6686/01

Edexcel GCE

Statistics

Unit S4 Mock paper

Advanced Subsidiary / Advanced

Time: 1 hour 30 minutes

Materials required for the examination	Items included with these question papers
Answer Book (AB04) Graph Paper (GP02) Mathematical Formulae	Nil

Candidates may use any calculator EXCEPT **those with a facility for symbolic algebra, differentiation and/or integration. Thus candidates may** NOT **use calculators such as Texas TI 89, TI 92, Casio CFX 9970G, Hewlett Packard HP 48G.**

Instructions to Candidates

In the boxes on the Answer Book provided, write the name of the Examining Body (Edexcel), your Centre Number, Candidate Number, the Unit Title (Statistics S4), the Paper Reference (6686), your surname, other names 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 including Statistical Formulae and Tables' is provided.

Full marks may be obtained for answers to ALL questions.

This paper has 6 questions. Pages 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.

Edexcel
Success through qualifications

1. The weights of the contents of jars of jam are normally distributed with a stated mean of 100 g. A random sample of 7 jars was taken and the contents of each jar, *x* grams, was weighed. The results are summarised by the following statistics.

$$\sum x = 710.9, \quad \sum x^2 = 72\,219.45.$$

Test at the 5% level of significance whether or not there is evidence that the mean weight of the contents of the jars is greater than 100 g. State your hypotheses clearly. **(8 marks)**

---

2. An engineer decided to investigate whether or not the strength of rope was affected by water. A random sample of 9 pieces of rope was taken and each piece was cut in half. One half of each piece was soaked in water over night, and then each piece of rope was tested to find its strength. The results, in coded units, are given in the table below

Rope no.	1	2	3	4	5	6	7	8	9
Dry rope	9.7	8.5	6.3	8.3	7.2	5.4	6.8	8.1	5.9
Wet rope	9.1	9.5	8.2	9.7	8.5	4.9	8.4	8.7	7.7

Assuming that the strength of rope follows a normal distribution, test whether or not there is any difference between the mean strengths of dry and wet rope. State your hypotheses clearly and use a 1% level of significance. **(8 marks)**

---

3. A certain vaccine is known to be only 70% effective against a particular virus; thus 30% of those vaccinated will actually catch the virus. In order to test whether or not a new and more expensive vaccine provides better protection against the same virus, a random sample of 30 people were chosen and given the new vaccine. If fewer than 6 people contracted the virus the new vaccine would be considered more effective than the current one.

(*a*) Write down suitable hypotheses for this test. **(1 mark)**

(*b*) Find the probability of making a Type I error. **(2 marks)**

(*c*) Find the power of this test if the new vaccine is

(i) 80% effective,

(ii) 90% effective. **(3 marks)**

An independent research organisation decided to test the new vaccine on a random sample of 50 people to see if it could be considered more than 70% effective. They required the probability of a Type I error to be as close as possible to 0.05.

(*d*) Find the critical region for this test. **(2 marks)**

(*e*) State the size of this critical region. **(1 mark)**

(*f*) Find the power of this test if the new vaccine is

(i) 80% effective,

(ii) 90% effective. **(2 marks)**

(*g*) Give one advantage and one disadvantage of the second test. **(2 marks)**

---

4. Gill, a member of the accounts department in a large company, is studying the expenses claims of company employees. She assumes that the claims, in £, follow a normal distribution with mean μ and variance σ^2. As a first stage in her investigation she took the following random sample of 10 claims.

30.85, 99.75, 142.73, 223.16, 75.43, 28.57, 53.90, 81.43, 68.62, 43.45.

(a) Find a 95% confidence interval for μ. **(6 marks)**

The chief accountant would like a 95% confidence interval where the difference between the upper confidence limit and the lower confidence limit is less than 20.

(*b*) Show that $\frac{\sigma^2}{n}$ < 26.03 (to 2 decimal places), where *n is* the size of the sample required to achieve this. **(3 marks)**

Gill decides to use her original sample of 10 to obtain a value for σ^2 so that the chance of her value being an underestimate is 0.01.

(*c*) Find such a value for σ^2. **(3 marks)**

(*d*) Use this value for σ^2 to estimate the size of sample the chief accountant requires. **(2 marks)**

5. An educational researcher is testing the effectiveness of a new method of teaching a topic in mathematics. A random sample of 10 children were taught by the new method and a second random sample of 9 children, of similar age and ability, were taught by the conventional method. At the end of the teaching, the same test was given to both groups of children.

The marks obtained by the two groups are summarised in the table below.

	New method	Conventional method
Mean ($\bar{x}$)	82.3	78.2
Standard deviation (s)	3.5	5.7
Number of students (n)	10	9

(*a*) Stating your hypotheses clearly and using a 5% level of significance, investigate whether or not

(i) the variance of the marks of children taught by the conventional method is greater than that of children taught by the new method, **(4 marks)**

(ii) the mean score of children taught by the conventional method is lower than the mean score of those taught by the new method. **(6 marks)**

[In each case you should give full details of the calculation of the test statistics.]

(*b*) State any assumptions you made in order to carry out these tests. **(1 mark)**

(*c*) Find a 95% confidence interval for the common variance of the marks of the two groups. **(5 marks)**

6. A statistics student is trying to estimate the probability, p, of rolling a 6 with a particular die. The die is rolled 10 times and the random variable X_1 represents the number of sixes obtained. The random variable $R_1 = \frac{X_1}{10}$ is proposed as an estimator of p.

(a) Show that R_1 is an unbiased estimator of p. **(1 mark)**

The student decided to roll the die again n times ($n > 10$) and the random variable X_2 represents the number of sixes in these n rolls. The random variable $R_2 = \frac{X_2}{n}$ and the random variable $Y = \frac{1}{2}(R_1 + R_2)$.

(b) Show that both R_2 and Y are unbiased estimators of p. **(2 marks)**

(c) Find Var (R_2) and Var (Y). **(3 marks)**

(d) State giving a reason which of the 3 estimators R_1, R_2 and Y are consistent estimators of p. **(2 marks)**

(e) For the case $n = 20$ state, giving a reason, which of the 3 estimators R_1, R_2 and Y you would recommend. **(4 marks)**

The student's teacher pointed out that a better estimator could be found based on the random variable $X_1 + X_2$.

(f) Find a suitable estimator and explain why it is better than R_1, R_2 and Y. **(6 marks)**

END