GCE Examinations

Statistics
Module S3

Advanced Subsidiary / Advanced Level

Paper A

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

1. A museum is open to the public for six hours a day from Monday to Friday every week. The number of visitors, V, to the museum on ten randomly chosen days were as follows:

182 172 113 99 168 183 135 129 150 108

(a) Calculate an unbiased estimate of the mean of V. **(2 marks)**

Assuming that V is normally distributed with a variance of 130,

(b) find a 95% confidence interval for the mean of V. **(4 marks)**

---

2. *(a)* Explain what is meant by a simple random sample. **(2 marks)**

(b) Explain briefly how you could use a table of random numbers to select a simple random sample of size 12 from a list of the 70 junior members of a tennis club. **(3 marks)**

(c) Give an example of a situation in which you might choose to take a stratified sample and explain why. **(2 marks)**

---

3. The time that a school pupil spends on French homework each week is normally distributed with a mean of 55 minutes and a standard deviation of 10 minutes.

The time that this pupil spends on English homework each week is normally distributed with a mean of 1 hour 30 minutes and a standard deviation of 18 minutes.

Find the probability that in a randomly chosen week

(a) the pupil spends more than 2 hours in total doing French and English homework, **(5 marks)**

(b) the pupil spends more than twice as long doing English homework as he spends doing French homework. **(6 marks)**

---

4. A group of 40 males and 40 females were asked which of three “Reality TV” shows they liked most – *Watched*, *Stranded* or *One-2-Win*. The results were as follows:

	Watched	*Stranded*	*One-2-Win*
Males	21	6	13
Females	15	10	15

Stating your hypotheses clearly, test at the 10% level whether or not there is a significant difference in the preferences of males and females.

(11 marks)

---

5. A marathon runner believes that she is more likely to win a medal at her national championships the higher the temperature is on the day of the race.

She records the temperature at the start of each of eight races against fields of a similar standard and her finishing position in each race. Her results are shown in the table below.

Temperature (°C)	16	9	11	5	7	21	12	15
Finishing position	2	15	5	19	10	4	6	11

(a) Calculate Spearman’s rank correlation coefficient for these data. **(6 marks)**

(b) Using a 5% level of significance and stating your hypotheses clearly, interpret your result.

(4 marks)

Another runner suggests that she should use her time in each race instead of her finishing position and calculate the product moment correlation coefficient for the data.

(c) Comment on this suggestion. **(2 marks)**

Turn over

6. The weight of a particular electrical component is normally distributed with a mean of 46.7 grams and a variance of 1.8 grams^2. The component is sold in boxes of 12.

(a) State the distribution of the mean weight of the components in one box. **(2 marks)**

(b) Find the probability that the mean weight of the components in a randomly chosen box is more than 47 grams.

(3 marks)

After a break in production the component manufacturer wishes to find out if the mean weight of the components has changed. A random sample of 30 components is found to have a mean weight of 46.5 grams.

(c) Assuming that the variance of the weight of the components is unchanged, test at the 5% level of significance if there has been any change in the mean weight of the components.

(7 marks)

---

7. A student collects data on whether competitors in local tennis tournaments are right, or left-handed. The table below shows the number of left-handed players who reached the last 16 for fifty tournaments.

No. of Left-handed Players	0	1	2	3	4	≥ 5
No. of Tournaments	4	12	18	11	5	0

The student believes that a binomial distribution with $n = 16$ and $p = 0.1$ could be a suitable model for these data.

(a) Stating your hypotheses clearly test the student's model at the 5% level of significance.

(13 marks)

To improve the model the student decides to estimate p using the data in the table. Using this value of p to calculate expected frequencies the student had 5 classes after combining and calculated that $\sum \frac{(O-E)^2}{E} = 2.127$

(b) Test at the 5% level of significance whether or not the binomial distribution is a suitable model for the number of left-handed players who reach the last 16 in local tennis tournaments.

(3 marks)

END