

GCE Examinations

Statistics

Module S1

Advanced Subsidiary / Advanced Level

Paper H

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

x	k	$k + 4$	$2k$
$P(X = x)$	$\frac{1}{8}$	$\frac{3}{8}$	$\frac{1}{2}$

(a) Find and simplify an expression in terms of k for $E(X)$. **(3 marks)**

Given that $E(X) = 9$,

(b) find the value of k . **(2 marks)**

2. (a) Explain briefly what is meant by a statistical model. **(2 marks)**

(b) State, with a reason, whether or not the normal distribution might be suitable for modelling each of the following:

(i) The number of children in a family;

(ii) The time taken for a particular employee to cycle to work each day using the same route;

(iii) The quarterly electricity bills for a particular house. **(6 marks)**

3. The probability that Ajita gets up before 6.30 am in the morning is 0.7

The probability that she goes for a run in the morning is 0.35

The probability that Ajita gets up after 6.30 am and does not go for a run is 0.22

Let A represent the event that Ajita gets up before 6.30 am and B represent the event that she goes for a run in the morning.

Find

(a) $P(A \cup B)$, **(2 marks)**

(b) $P(A \cap B')$, **(2 marks)**

(c) $P(B | A)$. **(3 marks)**

(d) State, with a reason, whether or not events A and B are independent. **(2 marks)**

4. A company produces jars of English Honey. The weight of the glass jars used is normally distributed with a mean of 122.3 g and a standard deviation of 2.6 g.

Calculate the probability that a randomly chosen jar will weigh

(a) less than 127 g, **(3 marks)**

(b) less than 121.5 g. **(3 marks)**

The weight of honey put into each jar by a machine is normally distributed with a standard deviation of 1.6 g. The machine operator can adjust the mean weight of the honey put into each jar without changing the standard deviation.

(c) Find, correct to 4 significant figures, the minimum that the mean weight can be set to such that at most 1 in 20 of the jars will contain less than 454 g. **(4 marks)**

5. The letters of the word DISTRIBUTION are written on separate cards. The cards are then shuffled and the top three are turned over.

Let the random variable V be the number of vowels that are turned over.

(a) Show that $P(V = 1) = \frac{21}{44}$. **(3 marks)**

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

(c) Find $E(V)$ and $\text{Var}(V)$. **(6 marks)**

Turn over

6. A cinema recorded the number of people at each showing of each film during a one-week period. The results are summarised in the table below.

Number of people	Number of showings
1 - 40	36
41 - 60	20
61 - 80	33
81 - 100	24
101 - 150	36
151 - 200	39
201 - 300	52

- (a) Draw a histogram on graph paper to illustrate these data. **(4 marks)**
- (b) Calculate estimates of the median and quartiles of these data. **(6 marks)**
- (c) Use your answers to part (b) to show that the data is positively skewed. **(3 marks)**
-

7. A new vaccine is tested over a six-month period in one health authority.

The table shows the number of new cases of the disease, d , reported in the m th month after the trials began.

m	1	2	3	4	5	6
d	102	69	61	58	52	48

A doctor suggests that a relationship of the form $d = a + bx$ where $x = \frac{1}{m}$ can be used to model the situation.

- (a) Tabulate the values of x corresponding to the given values of d and plot a scatter diagram of d against x . **(5 marks)**
- (b) Explain how your scatter diagram supports the suggested model. **(1 mark)**

You may use

$$\Sigma x = 2.45, \quad \Sigma d = 390, \quad \Sigma x^2 = 1.491, \quad \Sigma xd = 189.733$$

- (c) Find an equation of the regression line d on x in the form $d = a + bx$. **(7 marks)**
- (d) Use your regression line to estimate how many new cases of the disease there will be in the 13th month after the trial began. **(3 marks)**
- (e) Comment on the reliability of your answer to part (d). **(1 mark)**

END