GCE Examinations

Statistics
Module S1

Advanced Subsidiary / Advanced Level

Paper D

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)* Draw two separate scatter diagrams, each with eight points, to illustrate the relationship between x and y in the cases where they have a product moment correlation coefficient equal to

(i) exactly + 1,

(ii) about $^{-}$ 0.4. **(4 marks)**

(b) Explain briefly how the conclusion you would draw from a product moment correlation coefficient of + 0.3 would vary according to the number of pairs of data used in its calculation.

(2 marks)

---

2. A histogram was drawn to show the distribution of age in completed years of the participants on an outward-bound course.

There were 32 people aged 30 - 34 years on the course. The height of the rectangle representing this group was 19.2 cm and it was 1 cm in width.

Given that there were 28 people aged 35 - 39 years,

(a) find the height of the rectangle representing this group. **(3 marks)**

Given that the height of the rectangle representing people aged 40 - 59 years was 2.7 cm,

(b) find the number of people on the course in this age group. **(3 marks)**

---

3. The events A and B are such that

$$\mathrm{P}(A) = \tfrac{7}{12}, \ \mathrm{P}(A \cap B) = \tfrac{1}{4} \ \text{ and } \ \mathrm{P}(A \mid B) = \tfrac{2}{3}.$$

Find

(a) $\mathrm{P}(B)$, **(3 marks)**

(b) $\mathrm{P}(A \cup B)$, **(3 marks)**

(c) $\mathrm{P}(B \mid A')$. **(3 marks)**

© *Solomon Press*

4. The owner of a mobile burger-bar believes that hot weather reduces his sales.

To investigate the effect on his business he collected data on his daily sales, £P, and the maximum temperature, T °C, on each of 20 days. He then coded the data, using $x = T - 20$ and $y = P - 300$, and calculated the summary statistics given below.

$$\Sigma x = 57, \quad \Sigma y = 2222, \quad \Sigma x^2 = 401, \quad \Sigma y^2 = 305576, \quad \Sigma xy = 3871.$$

(a) Find an equation of the regression line of P on T. **(9 marks)**

The owner of the bar doesn't believe it is profitable for him to run the bar if he takes less than £460 in a day.

(b) According to your regression line at what maximum daily temperature, to the nearest degree Celsius, does it become unprofitable for him to run the bar? **(3 marks)**

---

5. The discrete random variable X has the probability function shown below.

$$P(X = x) = \begin{cases} kx, & x = 2,\ 3,\ 4,\ 5,\ 6, \\ 0, & \text{otherwise.} \end{cases}$$

(a) Find the value of k. **(2 marks)**

(b) Show that $\mathrm{E}(X) = \frac{9}{2}$. **(3 marks)**

Find

(c) $\mathrm{P}[X > \mathrm{E}(X)]$, **(2 marks)**

(d) $\mathrm{E}(2X - 5)$, **(2 marks)**

(e) $\mathrm{Var}(X)$. **(4 marks)**

Turn over

© *Solomon Press*

6. A geologist is analysing the size of quartz crystals in a sample of granite. She estimates that the longest diameter of 75% of the crystals is greater than 2 mm, but only 10% of the crystals have a longest diameter of more than 6 mm.

The geologist believes that the distribution of the longest diameters of the quartz crystals can be modelled by a normal distribution.

(a) Find the mean and variance of this normal distribution. **(9 marks)**

The geologist also estimated that only 2% of the longest diameters were smaller than 1 mm.

(b) Calculate the corresponding percentage that would be predicted by a normal distribution with the parameters you calculated in part *(a)*.

(3 marks)

(c) Hence, comment on the suitability of the normal distribution as a model in this situation.

(2 marks)

---

7. Jane and Tahira play together in a basketball team. The list below shows the number of points that Jane scored in each of 30 games.

39	19	28	30	18	21	23	15	34	24
29	17	43	12	24	25	41	19	26	40
45	23	21	32	37	24	18	15	24	36

(a) Construct a stem and leaf diagram for these data. **(3 marks)**

(b) Find the median and quartiles for these data. **(4 marks)**

(c) Represent these data with a boxplot. **(3 marks)**

Tahira played in the same 30 games and her lowest and highest points total in a game were 19 and 41 respectively. The quartiles for Tahira were 27, 31 and 35 respectively.

(d) Using the same scale draw a boxplot for Tahira's points totals. **(2 marks)**

(e) Compare and contrast the number of points scored per game by Jane and Tahira.

(3 marks)

END

© *Solomon Press*