GCE Examinations

Statistics
Module S1

Advanced Subsidiary / Advanced Level

Paper G

Time: 1 hour 30 minutes

Instructions and Information

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

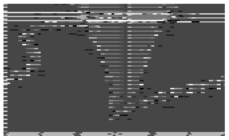
Full marks may be obtained for answers to ALL questions.

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. The discrete random variable Y has the following probability distribution.

y	$^{-}2$	$^{-}1$	0	1	2
$P(Y = y)$	0.1	0.15	0.2	0.3	0.25

Find

(a) $F(0.5)$, **(1 mark)**

(b) $P(^{-}1 < Y < 1.9)$, **(2 marks)**

(c) $E(Y)$, **(2 marks)**

(d) $E(3Y - 1)$. **(2 marks)**

---

2. A supermarket manager believes that those of her staff on lower rates of pay tend to work more hours of overtime.

(a) Suggest why this might be the case. **(1 mark)**

To investigate her theory the manager recorded the number of hours of overtime, h, worked by each of the store's 18 full-time staff during one week. She also recorded each employee's hourly rate of pay, £p, and summarised her results as follows:

$$\Sigma p = 86, \quad \Sigma h = 104.5, \quad \Sigma p^2 = 420.58, \quad \Sigma h^2 = 830.25, \quad \Sigma ph = 487.3$$

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

(c) Comment on the manager's hypothesis. **(2 marks)**

---

3. A magazine collected data on the total cost of the reception at each of a random sample of 80 weddings.

The data is grouped and coded using $y = \dfrac{C - 3250}{250}$, where C is the mid-point in pounds of each class, giving $\Sigma fy = 37$ and $\Sigma fy^2 = 2317$.

(a) Using these values, calculate estimates of the mean and standard deviation of the cost of the receptions in the sample. **(6 marks)**

(b) Explain why your answers to part *(a)* are only estimates. **(1 mark)**

The median of the data was £3050.

(c) Comment on the skewness of the data and suggest a reason for it. **(2 marks)**

---

4. The random variable A is normally distributed with a mean of 32.5 and a variance of 18.6

Find

(a) P($A < 38.2$), **(3 marks)**

(b) P($31 \leq A \leq 35$), **(4 marks)**

The random variable B is normally distributed with a standard deviation of 7.2

Given also that P($B > 110$) = 0.138,

(c) find the mean of B. **(4 marks)**

---

5. A group of children were each asked to try and complete a task to test hand-eye coordination. Each child repeated the task until he or she had been successful or had made four attempts.

The number of attempts made by the children in the group are summarised in the table below.

Number of attempts	1	2	3	4
Number of children	43	26	13	3

(a) Calculate the mean and standard deviation of the number of attempts made by each child.

(5 marks)

It is suggested that the number of attempts made by each child could be modelled by a discrete random variable X with the probability function

$$P(X = x) = \begin{cases} k(20 - x^2), & x = 1,\ 2,\ 3,\ 4, \\ 0, & \text{otherwise.} \end{cases}$$

(b) Show that $k = \frac{1}{50}$. **(3 marks)**

(c) Find E(X). **(2 marks)**

(d) Comment on the suitability of this model. **(1 mark)**

Turn over

6. Serving against his regular opponent, a tennis player has a 65% chance of getting his first serve in. If his first serve is in he then has a 70% chance of winning the point but if his first serve is not in, he only has a 45% chance of winning the point.

(a) Represent this information on a tree diagram. **(3 marks)**

For a point on which this player served to his regular opponent, find the probability that

(b) he won the point, **(3 marks)**

(c) his first serve went in given that he won the point, **(3 marks)**

(d) his first serve didn't go in given that he lost the point. **(5 marks)**

---

7. *Pipes-R-us* manufacture a special lightweight aluminium tubing.

The price £P, for each length, l metres, that the company sells is shown in the table.

l (metres)	0.5	0.8	1.0	1.5	2	4	6
P (£)	2.50	3.40	4.00	5.20	6.00	10.50	15.00

(a) Represent these data on a scatter diagram. **(3 marks)**

You may use

$$\Sigma l = 15.8, \quad \Sigma P = 46.6, \quad \Sigma l^2 = 60.14, \quad \Sigma lP = 159.77$$

(b) Find the equation of the regression line of P on l in the form $P = a + bl$. **(7 marks)**

(c) Give a practical interpretation of the constant b. **(1 mark)**

In response to customer demand *Pipes-R-us* decide to start selling tubes cut to specific lengths. Initially the company decides to use the regression line found in part *(b)* as a pricing formula for this new service.

(d) Calculate the price that *Pipes-R-us* should charge for 5.2 metres of the tubing. **(2 marks)**

(e) Suggest a reason why *Pipes-R-us* might not offer prices based on the regression line for any length of tubing. **(2 marks)**

END