GCE Examinations

Statistics
Module S1

Advanced Subsidiary / Advanced Level

Paper F

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. The weight in kilograms, w, of the 15 players in a rugby team was recorded and the results summarised as follows.

$$\Sigma w = 1145.3, \quad \Sigma w^2 = 88042.14.$$

(a) Calculate the mean and variance of the weight of the players. **(5 marks)**

Due to injury, one of the players who weighed 79.2 kg was replaced with another player who weighed 63.5 kg.

(b) Without further calculation state the effect of this change on the mean and variance of the weight of the players in the team. Explain your answers. **(4 marks)**

---

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

x	1	2	3	4	5
$P(X = x)$	a	b	$\frac{1}{4}$	$2a$	$\frac{1}{8}$

(a) Find an expression for b in terms of a. **(3 marks)**

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

Given that $E(X) = \frac{45}{16}$,

(c) find the values of a and b, **(4 marks)**

---

3. The time it takes girls aged 15 to complete an obstacle course is found to be normally distributed with a mean of 21.5 minutes and a standard deviation of 2.2 minutes.

(a) Find the probability that a randomly chosen 15 year-old girl completes the course in less than 25 minutes. **(3 marks)**

A 13 year-old girl completes the course in exactly 19 minutes.

(b) What percentage of 15 year-old girls would she beat over the course? **(3 marks)**

Anyone completing the course in less than 20 minutes is presented with a certificate of achievement. Three friends all complete the course one afternoon.

(c) What is the probability that exactly two of them get certificates? **(5 marks)**

© *Solomon Press*

4. The events A and B are such that

$$P(A) = 0.5,\ P(B) = 0.42 \text{ and } P(A \cup B) = 0.76$$

Find

(a) $P(A \cap B)$, **(3 marks)**

(b) $P(A' \cup B)$, **(3 marks)**

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

(d) Show that events A and B are not independent. **(3 marks)**

---

5. Each child in class 3A was given a packet of seeds to plant. The stem and leaf diagram below shows how many seedlings were visible in each child's tray one week after planting.

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

(a) Find the median and interquartile range for these data. **(5 marks)**

(b) Use the quartiles to describe the skewness of the data. Show your method clearly.

(3 marks)

The mean and standard deviation for these data were 27.2 and 10.3 respectively.

(c) Explaining your answer, state whether you would recommend using these values or your answers to part *(a)* to summarise these data.

(2 marks)

Outliers are defined to be values outside of the limits $Q_1 - 2s$ and $Q_3 + 2s$ where s is the standard deviation given above.

(d) Represent these data with a boxplot identifying clearly any outliers. **(6 marks)**

Turn over

© *Solomon Press*

6. A school introduced a new programme of support lessons in 1994 with a view to improving grades in GCSE English. The table below shows the number of years since 1994, n, and the corresponding percentage of students achieving A to C grades in GCSE English, p, for each year.

n	1	2	3	4	5	6
p (%)	35.2	37.1	40.6	39.0	43.4	44.8

(a) Represent these data on a scatter diagram. **(4 marks)**

You may use the following values.

$$\Sigma n = 21, \quad \Sigma p = 240.1, \quad \Sigma n^2 = 91, \quad \Sigma p^2 = 9675.41, \quad \Sigma np = 873.$$

(b) Find an equation of the regression line of p on n and draw it on your graph. **(9 marks)**

(c) Calculate the product moment correlation coefficient for these data and comment on the suitability of a linear model for the relationship between n and p during this period. **(4 marks)**

END

© *Solomon Press*