6683

Edexcel GCE

Statistics S1
(New Syllabus)
Advanced/Advanced Subsidiary
Tuesday 12 June 2001 − Afternoon
Time: 1 hour 30 minutes

Materials required for examination   **Items included with question papers**
Answer Book (AB16)   Nil
Graph Paper (ASG2)
Mathematical Formulae (Lilac)

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. Pages 6, 7 and 8 are blank.

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. Each of the 25 students on a computer course recorded the number of minutes x, to the nearest minute, spent surfing the internet during a given day. The results are summarised below.

$$\Sigma x = 1075, \ \Sigma x^2 = 44\,625.$$

(a) Find μ and σ for these data.

(3)

Two other students surfed the internet on the same day for 35 and 51 minutes respectively.

(b) Without further calculation, explain the effect on the mean of including these two students.

(2)

2. On a particular day in summer 1993 at 0800 hours the height above sea level, x metres, and the temperature, y °C, were recorded in 10 Mediterranean towns. The following summary statistics were calculated from the results.

$$\Sigma x = 7300, \ \Sigma x^2 = 6\,599\,600, \ S_{xy} = -13\,060, \ S_{yy} = 140.9.$$

(a) Find S_{xx}.

(2)

(b) Calculate, to 3 significant figures, the product moment correlation coefficient between x and y.

(2)

(c) Give an interpretation of your coefficient.

(1)

3. The continuous random variable Y is normally distributed with mean 100 and variance 256.

(a) Find $P(Y < 80)$.

(3)

(b) Find k such that $P(100 - k \le Y \le 100 + k) = 0.5l6$.

(5)

4. The discrete random variable X has the probability function shown in the table below.

x	-2	-1	0	1	2	3
$P(X = x)$	0.1	α	0.3	0.2	0.1	0.1

Find

(a) α,

(1)

(b) $P(-1 < X \le 2)$,

(2)

(c) $F(-0.4)$,

(1)

(d) $E(3X + 4)$,

(4)

(e) $Var(2X + 3)$.

(4)

---

5. A market researcher asked 100 adults which of the three newspapers A, B, C they read. The results showed that 30 read A, 26 read B, 21 read C, 5 read both A and B, 7 read both B and C, 6 read both C and A and 2 read all three.

(a) Draw a Venn diagram to represent these data.

(6)

One of the adults is then selected at random.

Find the probability that she reads

(b) at least one of the newspapers,

(2)

(c) only A,

(1)

(d) only one of the newspapers,

(2)

(e) A given that she reads only one newspaper.

(2)

---

6. Three swimmers Alan, Diane and Gopal record the number of lengths of the swimming pool they swim during each practice session over several weeks. The stem and leaf diagram below shows the results for Alan.

Lengths		2 $\vert$ 0 means 20
2	0 1 2 2	(4)
2	5 5 6 7 7 8 9	(7)
3	0 1 2 2 4	(5)
3	5 6 6 7 9	(5)
4	0 1 3 3 3 3 3 4 4 4	(10)
4	5 5 6 6 6 7 7 8 8 9 9 9	(12)
5	0 0 0	(3)

(*a*) Find the three quartiles for Alan's results.

(4)

The table below summarises the results for Diane and Gopal.

	Diane	Gopal
Smallest value	35	25
Lower quartile	37	34
Median	42	42
Upper quartile	53	50
Largest value	65	57

(*b*) Using the same scale and on the same sheet of graph paper draw box plots to represent the data for Alan, Diane and Gopal.

(8)

(*c*) Compare and contrast the three box plots.

(4)

---

7. A music teacher monitored the sight-reading ability of one of her pupils over a 10 week period. At the end of each week, the pupil was given a new piece to sight-read and the teacher noted the number of errors *y*. She also recorded the

number of hours x that the pupil had practised each week. The data are shown in the table below.

x	12	15	7	11	1	8	4	6	9	3
y	8	4	13	8	18	12	15	14	12	16

(*a*) Plot these data on a scatter diagram.

(3)

(*b*) Find the equation of the regression line of y on x in the form $y = a + bx$.

(You may use $\Sigma x^2 = 746$, $\Sigma xy = 749$.)

(9)

(*c*) Give an interpretation of the slope and the intercept of your regression line.

(2)

(*d*) State whether or not you think the regression model is reasonable

(i) for the range of x-values given in the table,

(ii) for all possible x-values.

In each case justify your answer either by giving a reason for accepting the model or by suggesting an alternative model.

(2)

END