Edexcel GCE

Statistics S1
Advanced/Advanced Subsidiary
Tuesday 5 November 2002 − Morning
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. (*a*) Explain briefly why statistical models are used when attempting to solve real-world problems.

(2)

(*b*) Write down the name of the distribution you would recommend as a suitable model for each of the following situations.

(i) The weight of marmalade in a jar.

(ii) The number on the uppermost face of a fair die after it has been rolled.

(2)

---

2. There are 125 sixth-form students in a college, of whom 60 are studying only arts subjects, 40 only science subjects and the rest a mixture of both.

Three students are selected at random, *without replacement*.

Find the probability that

(*a*) all three students are studying only arts subjects,

(4)

(*b*) exactly one of the three students is studying only science subjects.

(3)

---

3. The events A and B are independent such that $P(A) = 0.25$ and $P(B) = 0.30$.

Find

(*a*) $P(A \cap B)$,

(2)

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

(2)

(*c*) $P(A \,|\, B')$.

(4)

---

4. Strips of metal are cut to length L cm, where $L \sim N(\mu, 0.5^2)$.

(*a*) Given that 2.5% of the cut lengths exceed 50.98 cm, show that $\mu = 50$.

(5)

(*b*) Find $P(49.25 < L < 50.75)$.

(4)

Those strips with length either less than 49.25 cm or greater than 50.75 cm cannot be used.

Two strips of metal are selected at random.

(*c*) Find the probability that both strips cannot be used.

(2)

---

5. An agricultural researcher collected data, in appropriate units, on the annual rainfall x and the annual yield of wheat y at 8 randomly selected places.

The data were coded using $s = x - 6$ and $t = y - 20$ and the following summations were obtained.

$$\Sigma s = 48.5, \quad \Sigma t = 65.0, \quad \Sigma s^2 = 402.11, \quad \Sigma t^2 = 701.80, \quad \Sigma st = 523.23$$

(*a*) Find the equation of the regression line of t on s in the form $t = p + qs$.

(7)

(*b*) Find the equation of the regression line of y on x in the form $y = a + bx$, giving a and b to 3 decimal places.

(3)

The value of the product moment correlation coefficient between s and t is 0.943, to 3 decimal places.

(*c*) Write down the value of the product moment correlation coefficient between x and y. Give a justification for your answer.

(2)

Turn over

6. The discrete random variable X has the following probability distribution.

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

(a) Given that $E(X) = -0.2$, find the value of α and the value of β.

(6)

(b) Write down F(0.8).

(1)

(a) Evaluate Var(X).

(4)

Find the value of

(d) $E(3X - 2)$,

(2)

(e) Var $(2X + 6)$.

(2)

7. The following stem and leaf diagram shows the aptitude scores x obtained by all the applicants for a particular job.

Aptitude score		3	1 means 31
3	1 2 9		(3)
4	2 4 6 8 9		(5)
5	1 3 3 5 6 7 9		(7)
6	0 1 3 3 3 5 6 8 8 9		(10)
7	1 2 2 2 4 5 5 5 6 8 8 8 8 9		(14)
8	0 1 2 3 5 8 8 9		(8)
9	0 1 2		(3)

(*a*) Write down the modal aptitude score.

(1)

(*b*) Find the three quartiles for these data.

(3)

Outliers can be defined to be outside the limits $Q_1 - 1.0(Q_3 - Q_1)$ and $Q_3 + 1.0(Q_3 - Q_1)$.

(*c*) On a graph paper, draw a box plot to represent these data.

(7)

For these data, $\Sigma x = 3363$ and $\Sigma x^2 = 238\,305$.

(*d*) Calculate, to 2 decimal places, the mean and the standard deviation for these data.

(3)

(*e*) Use two different methods to show that these data are negatively skewed.

(4)

END