Paper Reference(s)

6683

Edexcel GCE

Statistics S1

Advanced/Advanced Subsidiary

Friday 11 June 2004 – Morning

Time: 1 hour 30 minutes

Materials required for examination	Items included with question papers
Answer Book (AB16) Graph Paper (ASG2) Mathematical Formulae (Lilac)	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 six 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. A fair die has six faces numbered 1, 2, 2, 3, 3 and 3. The die is rolled twice and the number showing on the uppermost face is recorded each time.

Find the probability that the sum of the two numbers recorded is at least 5.

(5)

---

2. A researcher thinks there is a link between a person's height and level of confidence. She measured the height h, to the nearest cm, of a random sample of 9 people. She also devised a test to measure the level of confidence c of each person. The data are shown in the table below.

h	179	169	187	166	162	193	161	177	168
c	569	561	579	561	540	598	542	565	573

[You may use $\Sigma h^2 = 272\,094$, $\Sigma c^2 = 2\,878\,966$, $\Sigma hc = 884\,484$]

(a) Draw a scatter diagram to illustrate these data.

(4)

(b) Find exact values of S_{hc} S_{hh} and S_{cc}.

(4)

(c) Calculate the value of the product moment correlation coefficient for these data.

(3)

(d) Give an interpretation of your correlation coefficient.

(1)

(e) Calculate the equation of the regression line of c on h in the form $c = a + bh$.

(3)

(f) Estimate the level of confidence of a person of height 180 cm.

(2)

(g) State the range of values of h for which estimates of c are reliable.

(1)

---

3. A discrete random variable X has a probability function as shown in the table below, where a and b are constants.

x	0	1	2	3
$P(X = x)$	0.2	0.3	b	a

Given that $E(X) = 1.7$,

(a) find the value of a and the value of b.

(5)

Find

(b) $P(0 < X < 1.5)$,

(1)

(c) $E(2X - 3)$.

(2)

(d) Show that $Var(X) = 1.41$.

(3)

(e) Evaluate $Var(2X - 3)$.

(2)

---

4. The attendance at college of a group of 18 students was recorded for a 4-week period.

The number of students actually attending each of 16 classes are shown below.

18	18	17	17
16	17	16	18
18	14	17	18
15	17	18	16

(*a*) (i) Calculate the mean and the standard deviation of the number of students attending these classes.

(ii) Express the mean as a percentage of the 18 students in the group.

(5)

In the same 4-week period, the attendance of a different group of 20, students is shown below.

20	16	18	19
15	14	14	15
18	15	16	17
16	18	15	14

(*b*) Construct a back-to-back stem and leaf diagram to represent the attendance in both groups.

(5)

(*c*) Find the mode, median and inter-quartile range for each group of students.

(6)

The mean percentage attendance and standard deviation for the second group of students are 81.25 and 1.82 respectively.

(*d*) Compare and contrast the attendance of these 2 groups of students.

(3)

---

5. A health club lets members use, on each visit, its facilities for as long as they wish. The club's records suggest that the length of a visit can be modelled by a normal distribution with mean 90 minutes. Only 20% of members stay for more than 125 minutes.

(*a*) Find the standard deviation of the normal distribution.

(4)

(*b*) Find the probability that a visit lasts less than 25 minutes.

(3)

The club introduce a closing time of 10:00 pm. Tara arrives at the club at 8:00 pm.

(*c*) Explain whether or not this normal distribution is still a suitable model for the length of her visit.

(2)

---

6. Three events A, B and C are defined in the sample space S. The events A and B are mutually exclusive and A and C are independent.

(*a*) Draw a Venn diagram to illustrate the relationships between the 3 events and the sample space.

(3)

Given that $P(A) = 0.2$, $P(B) = 0.4$ and $P(A \cup C) = 0.7$, find

(*b*) $P(A \mid C)$,

(2)

(*c*) $P(A \cup B)$,

(2)

(*d*) $P(C)$.

(4)

END