

1. On a particular day the height above sea level, x metres, and the mid-day temperature, $y^\circ\text{C}$, were recorded in 8 north European towns. These data are summarised below

$$S_{xx} = 3\,535\,237.5 \quad \sum y = 181 \quad \sum y^2 = 4305 \quad S_{xy} = -23\,726.25$$

- (a) Find S_{yy} (2)

- (b) Calculate, to 3 significant figures, the product moment correlation coefficient for these data. (2)

- (c) Give an interpretation of your coefficient. (1)

A student thought that the calculations would be simpler if the height above sea level, h , was measured in kilometres and used the variable $h = \frac{x}{1000}$ instead of x .

- (d) Write down the value of S_{hh} (1)

- (e) Write down the value of the correlation coefficient between h and y . (1)



Leave blank

3. The discrete random variable Y has probability distribution

y	1	2	3	4
$P(Y=y)$	a	b	0.3	c

where a , b and c are constants.

The cumulative distribution function $F(y)$ of Y is given in the following table

y	1	2	3	4
$F(y)$	0.1	0.5	d	1.0

where d is a constant.

(a) Find the value of a , the value of b , the value of c and the value of d . **(5)**

(b) Find $P(3Y + 2 \geq 8)$. **(2)**



Leave blank

4. Past records show that the times, in seconds, taken to run 100 m by children at a school can be modelled by a normal distribution with a mean of 16.12 and a standard deviation of 1.60

A child from the school is selected at random.

- (a) Find the probability that this child runs 100 m in less than 15 s. **(3)**

On sports day the school awards certificates to the fastest 30% of the children in the 100 m race.

- (b) Estimate, to 2 decimal places, the slowest time taken to run 100 m for which a child will be awarded a certificate. **(4)**



Leave blank

5. A class of students had a sudoku competition. The time taken for each student to complete the sudoku was recorded to the nearest minute and the results are summarised in the table below.

Time	Mid-point, x	Frequency, f
2 - 8	5	2
9 - 12		7
13 - 15	14	5
16 - 18	17	8
19 - 22	20.5	4
23 - 30	26.5	4

(You may use $\sum fx^2 = 8603.75$)

- (a) Write down the mid-point for the 9 - 12 interval. (1)
- (b) Use linear interpolation to estimate the median time taken by the students. (2)
- (c) Estimate the mean and standard deviation of the times taken by the students. (5)

The teacher suggested that a normal distribution could be used to model the times taken by the students to complete the sudoku.

- (d) Give a reason to support the use of a normal distribution in this case. (1)

On another occasion the teacher calculated the quartiles for the times taken by the students to complete a different sudoku and found

$$Q_1 = 8.5 \quad Q_2 = 13.0 \quad Q_3 = 21.0$$

- (e) Describe, giving a reason, the skewness of the times on this occasion. (2)



Leave blank

- 6. Jake and Kamil are sometimes late for school.
The events J and K are defined as follows

J = the event that Jake is late for school
 K = the event that Kamil is late for school

$P(J) = 0.25$, $P(J \cap K) = 0.15$ and $P(J' \cap K') = 0.7$

On a randomly selected day, find the probability that

- (a) at least one of Jake or Kamil are late for school, **(1)**

- (b) Kamil is late for school. **(2)**

Given that Jake is late for school,

- (c) find the probability that Kamil is late. **(3)**

The teacher suspects that Jake being late for school and Kamil being late for school are linked in some way.

- (d) Determine whether or not J and K are statistically independent. **(2)**

- (e) Comment on the teacher's suspicion in the light of your calculation in (d). **(1)**



Leave
blank

Question 6 continued

(This area contains horizontal lines for writing.)



Leave blank

7. A teacher took a random sample of 8 children from a class. For each child the teacher recorded the length of their left foot, f cm, and their height, h cm. The results are given in the table below.

f	23	26	23	22	27	24	20	21
h	135	144	134	136	140	134	130	132

(You may use $\sum f = 186$ $\sum h = 1085$ $S_{ff} = 39.5$ $S_{hh} = 139.875$ $\sum fh = 25291$)

(a) Calculate S_{fh} **(2)**

(b) Find the equation of the regression line of h on f in the form $h = a + bf$.
Give the value of a and the value of b correct to 3 significant figures. **(5)**

(c) Use your equation to estimate the height of a child with a left foot length of 25 cm. **(2)**

(d) Comment on the reliability of your estimate in (c), giving a reason for your answer. **(2)**

The left foot length of the teacher is 25 cm.

(e) Give a reason why the equation in (b) should not be used to estimate the teacher's height. **(1)**



Leave
blank

Question 7 continued

Lined area for writing the answer to Question 7 continued.



Leave blank

8. A spinner is designed so that the score S is given by the following probability distribution.

s	0	1	2	4	5
$P(S = s)$	p	0.25	0.25	0.20	0.20

- (a) Find the value of p . (2)
- (b) Find $E(S)$. (2)
- (c) Show that $E(S^2) = 9.45$ (2)
- (d) Find $\text{Var}(S)$. (2)

Tom and Jess play a game with this spinner. The spinner is spun repeatedly and S counters are awarded on the outcome of each spin. If S is even then Tom receives the counters and if S is odd then Jess receives them. The first player to collect 10 or more counters is the winner.

- (e) Find the probability that Jess wins after 2 spins. (2)
- (f) Find the probability that Tom wins after exactly 3 spins. (4)
- (g) Find the probability that Jess wins after exactly 3 spins. (3)



Leave blank

Question 8 continued

Ruled area for writing the answer to Question 8.

Q8

(Total 17 marks)

TOTAL FOR PAPER: 75 MARKS

END

