GCE Examinations

Statistics
Module S3

Advanced Subsidiary / Advanced Level

Paper F

Time: 1 hour 30 minutes

Instructions and Information

Candidates may use any calculator except those with a facility for symbolic algebra and / or calculus.

Full marks may be obtained for answers to ALL questions.

Mathematical and statistical formulae and tables are available.

This paper has 7 questions.

Advice to Candidates

You must show sufficient working to make your methods clear to an examiner. Answers without working will gain no credit.



Written by Shaun Armstrong & Chris Huffer

© Solomon Press

These sheets may be copied for use solely by the purchaser's institute.

1. A personnel manager has details on all company employees and wishes to consult a sample of them on a possible change to the company's hours of business. She decides to take a stratified sample based on different age groups.

(a) Give one advantage of using stratified sampling in this situation. **(1 mark)**

The manager needs to select a sample of size 10, without replacement, from a list of 65 employees aged 16 to 25. She numbers these employees from 01 to 65 in alphabetical order and uses the table of random numbers given in the formula book. She starts with the top of the sixth two-digit column and works down. The first two numbers she writes down are 30 and 47.

(b) Find the other eight numbers in the sample. **(3 marks)**

(c) Suggest another factor that might be useful to consider in deciding on the strata. **(1 mark)**

---

2. A Geography teacher is interested in the link between mathematical ability and the ability to visualise three-dimensional situations. He gives a group of 15 students a test and records each student's score, m, on the mathematics questions and each student's score, v, on the visio-spatial questions.

He calculates the following summary statistics:

$$S_{mm} = 3747.73, \quad S_{vv} = 2791.33, \quad S_{mv} = 2564.33$$

(a) Calculate the product moment correlation coefficient for these data. **(2 marks)**

(b) Stating your hypotheses clearly and using a 5% level of significance test the theory that students who are good at Mathematics tend to have better visio-spatial awareness. **(4 marks)**

---

3. A random variable X is distributed normally with a standard deviation of 6.8

Sixty observations of X are made and found to have a mean of 31.4

(a) Find a 90% confidence interval for the mean of X. **(4 marks)**

(b) How many observations of X would be needed in order to obtain a 90% confidence interval for the mean of X with a width of less than 1.5 **(5 marks)**

© *Solomon Press*

4. A paranormal investigator invites couples who believe they have a telepathic connection to participate in a trial. With each couple one person looks at a card with one of five shapes on it and the other person says which of the shapes they think it is. This is repeated six times and the number of correct answers recorded. The results from 120 couples are given below.

Number Correct	0	1	2	3	4	5	6
Number of Couples	26	56	28	8	2	0	0

The investigator wishes to see if this data fits a binomial distribution with parameters $n = 6$ and $p = \frac{1}{5}$ and calculates to 2 decimal places the expected frequencies given below.

Number Correct	0	1	2	3	4	5	6
Expected Frequency				9.83	1.84	0.18	0.01

(a) Find the other expected frequencies. **(3 marks)**

(b) Stating your hypotheses clearly, test at the 5% level of significance whether or not the distribution is an appropriate model.

(8 marks)

(c) Comment on your findings. **(1 mark)**

---

5. A Policy Unit wished to find out whether attitudes to the European Union varied with age. It conducted a survey asking 200 individuals to which of three age groups they belonged and whether they regarded themselves as generally pro-Europe or Eurosceptic. The results are shown in the table below.

	Pro-Europe	Eurosceptic
18 – 34 years	43	21
35 – 54 years	30	36
55 years or over	27	43

(a) Stating your hypotheses clearly, test at the 5% level of significance whether attitudes to Europe are associated with age.

(11 marks)

The survey also asked people if they voted at the last election. When the above test was repeated using only the results from those who had voted a value of 4.872 was calculated for $\sum \frac{(O-E)^2}{E}$. No classes were combined.

(b) Find if this value leads to a different result. **(2 marks)**

Turn over

© *Solomon Press*

6. Four swimmers, A, B, C and D, are to be used in a 4×100 metres freestyle relay. The time for each swimmer to complete a leg follows a normal distribution. The mean and standard deviation, in seconds, of the time for each swimmer to complete a leg and the order in which they are to swim are shown in the table below.

	mean	standard deviation
1st leg – A	63.1	1.2
2nd leg – B	65.7	1.5
3rd leg – C	65.4	1.8
4th leg – D	62.5	0.9

(a) Find the probability that the total time for first two legs is less than the total time for the last two.

(6 marks)

The total time for another team to complete this relay is normally distributed with a mean of 259.0 seconds and a standard deviation of 3.4 seconds. The two teams are to compete over four races.

(b) Find the probability that the first team wins all four races, assuming that the team's performances are not affected by previous results.

(8 marks)

---

7. A telephone company believes that, for young people, the average length of a telephone call on a land line is longer than on a mobile, due to the difference in price.

The company collected data on the time, t minutes, of 500 calls made by young people on mobiles and the data is summarised by

$$\Sigma t = 7335, \qquad \Sigma t^2 = 172\,040.$$

(a) Calculate unbiased estimates of the mean and variance of t. **(5 marks)**

For 200 calls made on land lines by the same young people, unbiased estimates of the mean and variance of the call length were 15.9 minutes and 108.5 minutes^2 respectively.

(b) Stating your hypotheses clearly, test at the 5% level whether or not there is evidence that longer calls are made on land lines than on mobiles.

(9 marks)

(c) Explain the importance of the central limit theorem in carrying out the test in part *(b)*.

(2 marks)

END

© *Solomon Press*