S1·3

Measures of variation

CHAPTER ASSIGNMENT · 199 MARKS

Name _____ Class _____ No. _____ Date _____

This review exercise covers material across Chapters 1 to 3: Representation of data, Measures of central tendency, and Measures of variation.

Practice questions from *Cambridge International AS & A Level Mathematics: Probability & Statistics 1 Coursebook* (Cambridge University Press, Dean Chalmers), Chapter 3 (Measures of variation), Exercises 3A to 3E, End-of-chapter review exercise 3 and Cross-topic review exercise 1, for classroom use. The Check yourself and Chapter quiz questions are Guru Jazzy's own.

Exercise 3A · The range, quartiles and percentiles

3A.1 Find the range and the interquartile range of the following sets of data:

(a) 5, 8, 13, 17, 22, 25, 30

(b) 7, 13, 21, 2, 37, 28, 17, 11, 2

(c) 42, 47, 39, 51, 73, 18, 83, 29, 41, 64

(d) 113, 97, 36, 81, 49, 41, 20, 66, 28, 32, 17, 107

(e) 4.6, 0, −2.6, 0.8, −1.9, −3.3, 5.2, −3.2

3A.2 A dataset is represented in a box-and-whisker plot on a scale from 0 to 4 with grid markings every 0.1 units: the lowest whisker is at 0.4, the lower quartile Q1 is at 1.4, the median Q2 is at 2.4, the upper quartile Q3 is at 3.15, and the highest whisker is at 3.7.

(a) Find the range and the interquartile range of the dataset.

(b) What type of skewness would you expect this set of data to have?

3A.3 The following stem-and-leaf diagram shows the marks out of 50 obtained by 15 students in a Science test. Key: 2 | 5 represents a mark of 25 out of 50.

```
Stem | Leaf
   0 | 9
   1 | 4
   2 | 5  8
   3 | 0  1  1  7  9
   4 | 3  5  6  8
   5 | 0  0
```

(a) Find the range and interquartile range of the marks.

(b) Illustrate the data in a box-and-whisker diagram on graph paper and include a scale.

(c) For this set of data, express Q3 in terms of Q1 and Q2.

3A.4 The numbers of fouls made in eight hockey matches and in eight football matches played at the weekend are shown in a back-to-back stem-and-leaf diagram. Key: 6 | 1 | 8 represents 16 fouls in hockey and 18 fouls in football.

Hockey (8): 11, 12, 14, 16, 18, 19, 21, 24.

Football (8): 10, 18, 19, 19, 21, 22, 23, 23.

(a) Is it true to say that the numbers of fouls in the two sports are equally varied? Explain your answer.

(b) Draw two box-and-whisker diagrams using the same scales. Write a sentence to compare the numbers of fouls committed in the two sports.

3A.5 Rishi and Daisy take the same seven tests in Mathematics, and both students' marks improve on successive tests. Their percentage marks are:

Rishi: 15, 24, 28, 33, 39, 42, 50

Daisy: 51, 65, 69, 72, 78, 83, 86

(a) Explain why it would not be useful to use the range or the interquartile range alone as measures for comparing the marks of the two students.

(b) Name two measures that could be used together to give a meaningful comparison of the two students' marks.

3A.6 The following table shows the maximum speeds, s km/h, of some vintage cars:

Maximum speed s (km/h)	No. vintage cars f
s < 35	0
35 <= s < 40	20
40 <= s < 45	65
45 <= s < 50	110
50 <= s < 55	27
55 <= s < 70	13
70 <= s < 75	5

(a) On the same sheet of graph paper, construct a cumulative frequency polygon and a box-and-whisker diagram to illustrate the data.

(b) Use your box-and-whisker diagram to assess what type of skewness the data have.

3A.7 Twenty adults are selected at random, and each is asked to state the number of trips abroad that they have made. The results are shown in a back-to-back stem-and-leaf diagram (Key: 3 | 1 | 1 represents 13 trips for a male and 11 trips for a female).

Males (9): 0, 0, 0, 2, 3, 8, 13, 14, 15 trips.

Females (11): 3, 4, 5, 11, 11, 12, 13, 16, 20, 21, 22 trips.

(a) (i) Draw box-and-whisker diagrams, using the same scales for males and for females. (ii) Interpret the key features of the data represented in your diagrams and compare the data for the two groups of adults.

(b) In a summary of the data, a student writes, 'The females have visited more countries than the males.' Is this statement justified? Give a reason to support your answer.

3A.8 The resistances, in ohms (Ω), of 100 conductors are represented in a cumulative frequency curve passing through (0, 0), (0.1, 8), (0.15, 20), (0.2, 45), (0.25, 75), (0.3, 88), (0.35, 95), (0.4, 98), (0.5, 100).

Find, to an appropriate degree of accuracy, an estimate of:

(a) the interquartile range

(b) the 90th percentile

(c) the percentile that is equal to 0.192 Ω

(d) the range of the middle 40% of the resistances.

3A.9 The areas, in cm^2, of 200 circuit boards are represented in a cumulative frequency graph from 0 to 56 cm^2, passing through (8, 0), (16, 20), (24, 60), (32, 120), (40, 170), (48, 192), (56, 200).

(a) State the greatest possible range of the data.

(b) Construct a box-and-whisker diagram to illustrate the data.

(c) Find the range of the middle 60% of the areas.

(d) An outlier is an extreme value that is more than 1.5 times the interquartile range above the upper quartile or more than 1.5 times the interquartile range below the lower quartile. Find the areas that define the outliers in this set of data, and estimate how many there are. How accurate is your answer?

3A.10 A company manufactures right-angled brackets for use in the construction industry. A sample of brackets are measured, and the number of degrees by which their angles deviate from a right angle are summarised:

Deviation from 90° d	No. brackets f
d < -1.5	0
-1.5 <= d < -1.0	24
-1.0 <= d < -0.5	46
-0.5 <= d < 0.0	61
0.0 <= d < 0.5	34
0.5 <= d < 1.0	34
1.0 <= d < 1.5	20
1.5 <= d < 2.5	17

(a) Draw a cumulative frequency polygon to illustrate these deviations.

(b) Estimate the median and the interquartile range of the bracket angles, giving both answers correct to 1 decimal place.

(c) A bracket is considered unsuitable for use if its angle deviates from a right angle by more than 1.2°. Estimate what percentage of this sample is unsuitable for use, giving your answer correct to the nearest integer.

3A.11 The following table shows the cumulative frequencies for values of x:

x	< 0	< 10	< 15	< 25	< 30	< 40
cf	0	12	30	90	102	120

Without drawing a cumulative frequency graph, find:

(a) the interquartile range

(b) the 85th percentile.

3A.12 Fifty 10-gram samples of a particular type of mushroom are collected by volunteers at a university and tested. The following table shows the mass of toxins, in hundredths of a gram, in these samples:

Mass (/0.01 g)	0-	4-	11-	17-	20-30
No. 10g samples	2	19	23	3	3

(a) Draw a cumulative frequency curve to illustrate the data.

(b) Use your curve to estimate, correct to 2 decimal places: (i) the interquartile range (ii) the range of the middle 80%.

(c) It was found that toxins made up between 0.75% and 2.25% of the mass of n of these samples. Use your curve to estimate the value of n.

(d) Make an assessment of the variation in the percentage of toxic material in these samples. Can you suggest any possible reasons for such variation?

3A.13 A 9-year study was carried out on the pollutants released when biomass fuels are used for cooking. Researchers offered nearly 1000 people living in 12 villages in southern China access to clean biogas and to improved kitchen ventilation. Some people took advantage of neither; some changed to clean fuels; some improved their kitchen ventilation; and some did both. Data on the concentrations of nitrogen dioxide (in mg/m^3) in these people's homes at the end of the study are summarized:

- Neither: min 0.10, Q1 0.40, median 0.58, Q3 0.78, max 0.95
- Clean fuels only: min 0.08, Q1 0.32, median 0.46, Q3 0.62, max 0.80
- Ventilation only: min 0.06, Q1 0.28, median 0.42, Q3 0.56, max 0.74
- Both: min 0.04, Q1 0.16, median 0.25, Q3 0.36, max 0.52

Study the data represented and write a brief analysis that summarises the results of this part of the study.

Exercise 3B · Variance and standard deviation

3B.1 Find the mean and the standard deviation for these sets of numbers:

(a) 27, 43, 29, 34, 53, 37, 19 and 58.

(b) −6.2, −8.5, 7.7, −4.3, 13.5 and −11.9.

3B.2 Last term Abraham sat three tests in each of his science subjects. His raw percentage marks for the tests, in the order they were completed, are listed:

Biology: 21, 33, 45

Chemistry: 41, 53, 65

Physics: 51, 63, 75

(a) Calculate the variance of Abraham's marks in each of the three subjects.

(b) Comment on the three values obtained in part a. Do the same comments apply to Abraham's mean mark for the tests in the three subjects? Justify your answer.

3B.3 The following table shows the number of pets owned by each of 35 families:

No. pets	0	1	2	3	4	5
No. families f	6	12	9	4	3	1

Find the mean and variance of the number of pets.

3B.4 The numbers of cobs produced by 360 maize plants are shown in the following table:

No. cobs	0	1	2	3	4
No. plants f	11	75	185	81	8

(a) Calculate the mean and the standard deviation.

(b) Find the interquartile range and give an example of what it tells us about this dataset that the standard deviation does not tell us.

3B.5 The times spent, in minutes, by 30 girls and by 40 boys on an assignment are detailed in the following table:

Time spent (min)	20-	30-	40-	60-80
No. girls f	6	14	7	3
No. boys f	15	11	7	7

(a) For the boys and for the girls, calculate estimates of the mean and standard deviation.

(b) It is required to make a comparison between the times spent by the two groups:

(i) What do the means tell us about the times spent?

(ii) Use the standard deviations to compare the times spent by the two groups.

3B.6 The lengths, correct to the nearest centimetre, of 50 rods are given in the following table:

Length (cm)	15-17	18-24	25-29	30-37
No. rods f	13	18	11	8

Calculate an estimate of the standard deviation of the lengths.

3B.7 For the dataset denoted by x in the following table, k is a constant:

x	15	16	17	18	19	20
f	2k	k + 5	k - 3	10	8	3

Find the value of k and calculate the variance of x, given that $\bar{x} = 17$.

3B.8 The following table illustrates the heights, in centimetres, of 150 children:

Height (cm)	No. children f
140 up to 144	a
144 up to 150	b
150 up to 160	69
160 up to 165	28

(a) Given that a calculated estimate of the mean height is exactly 153.14 cm, show that $142a + 147b = 7726$, and evaluate a and b.

(b) Calculate an estimate of the standard deviation of the heights.

3B.9 Kristina plans to raise money for charity. Her plan is to walk 217 km in 7 days so that she walks $k + 2n - n^2$ km on the nth day. Find the standard deviation of the daily distances she plans to walk, and compare this with the interquartile range.

3B.10 The mass of waste produced by a school during its three 13-week terms is given in tonnes, correct to 2 decimal places, in the following table:

Mass of waste (tonnes)	0.15-0.29	0.30-0.86	0.87-1.35	1.36-2.00
No. weeks f	5	8	20	6

(a) Calculate estimates of the mean and standard deviation of the mass of waste produced per week, giving both answers correct to 2 decimal places.

(b) No waste is produced in the 13 weeks of the year that the school is closed. If this additional data is included in the calculations, what effect does it have on the mean and on the standard deviation?

3B.11 The ages, in whole numbers of years, of a hotel's 50 staff are given in the following table. Calculated estimates of the mean and variance are 37.32 and 69.1176, respectively.

Age (years)	23-30	31-37	38-45	46-59
No. staff f	14	x	y	6

Exactly 1 year after these calculations were made, Gudrun became the 51st staff member and the mean age became exactly 38 years. Find Gudrun's age on the day of her recruitment, and determine what effect this had on the variance of the staff's ages. What assumptions must be made to justify your answers?

3B.12 In position 1, a 10-metre rod is placed 10 metres from a fixed point, P. Six small discs, A to F, are evenly spaced along the length of the rod. The rod is rotated anti-clockwise about its centre by $\alpha = 30°$ to position 2. The distances from P to the discs are denoted by x.

(a) What effect does the 30° rotation have on values of x? Investigate this by first considering the effect on the average distance from P to the discs.

(b) Find two values that can be used as measures of the change in the variation of x caused by the rotation.

(c) Use the values obtained in parts a and b to summarise the changes in the distances from P to the discs caused by the rotation.

(d) Can you prove that Σx^2 is constant for all values of α? (Hint: to do this, you need only to show that Σx^2 is constant for $0° < \alpha < 90°$).

Exercise 3C · Calculating from totals

3C.1 Given that:

(a) $\Sigma v^2 = 5480$, $\Sigma v = 288$ and $n = 64$, find the variance of v.

(b) $\Sigma w^2 = 4000$, $\bar{w} = 5.2$ and $n = 36$, find the standard deviation of w.

(c) $\Sigma x^2 f = 6120$, $\Sigma f = 40$ and the standard deviation of x is 12, find Σxf.

(d) $\Sigma xf = 2800$, $\Sigma f = 50$ and the variance of x is 100, find $\Sigma x^2 f$.

(e) $\Sigma t^2 = 193\,144$, $\Sigma t = 2324$ and that the standard deviation of t is 3, find the number of data values of t.



3C.2 A building is occupied by n companies. The number of people employed by these companies is denoted by x. Find the mean number of employees, given that $\Sigma x^2 = 8900$, $\Sigma x = 220$ and that the standard deviation of x is 18.

3C.3 Twenty-five values of p are such that $\Sigma p^2 = 6006$ and $\Sigma p = 388$, and 25 values of q are such that $\Sigma q^2 = 6114$ and $\Sigma q = 387$. Calculate the variance of the 50 values of p and q together.

3C.4 In a class of 30 students, the mean mass of the 14 boys is 63.5 kg and the mean mass of the girls is 57.3 kg. Calculate the mean and standard deviation of the masses of all the students together, given that the sums of the squares of the masses of the boys and girls are 58 444 kg^2 and 56 222 kg^2, respectively.

3C.5 The following table shows the front tyre pressure, in psi, of five 4-wheeled vehicles, A to E:

Vehicle	A	B	C	D	E
Front-left tyre	26	29	30	34	26
Front-right tyre	24	27	31	30	28

(a) Show that the variance of the pressure in all of these front tyres is 7.65 psi^2.

(b) Rear tyre pressures for these five vehicles are denoted by x. Given that $\Sigma x^2 = 7946$ and that the variance of the pressures in all of the front and rear tyres on these five vehicles together is 31.6275 psi^2, find the mean pressure in all the rear tyres.

3C.6 The totals $\Sigma x^2 = 7931$, $\Sigma x = 397$ and $\Sigma y = 499$ are given by 29 values of x and n values of y. All the values of x and y together have a variance of 52.

(a) Express Σy^2 in terms of n.

(b) Find the value of n for which $\Sigma y^2 - \Sigma x^2 = 10$.

3C.7 The five values in a dataset have a sum of 250 and standard deviation of 15. A sixth value is added to the dataset, such that the mean is now 40. Find the variance of the six values in the dataset.

3C.8 A group of 10 friends played a mini-golf competition. Eight of the friends tied for second place, each with a score of 34, and the other two friends tied for first place. Find the winning score, given that the standard deviation of the scores of all 10 friends was 1.2 and that the lowest score in golf wins.

3C.9 An author has written 15 children's books. The first eight books that she wrote contained between 240 and 250 pages each. The next six books contained between 180 and 190 pages each. Correct to 1 decimal place, the standard deviation of the number of pages in the 15 books together is 31.2. Show that it is not possible to determine a specific calculated estimate of the number of pages in the author's 15th book.

3C.10 A set of n pieces of data has mean $\bar{x}$ and standard deviation S. Another set of 2n pieces of data has mean $\bar{x}$ and standard deviation (1/2)S. Find the standard deviation of all these pieces of data together in terms of S.

Exercise 3D · Coded data

3D.1 Two years ago, the standard deviation of the masses of a group of men and a group of women were 8 kg and 6 kg, respectively. Today, all the men are 5 kg heavier and all the women are 3 kg lighter. Find the standard deviation for each group today.

3D.2 Twenty readings of y are summarised by the totals $\Sigma(y - 5)^2 = 890$ and $\Sigma(y - 5) = 130$. Find the standard deviation of y.

3D.3 The amounts of rainfall, r mm, at a certain location were recorded on 365 consecutive days and are summarised by $\Sigma(r - 3)^2 = 9950$ and $\Sigma(r - 3) = 1795.8$. Calculate the mean daily rainfall and the value of Σr^2.

3D.4 Exactly 20 years ago, the mean age of a group of boys was 15.7 years and the sum of the squares of their ages was 16 000. If the sum of the squares of their ages has increased by 8224 in this 20-year period, find the number of boys in the group.

3D.5 Readings from a device, denoted by y, are such that $\Sigma(y - 3)^2 = 2775$, $\Sigma y = 105$ and the standard deviation of y is 13. Find the number of readings that were taken.

3D.6 Mei measured the heights of her classmates and, after correctly analysing her data, she found the mean and standard deviation to be 163.8 cm and 7.6 cm. Decide whether or not these measures are valid, given the fact that Mei measured all the heights from the end of the tape measure, which is exactly 1.2 cm from the zero mark. Explain your answers.

3D.7 A transport company runs 21 coaches between two cities every week. In the past, the mean and variance of the journey times were 4 hours 35 minutes and 53.29 minutes^2. What would be the mean and standard deviation of the times if all the coaches departed 10 minutes later and arrived 5 minutes earlier than in the past? Are there any situations in which achieving this might actually be possible?

3D.8 During a sale, a boy bought six pairs of jeans, each with leg length x cm. He also bought four pairs of pants, each with leg length (x – 2) cm. The boy is quite short, so his father removed 4 cm from the length of each trouser leg. Find the variance of the leg lengths after his father made the alterations.

3D.9

(a) Find the mean and standard deviation of the first seven positive even integers.

(b) Without using a calculator, write down the mean and standard deviation of the first seven positive odd integers.

(c) Find an expression in terms of n for the variance of the first n positive even integers. What other measure can be found using this expression?

3D.10 Each year Upchester United plays against Upchester City in a local derby match. The number of goals scored in a match by United is denoted by u and the number of goals scored in a match by City is denoted by c. The number of goals scored in the past 15 matches are summarised by $\Sigma(u - 1)^2 = 25$, $\Sigma(u - 1) = 9$, $\Sigma c^2 = 39$ and $\Sigma c = 19$.

(a) How many goals have been scored altogether in these 15 matches?

(b) Show that $\Sigma u^2 = 58$.

(c) Find, correct to 3 decimal places, the variance of the number of goals scored by the two teams together in these 15 matches.

3D.11 Twenty values of x are summarised by $\Sigma(x - 1)^2 = 132$ and $\Sigma(x - 1) = 44$. Eighty values of y are summarised by $\Sigma(y + 1)^2 = 17\,704$ and $\Sigma(y + 1) = 1184$.

(a) Show that $\Sigma x = 64$ and that $\Sigma x^2 = 240$.

(b) Calculate the value of Σy and of Σy^2.

(c) Find the exact variance of the 100 values of x and y combined.

3D.12 The heights, x cm, of 200 boys and the heights, y cm, of 300 girls are summarised by the following totals:

$\Sigma(x - 160)^2 = 18\,240$, $\Sigma(x - 160) = 1820$, $\Sigma(y - 150)^2 = 20\,100$, $\Sigma(y - 150) = 2250$.

(a) Find the mean height of these 500 children.

(b) By first evaluating Σx^2 and Σy^2, find the variance of the heights of the 500 children, including appropriate units with your answer.

Exercise 3E · Further coding problems

3E.1 The range of prices of the newspapers sold at a kiosk is \$0.80. After 6 p.m. all prices are reduced by 20%. Find the range of the prices after 6 p.m.

3E.2 Find the standard deviation of x, given that $\Sigma 4x^2 = 14\,600$, $\Sigma 2x = 420$ and $n = 20$.

3E.3 The values of x given in the table have a standard deviation of 0.88. Find the standard deviation of the values of y.

x	2	4	6
f	a	b	c
y	7	13	19
f	a	b	c

3E.4 The temperatures, T °Celsius, at seven locations in the Central Kalahari Game Reserve were recorded at 4 p.m. one January afternoon: 32.1, 31.7, 31.2, 31.5, 31.9, 32.2 and 32.7.

(a) Evaluate $\Sigma 10(T - 30)$ and $\Sigma 100(T - 30)^2$.

(b) Use your answers to part a to calculate the standard deviation of T.

(c) By 5 p.m. the temperature at each location had dropped by exactly 0.75°C. Find the variance of the temperatures at 5 p.m.

3E.5 Building plots are offered for sale at $315 per square metre. The seller has to pay a lawyer's fee of $500 from the money received. Salome's plot is 240 square metres larger than Nadia's plot. How much more did the seller receive from Salome than from Nadia after paying the lawyer's fees?

3E.6 Temperatures in degrees Celsius (°C) can be converted to temperatures in degrees Fahrenheit (°F) using the formula $F = 1.8C + 32$.

(a) The temperatures yesterday had a range of 15°C. Express this range in degrees Fahrenheit.

(b) Temperatures elsewhere were recorded at hourly intervals in degrees Fahrenheit and were found to have mean 54.5 and variance 65.61. Find the mean and standard deviation of these temperatures in degrees Celsius.

3E.7 Ten items were selected from each of four sections at a supermarket. Details of the prices of those items, in dollars, on 1st April and on 1st June are shown in the following table:

Section	1st April Mean	1st April SD	1st June Mean	1st June SD
Bakery	2.50	0.40	2.25	0.36
Household	5.00	1.20	4.75	1.25
Tinned food	4.00	0.60	4.10	0.60
Fruit & veg	2.00	0.40	2.00	0.50

For which section's items could each of the following statements be true? Briefly explain each of your answers:

(a) The total cost of the items did not change.

(b) The price of each item changed by the same amount.

(c) The proportional change in the price of each item was the same.

3E.8 The lengths of 45 ropes used at an outdoor recreational centre can be extended by 30% when stretched. The sum of the squares of their stretched lengths is 0.0507 km^2 and their natural lengths, x metres, are summarised by $\Sigma(x - 20)^2 = 1200$. Find the mean natural length of these 45 ropes.

3E.9 Over a short period of time in 2016, the value of the pound sterling (£) fell by 15.25% against the euro (€). Find the percentage change in the value of the euro against the pound over this same period.

End-of-chapter review exercise 3

R3.1 Three boys and seven girls are asked how much money they have in their pockets. The boys have \$2.50 each and the mean amount that the 10 children have is \$3.90.

(a) Show that the girls have a total of \$31.50. [1]

(b) Given that the seven girls have equal amounts of money, find the standard deviation of the amounts that the 10 children have. [3]

R3.2 Jean Luc was asked to record the times of 20 athletes in a long distance race. He started his stopwatch when the race began and then went to sit in the shade, where he fell asleep. On waking, he found that x athletes had already completed the race but he was able to record the times taken by all the others. State the possible value(s) of x if Jean Luc was able to use his data to calculate:

(a) the variance of the times taken by the 20 athletes [1]

(b) the interquartile range of the times taken by the 20 athletes. [2]

R3.3 The quiz marks of nine students are written down in ascending order and it is found that the range and interquartile range are equal. Find the greatest possible number of distinct marks that were obtained by the nine students. [2]

R3.4 Two days before a skiing competition, the depths of snow, x metres, at 32 points on the course were measured and it was discovered that the numerical values of Σx and Σx^2 were equal.

(a) Given that the mean depth of snow was 0.885 m, find the standard deviation of x. [2]

(b) Snow fell the day before the competition, increasing the depth over the whole course by 1.5 cm. Explain what effect this had on the mean and on the standard deviation of x. [2]

R3.5 The percentage scores of a class of students in the three Mathematics tests they took this term have five-figure summaries:

First test: Min 30, Q1 43, Median 48, Q3 69, Max 80.

Second test: Min 35, Q1 44, Median 58, Q3 71, Max 93.

Third test: Min 33, Q1 49, Median 68, Q3 80, Max 95.

(a) Describe the progress made by the class in Mathematics tests this term. [2]

(b) Which of the tests has produced the least skewed set of scores? [1]

(c) What type of skew do the scores in each of the other two tests have? [2]

R3.6 The following table shows the mean and standard deviation of the lengths of 75 adult puff adders (Bitis arietans), which are found in Africa and on the Arabian peninsula:

Population	Frequency	Mean (cm)	SD (cm)
African	60	102.7	6.8
Arabian	15	78.8	4.2

(a) Find the mean length of the 75 puff adders. [3]

(b) The lengths of individual African puff adders are denoted by x_f and the lengths of individual Arabian puff adders by x_b. By first finding Σx_f^2 and Σx_b^2, calculate the standard deviation of the lengths of all 75 puff adders. [5]

R3.7 The scores obtained by 11 people throwing three darts each at a dartboard are 54, 46, 43, 52, 180, 50, 41, 56, 52, 49 and 54.

(a) Find the range, the interquartile range and the standard deviation of these scores. [4]

(b) Which measure in part a best summarises the variation of the scores? Explain why you have chosen this particular measure. [2]

R3.8 The heights, x cm, of a group of 28 people were measured. The mean height was found to be 172.6 cm and the standard deviation was found to be 4.58 cm. A person whose height was 161.8 cm left the group.

(i) Find the mean height of the remaining group of 27 people. [2]

(ii) Find Σx^2 for the original group of 28 people. Hence find the standard deviation of the heights of the remaining group of 27 people. [4]

R3.9 120 people were asked to read an article in a newspaper. The times taken, to the nearest second, by the people to read the article are summarised in the following table:

Time (seconds)	1-25	26-35	36-45	46-55	56-90
Number of people	4	24	38	34	20

Calculate estimates of the mean and standard deviation of the reading times. [5]

R3.10 The weights, in kilograms, of the 15 basketball players in each of two squads, A and B, are shown:

Squad A: 97, 98, 104, 84, 100, 109, 115, 99, 122, 82, 116, 96, 84, 107, 91

Squad B: 75, 79, 94, 101, 96, 77, 111, 108, 83, 84, 86, 115, 82, 113, 95

(i) Represent the data by drawing a back-to-back stem-and-leaf diagram with squad A on the left-hand side and squad B on the right-hand side. [4]

(ii) Find the interquartile range of the weights of the players in squad A. [2]

(iii) A new player joins squad B. The mean weight of the 16 players in squad B is now 93.9 kg. Find the weight of the new player. [3]

R3.11 The heights, x cm, of a group of 82 children are summarised as follows: $\Sigma(x - 130) = -287$, standard deviation of x = 6.9.

(i) Find the mean height. [2]

(ii) Find $\Sigma(x - 130)^2$. [2]

R3.12 A sample of 36 data values, x, gave $\Sigma(x - 45) = -148$ and $\Sigma(x - 45)^2 = 3089$.

(i) Find the mean and standard deviation of the 36 values. [3]

(ii) One extra data value of 29 was added to the sample. Find the standard deviation of all 37 values. [4]

R3.13 The ages, x years, of 150 cars are summarised by $\Sigma x = 645$ and $\Sigma x^2 = 8287.5$. Find $\Sigma(x - \bar{x})^2$, where $\bar{x}$ denotes the mean of x. [4]

R3.14 A set of data values is 152, 164, 177, 191, 207, 250 and 258. Compare the proportional change in the standard deviation with the proportional change in the interquartile range when the value 250 in the dataset is increased by 40%. [5]

R3.15 A shop has in its stock 80 rectangular celebrity posters. All of these posters have a width to height ratio of $1 : \sqrt{2}$, and their mean perimeter is 231.8 cm. Given that the sum of the squares of the widths is 200 120 cm^2, find the standard deviation of the widths of the posters. [4]

R3.16 At a village fair, visitors were asked to guess how many sweets are in a glass jar. The best six guesses were 180, 211, 230, 199, 214 and 166.

(a) Show that the mean of these guesses is 200, and use $SD = \sqrt{(\Sigma(x - \bar{x})^2/n)}$ to calculate the standard deviation. [4]

(b) The jar actually contained 202 sweets. Without further calculation, write down the mean and the standard deviation of the errors made by these six visitors. Explain why no further calculations are required to do this. [4]

R3.17 The number of women in senior management positions at a number of companies was investigated. The number of women at each of the 25 service companies and at each of the 16 industrial companies are denoted by w_S and w_I, respectively. The findings are summarised by the totals:

$\Sigma(w_S - 5)^2 = 28$, $\Sigma(w_S - 5) = 15$, $\Sigma(w_I - 3)^2 = 12$, $\Sigma(w_I - 3) = -4$.

(a) Show that there are, on average, more than twice as many women in senior management positions at the service companies than at the industrial companies. [3]

(b) Show that $\Sigma w_S^2 \neq (\Sigma w_S)^2$ and that $\Sigma w_I^2 \neq (\Sigma w_I)^2$. [5]

(c) Find the standard deviation of the number of women in senior management positions at all of these service and industrial companies together. [3]

R3.18 The ages, a years, of the five members of the boy-band AlphaArise are such that $\Sigma(a - 21)^2 = 11.46$ and $\Sigma(a - 21) = -6$.

The ages, b years, of the seven members of the boy-band BetaBeat are such that $\Sigma(b - 18)^2 = 10.12$ and $\Sigma(b - 18) = 0$.

(a) Show that the difference between the mean ages of the boys in the two bands is 1.8 years. [3]

(b) Find the variance of the ages of the 12 members of these two bands. [7]

Cross-topic review exercise 1

CT1.1 Two players, A and B, both played seven matches to reach the final of a tennis tournament. The number of games that each of them won in these matches are given in a back-to-back stem-and-leaf diagram (Key: 5 | 2 | 6 represents 25 games for A and 26 games for B):

Player A (7): 20, 22, 22, 25, 27, 28, 33 games.

Player B (7): 18, 19, 21, 21, 23, 24, 26 games.

(a) How many fewer games did player B win than player A? [1]

(b) Find the median number of games won by each player. [2]

(c) In a single stem-and-leaf diagram, show the number of games won by these two players in all of the 14 matches they played to reach the final. [3]

CT1.2 A total of 112 candidates took a multiple-choice test that had 40 questions. The numbers of correct answers given by the candidates are shown in the following table:

No. correct answers	0-9	10-15	16-25	26-30	31-39	40
No. candidates	18	24	27	23	19	1

(a) State which class contains the lower quartile and which class contains the upper quartile. Hence, find the least possible value of the interquartile range. [3]

(b) Copy and complete the following table, which shows the numbers of incorrect answers given by the candidates in the test:

No. incorrect answers	0	1-9	10-14	15-24	25-30	31-40
No. candidates	1					

[3]

(c) Calculate an estimate of the mean number of incorrectly answered questions. [3]

CT1.3 At a factory, 50-metre lengths of cotton thread are wound onto bobbins. Due to fraying, it is common for a length, l cm, of cotton to be removed after it has been wound onto a bobbin. The following table summarises the lengths of cotton thread removed from 200 bobbins:

Length removed (l cm)	0 <= l < 2.5	2.5 <= l < 5.0	5 <= l < 10
No. bobbins	137	49	14

(a) Calculate an estimate of the mean length of cotton removed. [3]

(b) Use your answer to part a to calculate, in metres, an estimate of the standard deviation of the length of cotton remaining on the 200 bobbins. [4]

CT1.4 People applying to a Computing college are given an aptitude test. Those who are accepted take a progress test 3 months after the course has begun. The following table gives the aptitude test scores, x, and the progress test scores, y, for a random sample of eight students, A to H:

Student	A	B	C	D	E	F	G	H
x	61	80	74	60	83	92	71	67
y	53	77	61	70	81	54	63	85

(a) Find the interquartile range of these aptitude test scores. [1]

(b) Use the summary totals $\Sigma x = 588$, $\Sigma x^2 = 44\,080$, $\Sigma y = 544$ and $\Sigma y^2 = 38\,030$ to calculate the variance of the aptitude and progress test scores when they are considered together. [3]

(c) The mean progress score for all the students at the college is 70 and the variance is 112. Any student who scores less than 1.5 standard deviations below the mean is sent a letter advising that improvement is needed. Which of the students A to H should the letter be sent to? [1]

CT1.5 The growth of 200 tomato plants, half of which were treated with a growth hormone, was monitored over a 5-day period and is summarised in cumulative frequency graphs:

- Treated: passes through (3.5, 0), (5, 12), (6, 40), (7, 80), (8, 100).
- Untreated: passes through (1, 0), (3, 38), (5, 55), (7, 68), (10, 100).

Use the graphs to describe two advantages of treating these tomato plants with the growth hormone. [2]

CT1.6 A survey of a random sample of 23 people recorded the number of unwanted emails they received in a particular week:

9, 18, 13, 18, 21, 17, 22, 27, 8, 11, 26, 26, 32, 17, 31, 20, 36, 15, 13, 25, 35, 29, 14

(a) Represent the data in a stem-and-leaf diagram. [3]

(b) Draw, on graph paper, a box-and-whisker diagram to represent the data. [4]

CT1.7 The volumes of water, $x \times 10^6$ litres, needed to fill six Olympic-sized pools are: 2.82, 2.50, 2.75, 3.14, 3.66 and 3.07.

(a) Find the value of $\Sigma(x - 2)$ and of $\Sigma(x - 2)^2$. [2]

(b) Use your answers to part a to find the mean and the standard deviation of the volumes of water, giving both answers correct to the nearest litre. [5]

CT1.8 The speeds of 72 coaches at a certain point on their journeys between two cities were recorded. The results are given in the following table:

Speed (km/h)	< 50	< 54	< 70	< 75	< 85
Cumulative frequency	0	9	41	54	72

(a) State the number of coaches whose speeds were between 54 and 70 km/h. [1]

(b) A student has illustrated the data in a cumulative frequency polygon. Find the two speeds between which the polygon has the greatest gradient. [1]

(c) Calculate an estimate of the lower boundary of the speeds of the fastest 25 coaches. [3]

CT1.9 The following table shows the masses, m grams, of 100 unsealed bags of plain potato crisps:

Mass (m grams)	34.6 <= m < 35.4	35.4 <= m < 36.2	36.2 <= m < 37.2
No. bags f	20	30	50

(a) Show that the heights of the columns in a histogram illustrating these data must be in the ratio 2 : 3 : 4. [2]

(b) Calculate estimates of the mean and of the standard deviation of the masses. [4]

(c) Before each bag is sealed, 0.05 grams of salt is added. Find the variance of the masses of the sealed bags of salted potato crisps. [1]

CT1.10 The masses, in carats, of a sample of 200 pearls are summarised in a cumulative frequency graph passing through (0, 0), (20, 50), (40, 68), (60, 84), (80, 115), (90, 160), (100, 190), (110, 200). One carat is equivalent to 200 milligrams.

(a) Use the graph to estimate, in carats: (i) the median mass of the pearls [1] (ii) the interquartile range of the masses. [2]

(b) To qualify as a 'paragon', a pearl must be flawless and weigh at least 20 grams. Use the graph to estimate the largest possible number of paragons in the sample. [2]

CT1.11 The amounts spent, S dollars, by six customers at a hairdressing salon yesterday were as follows: 12.50, 15.75, 41.30, 34.20, 10.80, 40.85. Each of the customers paid with a $50 note and each received the correct change, which is denoted by $C.

(a) Find, in dollars, the value of $\bar{S} + \bar{C}$ and of $\bar{S} - \bar{C}$. [3]

(b) Explain why the standard deviation of S and the standard deviation of C are identical. [3]

CT1.12 The numbers of goals, x, scored by a team in each of its previous 25 games are summarised by the totals $\Sigma(x - 1)^2 = 30$ and $\Sigma(x - 1) = 12$.

(a) Find the mean number of goals that the team scored per game. [2]

(b) Find the value of Σx^2. [3]

(c) Find the value of a and of b in the following table, which shows the frequencies of the numbers of goals scored by the team:

No. goals	0	1	2	3	> 3
No. games f	5	a	b	4	0

[2]

CT1.13 The lengths of some insects of the same type from two countries, X and Y, were measured. The results for 80 insects from Country X and 80 insects from Country Y are summarised:

Country X: Leaves on stems 80 to 84:

Stem 80 (10 leaves): 2, 3, 4, 4, 4, 6, 6, 6, 7, 9

Stem 81 (18 leaves): 1, 1, 3, 3, 4, 4, 5, 5, 6, 7, 8, ...

Stem 82 (16 leaves): 0, 2, 2, 3, 3, 3, 4, 4, 5, 5, 5, 6, 6, 7, 7, 8, 8, 9, 9

Stem 83 (16 leaves): 0, 0, 2, 2, 2, 3, 5, 5, 6, 7, 7, 8, 8, 9, 9, 9

Stem 84 (11 leaves): 0, 0, 1, 1, 1, 2, 2, 2, 3, 3, 3, 5, 5, 5, 6, 7, 8

Country Y: Leaves on stems 81 to 86:

Stem 81 (13 leaves): 1, 1, 2, 2, 3, 3, 3, 5, 5, 6, 7, 8, 9

Stem 82 (12 leaves): 1, 2, r, 3, 3, 5, 5, 6, 6, 7, 8, 8

Stem 83 (11 leaves): 0, 1, 2, 2, 3, 5, 5, 5, 8, 9, 9

Stem 84 (15 leaves): 0, 0, 1, 2, 3, 3, 3, q, 4, 5, 6, 6, 7, 8, 8

Stem 85 (15 leaves): 0, 0, 1, 2, 4, 4, 5, 5, 6, 6, 7, 7, 7, 8, 9

Stem 86 (14 leaves): 0, 1, 2, 2, 4, 4, 4, 4, 5, 5, 6, 6, 7, 7, 7, 8, 9

Key: 5 | 81 | 3 means an insect from country X has length 0.815 cm and an insect from country Y has length 0.813 cm.

(i) Find the median and interquartile range of the lengths of the insects from country X. [2]

(ii) The interquartile range of the lengths of the insects from country Y is 0.028 cm. Find the values of q and r. [2]

(iii) Represent the data by means of a pair of box-and-whisker plots in a single diagram on graph paper. [4]

(iv) Compare the lengths of the insects from the two countries. [2]

Check yourself

CY.1 Five numbers are 3, 7, 8, 12, 15.

(a) Find the mean and variance of these five numbers. [3]

(b) If each number is transformed by $y = 4x - 5$, state the mean and standard deviation of y. [2]

CY.2 A dataset of 20 observations of x has summary totals $\Sigma(x - 12) = 40$ and $\Sigma(x - 12)^2 = 720$.

(a) Find the mean and standard deviation of x. [4]

(b) Hence find the value of Σx^2. [2]

CY.3 Group A contains 15 students with mean mark 64 and standard deviation 8. Group B contains 25 students with mean mark 72 and standard deviation 6.

(a) Find the combined mean mark of all 40 students. [2]

(b) Show that the sum of squares of the marks for Group A is 62 400 and find the corresponding sum of squares for Group B. [3]

(c) Hence calculate the standard deviation of the marks of all 40 students combined. [3]

CY.4 The continuous variable T represents reaction times in milliseconds. A cumulative frequency polygon shows that Q1 = 185 ms, median Q2 = 210 ms, and Q3 = 245 ms.

(a) Calculate the interquartile range of T. [1]

(b) Determine whether a recorded time of 340 ms is an outlier. [2]

(c) State, with a reason, the skewness of this distribution. [2]

CY.5 Prove that for any n data values $x_1, x_2, \ldots, x_n$ with mean $\bar{x}$:

$\Sigma(x - \bar{x})^2 = \Sigma x^2 - n(\bar{x})^2$. [4]

Chapter quiz

Q1 Find the range of the numbers -4.5, 2.1, -1.8, 6.7, 0, 3.4.

A	B	C	D
11.2	8.5	9.0	10.1

Q2 Seven numbers are 4, 7, 9, 12, 15, 18, 22. Find the interquartile range (IQR).

A	B	C	D
11	7	18	9

Q3 In a box plot, Q1 = 24 and Q3 = 40. Which value is classified as an outlier?

A	B	C	D
65	63	1	2

Q4 For a dataset, $\Sigma x = 120$, $\Sigma x^2 = 1840$, and $n = 10$. Find the variance.

A	B	C	D
40	16	144	4

Q5 For a sample of 25 values, $\Sigma(x - 8)^2 = 225$. Find the standard deviation if the mean is 8.

A	B	C	D
3	9	15	2.5

Q6 If every value in a dataset is increased by 5, what happens to the standard deviation?

A	B	C	D
It remains unchanged	It increases by 5	It increases by 25	It is multiplied by 5

Q7 The variable y is coded as $y = 3x - 4$. If $\text{Var}(x) = 16$, find the standard deviation of y.

A	B	C	D
12	44	48	36

Q8 The variable u is defined by $u = (x - 20)/5$. If $\text{Var}(u) = 4$, find $\text{Var}(x)$.

A	B	C	D
100	20	4	80

Q9 Given $n = 20$, $\Sigma x = 100$, and $\text{Var}(x) = 15$, find the value of Σx^2.

A	B	C	D
800	500	600	1000

Q10 Set A has 10 values with mean 12, Set B has 10 values with mean 18. Find the combined mean.

A	B	C	D
15	14	16	13.5

Q11 Why is the interquartile range (IQR) preferred over the range when outliers are present?

A	B	C	D
IQR ignores extreme values	IQR uses every value in data	IQR is always larger	IQR equals the standard deviation

---

Q12 A box plot has Q1 = 15, Q2 = 28, and Q3 = 35. What is the skewness of this distribution?

A	B	C	D
Negative (left-skewed)	Positive (right-skewed)	Symmetrical	Zero skew

---

End of assignment · Check your answers against the chapter at gurujazzy.com/subjects/as-a-level-mathematics/s1-3-measures-of-variation/