GCE Examinations

Statistics
Module S1

Advanced Subsidiary / Advanced Level

Paper E

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. Joel buys a box of second-hand Jazz and Blues CDs at a car boot sale.

In the box there are 30 CDs, 8 of which were recorded live. 16 of the CDs are predominantly Jazz and 13 of these were recorded in the studio. This information is shown in the following table.

	Studio	Live	Total
Jazz	13		16
Blues			
Total		8	30

(a) Copy and complete the table above. **(2 marks)**

Joel picks a CD at random to play first.

Find the probability that it is

(b) a Blues CD that was recorded live, **(1 mark)**

(c) a Jazz CD, given that it was recorded in the studio. **(2 marks)**

---

2. The discrete random variable Q has the following probability distribution.

q	1	2	3	4	5
$P(Q = q)$	$\frac{1}{5}$	$\frac{1}{5}$	$\frac{1}{5}$	$\frac{1}{5}$	$\frac{1}{5}$

(a) Write down the name of this distribution. **(1 mark)**

The discrete random variable R has the following probability distribution.

r	14	24	34	44	54
$P(R = r)$	$\frac{1}{5}$	$\frac{1}{5}$	$\frac{1}{5}$	$\frac{1}{5}$	$\frac{1}{5}$

(b) State the relationship between R and Q in the form $R = aQ + b$. **(2 marks)**

Given that $E(Q) = 3$ and $Var(Q) = 2$,

(c) find $E(R)$ and $Var(R)$. **(4 marks)**

© *Solomon Press*

3. The random variable X is normally distributed with a mean of 42 and a variance of 18.

Find

(a) $P(X \leq 45)$, **(3 marks)**

(b) $P(32 \leq X \leq 38)$, **(4 marks)**

(c) the value of x such that $P(X \leq x) = 0.95$ **(4 marks)**

---

4. The ages of 300 houses in a village are recorded giving the following table of results.

Age (years)	Number of houses
0 -	36
20 -	92
40 -	74
60 -	39
100 -	14
200 -	27
300 - 500	18

Use linear interpolation to estimate for these data

(a) the median, **(3 marks)**

(b) the limits between which the middle 80% of the ages lie. **(5 marks)**

An estimate of the mean of these data is calculated to be 86.6 years.

(c) Explain why the mean and median are so different and hence say which you consider best represents the data.

(3 marks)

Turn over

© *Solomon Press*

5. The discrete random variable Y has the following cumulative distribution function.

y	0	1	2	3	4
F(Y)	0.05	0.15	0.35	0.75	1

(a) Write down the probability distribution of Y. **(2 marks)**

(b) Find $P(1 \leq Y < 3)$. **(2 marks)**

(c) Show that $E(Y) = 2.7$ **(2 marks)**

(d) Find $E(2Y + 4)$. **(2 marks)**

(e) Find $Var(Y)$. **(4 marks)**

---

6. A software company sets exams for programmers who wish to qualify to use their packages. Past records show that 55% of candidates taking the exam for the first time will pass, 60% of those taking it for the second time will pass, but only 40% of those taking the exam for the third time will pass. Candidates are not allowed to sit the exam more than three times.

A programmer decides to keep taking the exam until he passes or is allowed no further attempts. Find the probability that he will

(a) pass the exam on his second attempt, **(2 marks)**

(b) pass the exam. **(3 marks)**

Another programmer already has the qualification.

(c) Find, correct to 3 significant figures, the probability that she passed first time. **(3 marks)**

At a particular sitting of the exam there are 400 candidates.

The ratio of those sitting the exam for the first time to those sitting it for the second time to those sitting it for the third time is 5 : 3 : 2

(d) How many of the 400 candidates would be expected to pass? **(4 marks)**

© *Solomon Press*

7. A doctor wished to investigate the effects of staying awake for long periods on a person's ability to complete simple tasks. She recorded the number of times, n, that a subject could clinch his or her fist in 30 seconds after being awake for h hours.

The results for one subject were as follows.

h (hours)	16	17	18	19	20	21	22	23	24
n	116	114	109	101	94	94	86	81	80

(a) Plot a scatter diagram of n against h for these results. **(4 marks)**

You may use

$\Sigma h = 180, \quad \Sigma n = 875, \quad \Sigma h^2 = 3660, \quad \Sigma hn = 17204.$

(b) Obtain the equation of the regression line of n on h in the form $n = a + bh$. **(7 marks)**

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

(d) Explain why this regression line would be unlikely to be appropriate for values of h between 0 and 16.

(2 marks)

Another subject underwent the same tests giving rise to a regression line of $n = 213.4 - 5.87h$.

(e) After how many hours of being awake together would you expect these two subjects to be able to clench their fists the same number of times in 30 seconds?

(3 marks)

END

© *Solomon Press*