GCE Examinations

Statistics
Module S1

Advanced Subsidiary / Advanced Level

Paper C

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)* Explain briefly what you understand by a statistical model. **(2 marks)**

A zoologist is analysing data on the weights of adult female otters.

(b) Name a distribution that you think might be suitable for modelling such data.
(1 mark)

(c) Describe two features that you would expect to find in the distribution of the weights of adult female otters and that led to your choice in part *(b)*.
(2 marks)

(d) Why might your choice in part *(b)* not be suitable for modelling the weights of all adult otters?
(1 mark)

---

2. For a geography project a student studied weather records kept by her school since 1993. To see if there was any evidence of global warming she worked out the mean temperature in degrees Celsius at noon for the month of June in each year.

Her results are shown in the table below.

Year	1993	1994	1995	1996	1997	1998	1999	2000
Mean temperature (°C)	21.9	24.1	20.7	23.0	24.2	22.1	22.6	23.9

(a) Plot a scatter diagram showing these data. **(3 marks)**

The student wanted to investigate further whether or not her data provided evidence of an increase in temperature in June each year. Using Y for the number of years since 1993 and T for the mean temperature, she calculated the following summary statistics.

$$\Sigma Y = 28, \quad \Sigma T = 182.5, \quad \Sigma Y^2 = 140, \quad \Sigma T^2 = 4173.93, \quad \Sigma YT = 644.7.$$

(b) Calculate the product moment correlation coefficient for these data. **(5 marks)**

(c) Comment on your result in relation to the student's enquiry. **(1 mark)**

© *Solomon Press*

3. In a study of 120 pet-owners it was found that 57 owned at least one dog and of these 16 also owned at least one cat. There were 35 people in the group who didn't own any cats or dogs.

As an incentive to take part in the study, one participant is chosen at random to win a year's free supply of pet food.

Find the probability that the winner of this prize

(a) owns a dog but does not own a cat, **(2 marks)**

(b) owns a cat, **(4 marks)**

(c) does not own a cat given that they do not own a dog. **(4 marks)**

---

4. An internet service provider runs a series of television adverts at weekly intervals.
To investigate the effectiveness of the adverts the company recorded the viewing figures in millions, v, for the programme in which the advert was shown, and the number of new customers, c, who signed up for their service the next day.

The results are summarised as follows.

$$\bar{v} = 4.92, \quad \bar{c} = 104.4, \quad S_{vc} = 594.05, \quad S_{vv} = 85.44.$$

(a) Calculate the equation of the regression line of c on v in the form $c = a + bv$. **(4 marks)**

(b) Give an interpretation of the constants a and b in this context. **(2 marks)**

(c) Estimate the number of customers that will sign up with the company the day after an advert is shown during a programme watched by 3.7 million viewers. **(2 marks)**

(d) State two other factors besides viewing figures that will affect the success of an advert in gaining new customers for the company. **(2 marks)**

Turn over

© *Solomon Press*

5. The time taken in minutes, T, for a mechanic to service a bicycle follows a normal distribution with a mean of 25 minutes and a variance of 16 minutes^2.

Find

(a) $P(T < 28)$, **(3 marks)**

(b) $P(|T - 25| < 5)$. **(4 marks)**

One afternoon the mechanic has 3 bicycles to service.

(c) Find the probability that he will take less than 23 minutes on each of the three bicycles.

(4 marks)

---

6. The number of people visiting a new art gallery each day is recorded over a three-month period and the results are summarised in the table below.

Number of visitors	Number of days
400 - 459	3
460 - 479	8
480 - 499	13
500 - 519	12
520 - 539	18
540 - 559	11
560 - 599	9
600 - 699	5

(a) Draw a histogram on graph paper to illustrate these data. **(4 marks)**

In order to calculate summary statistics for the data it is coded using $y = \frac{x - 509.5}{10}$, where x is the mid-point of each class.

(b) Find $\sum fy$. **(3 marks)**

You may assume that $\sum fy^2 = 2041$.

(c) Using these values for $\sum fy$ and $\sum fy^2$, calculate estimates of the mean and standard deviation of the number of visitors per day.

(6 marks)

© Solomon Press

7. A bag contains 4 red and 2 blue balls, all of the same size. A ball is selected at random and removed from the bag. This is repeated until a blue ball is pulled out of the bag.

The random variable B is the number of balls that have been removed from the bag.

(a) Show that $P(B = 2) = \frac{4}{15}$. **(2 marks)**

(b) Find the probability distribution of B. **(4 marks)**

(c) Find $E(B)$. **(3 marks)**

The bag and the same 6 balls are used in a game at a funfair. One ball is removed from the bag at a time and a contestant wins 50 pence if one of the first two balls picked out is blue.

(d) What are the expected winnings from playing this game once? **(4 marks)**

For £1, a contestant gets to play the game three times.

(e) What is the expected profit or loss from the three games? **(3 marks)**

END

© *Solomon Press*