GCE Examinations

Statistics
Module S1

Advanced Subsidiary / Advanced Level

Paper K

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 6 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. There are 16 competitors in a table-tennis competition, 5 of which come from Racknor Comprehensive School. Prizes are awarded to the competitors finishing in each of first, second and third place.

Assuming that all the competitors have an equal chance of success, find the probability that the students from Racknor Comprehensive

(a) win no prizes, **(3 marks)**

(b) win the 1st and 3rd place prizes but not the 2nd place prize, **(3 marks)**

(c) win exactly one of the prizes. **(4 marks)**

---

2. A statistics student gave a questionnaire to a random sample of 50 pupils at his school. The sample included pupils aged from 11 to 18 years old.

The student summarised the data on age in completed years, A, and the number of hours spent doing homework in the previous week, H, giving the following:

$$\Sigma A = 703, \quad \Sigma H = 217, \quad \Sigma A^2 = 10\,131, \quad \Sigma H^2 = 1338.5, \quad \Sigma AH = 3253.5$$

(a) Calculate the product moment correlation coefficient for these data and explain what is shown by your result.

(6 marks)

The student also asked each pupil how many hours of paid work they had done in the previous week. He then calculated the product moment correlation coefficient for the data on hours doing homework and hours doing paid work, giving a value of $r = 0.5213$

The student concluded that paid work did not interfere with homework as pupils doing more paid work also tended to do more homework.

(b) Explain why this conclusion may not be valid. **(2 marks)**

(c) Explain briefly how the student could more effectively investigate the effect of paid work on homework.

(2 marks)

© *Solomon Press*

3. A soccer fan collected data on the number of minutes of league football, m, played by each team in the four main divisions before first scoring a goal at the start of a new season. Her results are shown in the table below.

m (minutes)	Number of teams
$0 \leq m < 40$	36
$40 \leq m < 80$	28
$80 \leq m < 120$	10
$120 \leq m < 160$	4
$160 \leq m < 200$	5
$200 \leq m < 300$	4
$300 \leq m < 400$	2
$400 \leq m < 600$	3

(a) Calculate estimates of the mean and standard deviation of these data. **(8 marks)**

(b) Explain why the mean and standard deviation might not be the best summary statistics to use with these data.

(2 marks)

(c) Suggest alternative summary statistics that would better represent these data.

(1 mark)

---

4. Alan runs on a treadmill each day for as long as he can at 7 miles per hour. The length of time for which he runs is normally distributed with a mean of 21.6 minutes and a standard deviation of 1.8 minutes.

(a) Calculate the probability that on any one day Alan will run for less than 20 minutes.

(3 marks)

(b) Estimate the number of times in a ninety-day period that Alan will run for more than 24 minutes.

(4 marks)

(c) On a particular day Alan is still running after 22 minutes. Find the probability that he will stop running in the next 2 minutes.

(6 marks)

Turn over

© *Solomon Press*

5. In a survey unemployed people were asked how many months it had been, to the nearest month, since they were last employed on a full-time basis. The data collected is summarised in this stem and leaf diagram.

Number of months	(2 | 1 means 21 months)	Totals
0	1 1 2 2 4 4 4 6 7 7 9	(11)
1	0 2 3 5 5 6 8 9	( )
2	1 5 6 8	( )
3	0 7 9	( )
4	5	( )
5	2 7	(2)
6	3	(1)
7	0	(1)

(a) Write down the values needed to complete the totals column on the stem and leaf diagram.

(1 mark)

(b) State the mode of these data. **(1 mark)**

(c) Find the median and quartiles of these data. **(4 marks)**

Given that any values outside of the limits $Q_1 - 1.5(Q_3 - Q_1)$ and $Q_3 + 1.5(Q_3 - Q_1)$ are to be regarded as outliers,

(d) determine if there are any outliers in these data, **(3 marks)**

(e) draw a box plot representing these data on graph paper, **(3 marks)**

(f) describe the skewness of these data and suggest a reason for it. **(2 marks)**

© *Solomon Press*

6. In a game two spinners are used. The score on the first spinner is given by the random variable A, which has the following probability distribution:

a	1	2	3
$P(A = a)$	$\frac{1}{3}$	$\frac{1}{3}$	$\frac{1}{3}$

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

(b) Write down $E(A)$. **(1 mark)**

The score on the second spinner is given by the random variable B, which has the following probability distribution:

b	1	2	3
$P(B = b)$	$\frac{1}{2}$	$\frac{1}{4}$	$\frac{1}{4}$

(c) Find $E(B)$. **(2 marks)**

On each player's turn in the game, both spinners are used and the scores on the two spinners are added together. The total score on the two spinners is given by the random variable C.

(d) Show that $P(C = 2) = \frac{1}{6}$. **(3 marks)**

(e) Find the probability distribution of C. **(6 marks)**

(f) Show that $E(C) = E(A) + E(B)$. **(4 marks)**

END

© *Solomon Press*