Mathematics

Edexcel IAL

S1

Worksheets

Measures of Location and Spread

Eng. Nagy Elraheb

Mathematical Modelling:

Exercise:

1   Briefly explain the role of statistical tests in the process of mathematical modelling.

2   Describe how to refine the process of designing a mathematical model.

3   It is generally accepted that there are seven stages involved in creating a mathematical model. They are summarised below. Write down the missing stages.

Stage 1:

Stage 2: A mathematical model is devised

Stage 3: Model used to make predictions

Stage 4:

Stage 5: Comparisons are made against the devised model

Stage 6:

Stage 7: Model is refined

Measures of Location and Spread

Exercise 1:

1 State whether each of the following variables is qualitative or quantitative:

  a The height of a building

  b The colour of a jumper

  c Time spent waiting in a queue

  d Shoe size

  e Names of students in a school

2 State which of the following statements are true:

  a The weight of apples is discrete data.

  b The number of apples on the trees in an orchard is discrete data.

  c The amount of time it takes a train to make a journey is continuous data.

  d Simhal collected data on car colours by standing at the end of her road and writing down the car colours. The data she collected is quantitative.

3 The distribution of the lifetimes of torch batteries are shown in the grouped frequency table below.

  a Write down the class boundaries for the second group.

  b Work out the midpoint of the fifth group.

Lifetime (Nearest 0.1 of an hour)	Frequency
5.0–5.9	5
6.0–6.9	8
7.0–7.9	10
8.0–8.9	22
9.0–9.9	10
10.0–10.9	2

4 The grouped frequency table below shows the distributions of the weights of 16-week-old kittens.

  a Write down the class boundaries for the third group.

  b Work out the midpoint of the second group.

Weight (kg)	Frequency
1.2–1.3	8
1.3–1.4	28
1.4–1.5	32
1.5–1.6	22

Exercise 2:

1 Priyanka collected wild mushrooms every day for a week. When she got home each day she weighed them to the nearest 100 g. The weights are shown below:

    500   700   400   300   900   700   700

  a Write down the mode for these data.

  b Calculate the mean for these data.

  c Find the median for these data.

On the next day, Priyanka collected 650 g of wild mushrooms.

  d Write down the effect this will have
    on the mean, the mode and the median.

> **Hint** Try to answer part **d** without recalculating the averages. You could recalculate to check your answer.

2 Taha collects six pieces of data, x_1, x_2, x_3, x_4, x_5 and x_6. He works out that Σx is 256.2

  a Calculate the mean for these data.

Taha collects another piece of data. It is 52.

  b Write down the effect this piece of data will have on the mean.

3 The daily mean visibility, v metres, for Kuala Lumpur in May and June was recorded each day. The data are summarised as follows:

    May: $n = 31$, $\Sigma v = 724\,000$

    June: $n = 30$, $\Sigma v = 632\,000$

  a Calculate the mean visibility in each month.

  b Calculate the mean visibility for the total recording period.

4 A small workshop records how long it takes, in minutes, for each of their workers to make a certain item. The times are shown in the table.

Worker	A	B	C	D	E	F	G	H	I	J
Time in minutes	7	12	10	8	6	8	5	26	11	9

  a Write down the mode for these data.

  b Calculate the mean for these data.

  c Find the median for these data.

  d The manager wants to give the workers an idea of the average time they took. Write down, with a reason, which of the answers to **a**, **b** and **c** she should use.

5 The frequency table shows the number of breakdowns, b, per month recorded by a lorry firm over a certain period of time.

Breakdowns	0	1	2	3	4	5
Frequency	8	11	12	3	1	1

  a Write down the modal number of breakdowns.

  b Find the median number of breakdowns.

  c Calculate the mean number of breakdowns.

  d In a brochure about how many loads reach their destination on time, the firm quotes one of the answers to **a**, **b** or **c** as the number of breakdowns per month for its vehicles. Write down which of the three answers the firm should quote in the brochure.

6 The table shows the frequency distribution for the number of petals in the flowers of a group of celandines.

Number of petals	5	6	7	8	9
Frequency	8	57	29	3	1

Calculate the mean number of petals.

7 A scientist is investigating how many eggs the endangered kakapo bird lays in each brood cycle. The results are given in this frequency table.

Number of eggs	1	2	3
Frequency	7	p	2

Problem-solving

Use the formula for the mean of an ungrouped frequency table to write an equation involving p.

If the mean number of eggs is 1.5, find the value of p.

Exercise 3:

1 The weekly wages (to the nearest €) of the production line workers in a small factory are shown in the table.

a Write down the modal class.

b Calculate an estimate of the mean wage.

c Write down the interval containing the median.

Weekly wage (€)	Frequency
175–225	4
226–300	8
301–350	18
351–400	28
401–500	7

2 The noise levels at 30 locations near an outdoor concert venue were measured to the nearest decibel. The data collected are shown in the grouped frequency table.

Noise (decibels)	65–69	70–74	75–79	80–84	85–89	90–94	95–99
Frequency	1	4	6	6	8	4	1

a Calculate an estimate of the mean noise level. (1 mark)

b Explain why your answer to part **a** is an estimate. (1 mark)

3 The table shows the daily mean temperatures in Addis Ababa for the 30 days of June one year.

Temperature (°C)	$8 \leqslant t < 10$	$10 \leqslant t < 12$	$12 \leqslant t < 14$	$14 \leqslant t < 16$	$16 \leqslant t < 18$	$18 \leqslant t < 20$	$20 \leqslant t < 22$
Frequency	1	2	4	4	10	4	5

a Write down the modal class. (1 mark)

b Calculate an estimate for the mean daily mean temperature. (1 mark)

4 Two shops (A and B) recorded the ages of their workers.

Age of worker	16–25	26–35	36–45	46–55	56–65	66–75
Frequency A	5	16	14	22	26	14
Frequency B	4	12	10	28	25	13

By comparing estimated means for each shop, determine which shop is better at employing older workers.

Exercise 4:

1 The daily mean pressure (hPa) during the last 16 days of July in Perth is recorded. The data are given below:

> 1024 1022 1021 1013 1009 1018 1017 1024
>
> 1027 1029 1031 1025 1017 1019 1017 1014

Notation hPa (hectopascal) is the SI unit used to measure atmospheric pressure in weather and meteorology.

a Find the median pressure for that period.

b Find the lower and upper quartiles.

2 Zaynep records the number of books in the collections of students in her year. The results are in the table below.

Number of books	35	36	37	38	39
Frequency	3	17	29	34	12

Find Q_1, Q_2 and Q_3.

Hint This is an ungrouped frequency table so you do not need to use interpolation. Use the rules for finding the median and quartiles of **discrete** data.

3 A hotel is worried about the reliability of its lift. It keeps a weekly record of the number of times it breaks down over a period of 26 weeks. The data collected are summarised in the table opposite.

Use interpolation to estimate the median number of breakdowns.

Number of breakdowns	Frequency
0–1	18
2–3	7
4–5	1

(2 marks)

4 The weights of 31 cows were recorded to the nearest kilogram. The weights are shown in the table.

a Find an estimate for the median weight.

b Find the lower quartile, Q_1.

c Find the upper quartile, Q_3.

Weight of cow (kg)	300–349	350–399	400–449	450–499	500–549
Frequency	3	6	10	7	5

d Interpret the meaning of the value you have found for the upper quartile in part **c**.

5 A roadside assistance company kept a record over a week of the amount of time, in minutes, people were kept waiting for assistance. The times are shown below.

Time waiting, t (minutes)	$20 \leqslant t < 30$	$30 \leqslant t < 40$	$40 \leqslant t < 50$	$50 \leqslant t < 60$	$60 \leqslant t < 70$
Frequency	6	10	18	13	2

   a Find an estimate for the mean time spent waiting.   **(1 mark)**

   b Calculate the 65th percentile.   **(2 marks)**

The firm writes the following statement for an advertisement:

> Only 10% of our customers have to wait longer than 56 minutes.

   c By calculating a suitable percentile, comment on the validity of this claim.   **(3 marks)**

6 The table shows the recorded wingspans, in metres, of 100 endangered California condor birds.

Wingspan, w (m)	$1.0 \leqslant w < 1.5$	$1.5 \leqslant w < 2.0$	$2.0 \leqslant w < 2.5$	$2.5 \leqslant w < 3.0$	$3.0 \leqslant w$
Frequency	4	20	37	28	11

   a Estimate the 80th percentile and interpret the value.   **(3 marks)**

   b State why it is not possible to estimate the 90th percentile.   **(1 mark)**

Exercise 5:

1 The lengths of a number of slow worms were measured, to the nearest mm. The results are shown in the table.

Length of slow worms (mm)	Frequency
125–139	4
140–154	4
155–169	2
170–184	7
185–199	20
200–214	24
215–229	10

   a Work out how many slow worms were measured.

   b Estimate the interquartile range for the lengths of the slow worms.

   c Calculate an estimate for the mean length of the slow worms.

   d Estimate the number of slow worms whose length is more than one interquartile range above the mean.

Problem-solving

For part **d**, work out $\bar{x}$ + IQR, and determine which class interval it falls in. Then use proportion to work out how many slow worms from that class interval you need to include in your estimate.

2 The table shows the monthly income for workers in a factory.

Monthly income, x (\$)	$900 \leqslant x < 1000$	$1000 \leqslant x < 1100$	$1100 \leqslant x < 1200$	$1200 \leqslant x < 1300$
Frequency	3	24	28	15

    a Calculate the 34% to 66% interpercentile range.   **(3 marks)**

    b Estimate the number of data values that fall within this range.   **(2 marks)**

3 A train travelled from Manchester to Liverpool. The times, to the nearest minute, it took for the journey were recorded over a certain period. The times are shown in the table.

Journey time (minutes)	15–16	17–18	19–20	21–22
Frequency	5	10	35	10

    a Calculate the 5% to 95% interpercentile range.   **(3 marks)**

    b Estimate the number of data values that fall within this range.   **(1 mark)**

4 The daily mean temperature (°C) in Santiago for each of the first ten days of June is given below:

      14.3   12.7   12.4   10.9   9.4   13.2   12.1   10.3   10.3   10.6

    a Calculate the median and interquartile range.   **(2 marks)**

    The median daily mean temperature in Santiago during the first 10 days of May was 9.9 °C and the interquartile range was 3.9 °C.

    b Compare the data for May with the data for June.   **(2 marks)**

    The 10% to 90% interpercentile range for the daily mean temperature in Santiago during July was 5.4 °C.

    c Estimate the number of days in July on which the daily mean temperature fell within this range.   **(1 mark)**

Exercise 6:

1 The summary data for a variable x are: $\Sigma x = 24$ $\Sigma x^2 = 78$ $n = 8$

    Find:

    a the mean

    b the variance σ^2

    c the standard deviation σ.

2 Ten collie dogs are weighed (w kg). The summary data for the weights are:

$\Sigma w = 241$ $\Sigma w^2 = 5905$

Use this summary data to find the standard deviation of the collies' weights.   **(2 marks)**

3 Eight students' heights (h cm) are measured. They are as follows:

165   170   190   180   175   185   176   184

a Work out the mean height of the students.

b Given $\Sigma h^2 = 254\,307$, work out the variance. Show all your working.

c Work out the standard deviation.

4 For a set of 10 numbers: $\Sigma x = 50$ $\Sigma x^2 = 310$

For a different set of 15 numbers: $\Sigma x = 86$ $\Sigma x^2 = 568$

Find the mean and the standard deviation of the combined set of 25 numbers.

5 Nahab asks the students in his year group how much allowance they get per week. The results, rounded to the nearest Omani Riyals, are shown in the table.

Number of OMR	8	9	10	11	12
Frequency	14	8	28	15	20

a Work out the mean and standard deviation of the allowance. Give units with your answer.   **(3 marks)**

b How many students received an allowance amount more than one standard deviation above the mean?   **(2 marks)**

6 In a student group, a record was kept of the number of days of absence each student had over one particular term. The results are shown in the table.

Number of days absent	0	1	2	3	4
Frequency	12	20	10	7	5

Work out the standard deviation of the number of days absent.   **(2 marks)**

7 A certain type of machine contained a part that tended to wear out after different amounts of time. The time it took for 50 of the parts to wear out was recorded. The results are shown in the table.

Lifetime, h (hours)	$5 < h \leqslant 10$	$10 < h \leqslant 15$	$15 < h \leqslant 20$	$20 < h \leqslant 25$	$25 < h \leqslant 30$
Frequency	5	14	23	6	2

The manufacturer makes the following claim:

90% of the parts tested lasted longer than one standard deviation below the mean.

You need to calculate estimates for the mean and the standard deviation, then estimate the number of parts that lasted longer than one standard deviation below the mean.

Comment on the accuracy of the manufacturer's claim, giving relevant numerical evidence.

(5 marks)

8 The daily mean wind speed, x (knots) in Chicago is recorded. The summary data are:

$$\Sigma x = 243 \qquad \Sigma x^2 = 2317$$

a Work out the mean and the standard deviation of the daily mean wind speed. **(2 marks)**

The highest recorded wind speed was 17 knots and the lowest recorded wind speed was 4 knots.

b Estimate the number of days in which the wind speed was greater than one standard deviation above the mean. **(2 marks)**

c State one assumption you have made in making this estimate. **(1 mark)**

Challenge

The manager at a local bakery calculates the mean and standard deviation of the number of loaves of bread bought per person in a random sample of her customers as 0.787 and 0.99 respectively. If each loaf costs $1.04, calculate the mean and standard deviation of the amount spent on loaves per person.

Exercise 7

1 A set of data values, x, is shown below:

| 110 | 90 | 50 | 80 | 30 | 70 | 60 |

a Code the data using the coding $y = \dfrac{x}{10}$

b Calculate the mean of the coded data values.

c Use your answer to part **b** to calculate the mean of the original data.

2 A set of data values, x, is shown below:

| 52 | 73 | 31 | 73 | 38 | 80 | 17 | 24 |

a Code the data using the coding $y = \dfrac{x - 3}{7}$

b Calculate the mean of the coded data values.

c Use your answer to part **b** to calculate the mean of the original data.

3 The coded mean price of televisions in a shop was worked out. Using the coding $y = \dfrac{x - 65}{200}$

the mean price was 1.5. Find the true mean price of the televisions.   **(2 marks)**

4 The coding $y = x - 40$ gives a standard deviation for y of 2.34

Write down the standard deviation of x.

> **Watch out** Adding or subtracting constants does not affect how spread out the data are, so you can ignore the '−40' when finding the standard deviation for x.

5 A study was performed to investigate how long a mobile phone battery lasts if the phone is not used. The grouped frequency table shows the battery life (b hours) of a random sample of 100 different mobile phones.

Battery life (b hours)	Frequency (f)	Midpoint (x)	$y = \dfrac{x - 14}{2}$
11–21	11		
21–27	24		
27–31	27		
31–37	26		
37–43	12		

a Copy and complete the table.

b Use the coding $y = \dfrac{x - 14}{2}$ to calculate an estimate of the mean battery life.

6 The lifetime, x, in hours, of 70 light bulbs is shown in the table below.

Lifetime, x (hours)	$20 < x \leqslant 22$	$22 < x \leqslant 24$	$24 < x \leqslant 26$	$26 < x \leqslant 28$	$28 < x \leqslant 30$
Frequency	3	12	40	10	5

The data are coded using $y = \dfrac{x - 1}{20}$

a Estimate the mean of the coded values $\bar{y}$.

b Hence find an estimate for the mean lifetime of the light bulbs, $\bar{x}$.

c Estimate the standard deviation of the lifetimes of the light bulbs.

> **Problem-solving**
>
> Code the midpoints of each class interval. The midpoint of the $22 < x \leqslant 24$ class interval is 23, so the coded midpoint will be $\dfrac{23 - 1}{20} = 1.1$

7 The weekly income, i, of 100 workers was recorded.

The data were coded using $y = \dfrac{i - 90}{100}$ and the following summations were obtained:

$$\Sigma y = 131, \ \Sigma y^2 = 176.84$$

Estimate the standard deviation of the actual workers' weekly income.   **(2 marks)**

8 A meteorologist collected data on the annual rainfall, x mm, at six randomly selected places.

The data were coded using $s = 0.01x - 10$ and the following summations were obtained:

$$\Sigma s = 16.1, \ \Sigma s^2 = 147.03$$

Work out an estimate for the standard deviation of the actual annual rainfall.   **(2 marks)**

9 The daily mean pressure, p hPa, in Accra during August is recorded.

The data are coded using $c = \dfrac{p}{2} - 500$ and the following

summary statistics were obtained:

$$n = 30 \qquad \bar{c} = 10.15 \qquad S_{cc} = 296.4$$

Find the mean and standard deviation of the daily mean pressure.   **(4 marks)**