6685

Edexcel GCE

Statistics S3

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)	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 S3), the paper reference (6685), 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. There are 64 girls and 56 boys in a school.

Explain briefly how you could take a random sample of 15 pupils using

(*a*) a simple random sample,

(3)

(*b*) a stratified sample.

(3)

---

2. A random sample of 8 students sat examinations in Geography and Statistics. The product moment correlation coefficient between their results was 0.572 and the Spearman rank correlation coefficient was 0.655.

(*a*) Test both of these values for positive correlation. Use a 5% level of significance.

(6)

(*b*) Comment on your results.

(2)

---

3. It is known from past evidence that the weight of coffee dispensed into jars by machine *A* is normally distributed with mean μ_A and standard deviation 2.5 g. Machine *B* is known to dispense the same nominal weight of coffee into jars with mean μ_B and standard deviation 2.3 g. A random sample of 10 jars filled by machine *A* contained a mean weight of 249 g of coffee. A random sample of 15 jars filled by machine *B* contained a mean weight of 251 g.

(*a*) Test, at the 5% level of significance, whether or not there is evidence that the population mean weight dispensed by machine B is greater than that of machine A.

(7)

(*b*) Write down an assumption needed to carry out this test.

(1)

---

4. Kylie regularly travels from home to visit a friend. On 10 randomly selected occasions the journey time *x* minutes was recorded. The results are summarised as follows.

$$\Sigma x = 753, \qquad \Sigma x^2 = 57\ 455.$$

(*a*) Calculate unbiased estimates of the mean and the variance of the population of journey times.

(3)

After many journeys, a random sample of 100 journeys gave a mean of 74.8 minutes and a variance of 84.6 minutes^2.

(*b*) Calculate a 95% confidence interval for the mean of the population of journey times.

(5)

(*c*) Write down two assumptions you made in part (*b*).

(2)

5. A random sample of 500 adults completed a questionnaire on how often they took part in some form of exercise. They gave a response of 'never', 'sometimes' or 'regularly'. Of those asked, 52% were females of whom 10% never exercised and 35% exercised regularly. Of the males, 12.5% never exercised and 55% sometimes exercised.

Test, at the 5% level of significance, whether or not there is any association between gender and the amount of exercise. State your hypotheses clearly.

(12)

---

6 Three six-sided dice, which were assumed to be fair, were rolled 250 times. On each occasion the number X of sixes was recorded. The results were as follows.

Number of sixes	0	1	2	3
Frequency	125	109	13	3

(*a*) Write down a suitable model for X.

(2)

(*b*) Test, at the 1% level of significance, the suitability of your model for these data.

(11)

(*c*) Explain how the test would have been modified if it had not been assumed that the dice were fair.

(2)

---

7. The random variable D is defined as

$$D = A - 3B + 4C$$

where $A \sim N(5, 2^2)$, $B \sim N(7, 3^2)$ and $C \sim N(9, 4^2)$, and A, B and C are independent.

(*a*) Find $P(D < 44)$.

(9)

The random variables B_1, B_2 and B_3 are independent and each has the same distribution as B. The random variable X is defined as

$$X = A - \sum_{i=1}^{3} B_i + 4C.$$

(*b*) Find $P(X > 0)$.

(7)

END