

SECTION A: STATISTICS

Answer ALL questions. Write your answers in the spaces provided.

1. The number of hours of sunshine each day, y , for the month of July at Heathrow are summarised in the table below.

Hours	$0 \leq y < 5$	$5 \leq y < 8$	$8 \leq y < 11$	$11 \leq y < 12$	$12 \leq y < 14$
Frequency	12	6	8	3	2

A histogram was drawn to represent these data. The $8 \leq y < 11$ group was represented by a bar of width 1.5 cm and height 8 cm.

- (a) Find the width and the height of the $0 \leq y < 5$ group. (3)

- (b) Use your calculator to estimate the mean and the standard deviation of the number of hours of sunshine each day, for the month of July at Heathrow.
Give your answers to 3 significant figures. (3)

The mean and standard deviation for the number of hours of daily sunshine for the same month in Hurn are 5.98 hours and 4.12 hours respectively.
Thomas believes that the further south you are the more consistent should be the number of hours of daily sunshine.

- (c) State, giving a reason, whether or not the calculations in part (b) support Thomas' belief. (2)

- (d) Estimate the number of days in July at Heathrow where the number of hours of sunshine is more than 1 standard deviation above the mean. (2)

Helen models the number of hours of sunshine each day, for the month of July at Heathrow by $N(6.6, 3.7^2)$.

- (e) Use Helen's model to predict the number of days in July at Heathrow when the number of hours of sunshine is more than 1 standard deviation above the mean. (2)

- (f) Use your answers to part (d) and part (e) to comment on the suitability of Helen's model. (1)

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

Question 1 continued

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

(Total for Question 1 is 13 marks)

2. A meteorologist believes that there is a relationship between the daily mean windspeed, w kn, and the daily mean temperature, t °C. A random sample of 9 consecutive days is taken from past records from a town in the UK in July and the relevant data is given in the table below.

t	13.3	16.2	15.7	16.6	16.3	16.4	19.3	17.1	13.2
w	7	11	8	11	13	8	15	10	11

The meteorologist calculated the product moment correlation coefficient for the 9 days and obtained $r = 0.609$

- (a) Explain why a linear regression model based on these data is unreliable on a day when the mean temperature is 24 °C (1)
- (b) State what is measured by the product moment correlation coefficient. (1)
- (c) Stating your hypotheses clearly test, at the 5% significance level, whether or not the product moment correlation coefficient for the population is greater than zero. (3)

Using the same 9 days a location from the large data set gave $\bar{t} = 27.2$ and $\bar{w} = 3.5$

- (d) Using your knowledge of the large data set, suggest, giving your reason, the location that gave rise to these statistics. (1)

DO NOT WRITE IN THIS AREA

Question 2 continued

Ruled area for writing the answer to Question 2, consisting of 25 horizontal lines.

(Total for Question 2 is 6 marks)

3. A machine cuts strips of metal to length L cm, where L is normally distributed with standard deviation 0.5 cm.

Strips with length either less than 49 cm or greater than 50.75 cm **cannot** be used.

Given that 2.5% of the cut lengths exceed 50.98 cm,

(a) find the probability that a randomly chosen strip of metal **can** be used.

(5)

Ten strips of metal are selected at random.

(b) Find the probability fewer than 4 of these strips **cannot** be used.

(2)

A second machine cuts strips of metal of length X cm, where X is normally distributed with standard deviation 0.6 cm

A random sample of 15 strips cut by this second machine was found to have a mean length of 50.4 cm

(c) Stating your hypotheses clearly and using a 1% level of significance, test whether or not the mean length of all the strips, cut by the second machine, is greater than 50.1 cm

(5)

4. Given that

$$P(A) = 0.35 \quad P(B) = 0.45 \quad \text{and} \quad P(A \cap B) = 0.13$$

find

(a) $P(A' | B')$ (2)

(b) Explain why the events A and B are not independent. (1)

The event C has $P(C) = 0.20$

The events A and C are mutually exclusive and the events B and C are statistically independent.

(c) Draw a Venn diagram to illustrate the events A , B and C , giving the probabilities for each region. (5)

(d) Find $P([B \cup C]')$ (2)

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

Question 4 continued

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

(Total for Question 4 is 10 marks)

5. A company sells seeds and claims that 55% of its pea seeds germinate.
- (a) Write down a reason why the company should not justify their claim by testing all the pea seeds they produce. (1)

A random selection of the pea seeds is planted in 10 trays with 24 seeds in each tray.

- (b) Assuming that the company's claim is correct, calculate the probability that in at least half of the trays 15 or more of the seeds germinate. (3)
- (c) Write down two conditions under which the normal distribution may be used as an approximation to the binomial distribution. (1)

A random sample of 240 pea seeds was planted and 150 of these seeds germinated.

- (d) Assuming that the company's claim is correct, use a normal approximation to find the probability that at least 150 pea seeds germinate. (3)
- (e) Using your answer to part (d), comment on whether or not the proportion of the company's pea seeds that germinate is different from the company's claim of 55% (1)

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

Question 5 continued

(Total for Question 5 is 9 marks)

TOTAL FOR SECTION A IS 50 MARKS
