly, draw a box plot to represent Keith's data. **(8)**

- (c) Comment on the skewness of the two distributions. **(3)**
-

2. An experiment carried out by a student yielded pairs of (x, y) observations such that

$$\bar{x} = 36, \quad \bar{y} = 28.6, \quad S_{xx} = 4402, \quad S_{xy} = 3477.6$$

- (a) Calculate the equation of the regression line of y on x in the form $y = a + bx$. Give your values of a and b to 2 decimal places. **(3)**

- (b) Find the value of y when $x = 45$. **(1)**
-

3. The random variable $X \sim N(\mu, \sigma^2)$.

It is known that

$$P(X \leq 66) = 0.0359 \quad \text{and} \quad P(X \geq 81) = 0.1151.$$

- (a) In the space below, give a clearly labelled sketch to represent these probabilities on a Normal curve. **(1)**

- (b) (i) Show that the value of σ is 5.

- (ii) Find the value of μ . **(8)**

- (c) Find $P(69 \leq X \leq 83)$. **(3)**
-

4. The discrete random variable X has probability function

$$P(X=x) = \begin{cases} 0.2, & x = -3, -2, \\ \alpha, & x = -1, 0 \\ 0.1, & x = 1, 2. \end{cases}$$

Find

- (a) α , (2)
- (b) $P(-1 \leq X < 2)$, (1)
- (c) $F(0.6)$, (1)
- (d) the value of a such that $E(aX + 3) = 1.2$, (4)
- (e) $\text{Var}(X)$, (4)
- (f) $\text{Var}(3X - 2)$. (2)
-

5. The events A and B are such that $P(A) = \frac{1}{2}$, $P(B) = \frac{1}{3}$ and $P(A \cap B) = \frac{1}{4}$.

- (a) Using the space below, represent these probabilities in a Venn diagram. (4)

Hence, or otherwise, find

- (b) $P(A \cup B)$, (1)
- (c) $P(A \mid B')$ (2)
-

6. Students in Mr Brawn's exercise class have to do press-ups and sit-ups. The number of press-ups x and the number of sit-ups y done by a random sample of 8 students are summarised below.

$$\begin{aligned}\Sigma x &= 272, & \Sigma x^2 &= 10\,164, & \Sigma xy &= 11\,222, \\ \Sigma y &= 320, & \Sigma y^2 &= 13\,464.\end{aligned}$$

- (a) Evaluate S_{xx} , S_{yy} and S_{xy} . (4)
- (b) Calculate, to 3 decimal places, the product moment correlation coefficient between x and y . (3)
- (c) Give an interpretation of your coefficient. (2)
- (d) Calculate the mean and the standard deviation of the number of press-ups done by these students. (4)

Mr Brawn assumes that the number of press-ups that can be done by any student can be modelled by a normal distribution with mean μ and standard deviation σ . Assuming that μ and σ take the same values as those calculated in part (d),

- (e) find the value of a such that $P(\mu - a < X < \mu + a) = 0.95$. (3)
- (f) Comment on Mr Brawn's assumption of normality. (2)
-

7. A college organised a 'fun run'. The times, to the nearest minute, of a random sample of 100 students who took part are summarised in the table below.

Time	Number of students
40–44	10
45–47	15
48	23
49–51	21
52–55	16
56–60	15

- (a) Give a reason to support the use of a histogram to represent these data. **(1)**
- (b) Write down the upper class boundary and the lower class boundary of the class 40–44. **(1)**
- (c) On graph paper, draw a histogram to represent these data. **(4)**

END