
Stats1 Chapter 7: Data Measures

Chapter Practice

Key Points

- 1 The **mode** or **modal class** is the value or class that occurs most often.
- 2 The **median** is the middle value when the data values are put in order.
- 3 The **mean** can be calculated using the formula $\bar{x} = \frac{\sum x}{n}$.
- 4 For data given in a frequency table, the mean can be calculated using the formula $\bar{x} = \frac{\sum xf}{\sum f}$.
- 5 To find the **lower quartile** for discrete data, divide n by 4. If this is a whole number, the lower quartile is halfway between this data point and the one above. If it is not a whole number, round *up* and pick this data point.
- 6 To find the **upper quartile** for discrete data, find $\frac{3}{4}$ of n . If this is a whole number, the upper quartile is halfway between this data point and the one above. If it is not a whole number, round *up* and pick this data point.
- 7 The **range** is the difference between the largest and smallest values in the data set.
- 8 The **interquartile range** (IQR) is the difference between the upper quartile and the lower quartile, $Q_3 - Q_1$.
- 9 The **interpercentile range** is the difference between the values for two given percentiles.

Key Points

10 Variance $= \frac{\sum (x - \bar{x})^2}{n} = \frac{\sum x^2}{n} - \left(\frac{\sum x}{n} \right)^2 = \frac{S_{xx}}{n}$ where $S_{xx} = \sum (x - \bar{x})^2 = \sum x^2 - \frac{(\sum x)^2}{n}$

11 The **standard deviation** is the square root of the variance:

$$\sigma = \sqrt{\frac{\sum (x - \bar{x})^2}{n}} = \sqrt{\frac{\sum x^2}{n} - \left(\frac{\sum x}{n} \right)^2} = \sqrt{\frac{S_{xx}}{n}}$$

12 You can use these versions of the formulae for variance and standard deviation for grouped data that is presented in a frequency table:

$$\sigma^2 = \frac{\sum f(x - \bar{x})^2}{\sum f} = \frac{\sum fx^2}{\sum f} - \left(\frac{\sum fx}{\sum f} \right)^2 \quad \sigma = \sqrt{\frac{\sum f(x - \bar{x})^2}{\sum f}} = \sqrt{\frac{\sum fx^2}{\sum f} - \left(\frac{\sum fx}{\sum f} \right)^2}$$

where f is the frequency for each group and $\sum f$ is the total frequency.

13 If data is coded using the formula $y = \frac{x - a}{b}$

- the mean of the coded data is given by $\bar{y} = \frac{\bar{x} - a}{b}$
- the standard deviation of the coded data is given by $\sigma_y = \frac{\sigma_x}{b}$ where σ_x is the standard deviation of the original data.

Chapter Exercises

- 1 The mean science mark for one group of eight students is 65. The mean mark for a second group of 12 students is 72. Calculate the mean mark for the combined group of 20 students.

- 2 The data shows the prices (x) of six shares on a particular day in the year 2007:

807 967 727 167 207 767

a Code the data using the coding $y = \frac{x - 7}{80}$.

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 of employees' annual earnings (£ x) for a store is 18. The coding used was

$y = \frac{x - 720}{1000}$. Work out the uncoded mean earnings.

- 4 Different teachers using different methods taught two groups of students. Both groups of students sat the same examination at the end of the course. The students' marks are shown in the grouped frequency table.

Exam mark	20–29	30–39	40–49	50–59	60–69	70–79	80–89
Frequency group A	1	3	6	6	11	10	8
Frequency group B	1	2	4	13	15	6	3

- a Work out an estimate of the mean mark for group A and an estimate of the mean mark for group B.
- b Write down whether or not the answer to a suggests that one method of teaching is better than the other. Give a reason for your answer.

Chapter Exercises

- 5 The lifetimes of 80 batteries, to the nearest hour, are shown in the table below.

Lifetime (hours)	6–10	11–15	16–20	21–25	26–30
Frequency	2	10	18	45	5

a Write down the modal class for the lifetime of the batteries.

b Use interpolation to find the median lifetime of the batteries.

The midpoint of each class is represented by x and its corresponding frequency by f , giving $\Sigma fx = 1645$.

c Calculate an estimate of the mean lifetime of the batteries.

Another batch of 12 batteries is found to have an estimated mean lifetime of 22.3 hours.

d Estimate the mean lifetime for all 92 batteries.

- 6 A frequency distribution is shown below.

Class interval	1–20	21–40	41–60	61–80	81–100
Frequency	5	10	15	12	8

Use interpolation to find an estimate for the interquartile range.

- 7 A frequency distribution is shown below.

Class interval	1–10	11–20	21–30	31–40	41–50
Frequency	10	20	30	24	16

a Use interpolation to estimate the value of the 30th percentile.

b Use interpolation to estimate the value of the 70th percentile.

c Hence estimate the 30% to 70% interpercentile range.

Chapter Exercises

- 8 The times it took a random sample of runners to complete a race are summarised in the table.

Time taken (t minutes)	20–29	30–39	40–49	50–59	60–69
Frequency	5	10	36	20	9

- a Use interpolation to estimate the interquartile range. (3 marks)

The midpoint of each class was represented by x and its corresponding frequency by f giving:

$$\Sigma fx = 3740 \quad \Sigma fx^2 = 183\,040$$

- b Estimate the variance and standard deviation for this data. (3 marks)

- 9 The heights of 50 clover flowers are summarised in the table.

Heights in mm (x)	$90 \leq x < 95$	$95 \leq x < 100$	$100 \leq x < 105$	$105 \leq x < 110$	$110 \leq x < 115$
Frequency	5	10	26	8	1

- a Find Q_1 . b Find Q_2 . c Find the interquartile range.

- d Use $\Sigma fx = 5075$ and $\Sigma fx^2 = 516\,112.5$ to find the standard deviation.

- 10 The daily mean temperature is recorded in Camborne during September 2015.

Temperature, t ($^{\circ}\text{C}$)	$11 < t \leq 13$	$13 < t \leq 15$	$15 < t \leq 17$
Frequency	12	14	4

© Crown Copyright Met Office

- a Use your calculator to find estimates for the mean and standard deviation of the temperatures. (3 marks)

- b Use linear interpolation to find an estimate for the 10% to 90% interpercentile range. (3 marks)

- c Estimate the number of days in September 2015 where the daily mean temperature in Camborne is more than one standard deviation greater than the mean. (2 marks)

Chapter Exercises

- 11 The daily mean windspeed, w knots was recorded at Heathrow during May 2015. The data were coded using $z = \frac{w - 3}{2}$.

Summary statistics were calculated for the coded data:

$$n = 31 \quad \Sigma z = 106 \quad S_{zz} = 80.55$$

- a Find the mean and standard deviation of the coded data. (2 marks)
- b Work out the mean and standard deviation of the daily mean windspeed at Heathrow during May 2015. (2 marks)

- 12 20 endangered forest owlets were caught for ringing. Their wingspans (x cm) were measured to the nearest centimetre.

The following summary statistics were worked out:

$$\Sigma x = 316 \quad \Sigma x^2 = 5078$$

- a Work out the mean and the standard deviation of the wingspans of the 20 birds. (3 marks)
- One more bird was caught. It had a wingspan of 13 centimetres.
- b Without doing any further calculation, say how you think this extra wingspan will affect the mean wingspan. (1 mark)

20 giant ibises were also caught for ringing. Their wingspans (y cm) were also measured to the nearest centimetre and the data coded using $z = \frac{y - 5}{10}$.

The following summary statistics were obtained from the coded data:

$$\Sigma z = 104 \quad S_{zz} = 1.8$$

- c Work out the mean and standard deviation of the wingspans of the giant ibis. (5 marks)

Chapter Exercises

Challenge

A biologist recorded the heights, x cm, of 20 plant seedlings. She calculated the mean and standard deviation of her results:

$$\bar{x} = 3.1 \text{ cm} \quad \sigma = 1.4 \text{ cm}$$

The biologist subsequently discovered she had written down one value incorrectly. She replaced a value of 2.3 cm with a value of 3.2 cm.

Calculate the new mean and standard deviation of her data.

Chapter Answers

- 1 69.2
- 2 a 10, 12, 9, 2, 2.5, 9.5 b 7.5 c 607
- 3 £18 720
- 4 a Group A 63.4, group B 60.2
b The method used for group A may be better.
- 5 a 21 to 25 hours b 21.6 hours
c 20.6 hours d 20.8 hours
- 6 37.5
- 7 a 20.5 b 34.7 c 14.2
- 8 a 13.1
b Variance 102, standard deviation 10.1 minutes
- 9 a 98.75 mm b 104 mm c 5.58 mm d 4.47 mm
- 10 a Mean 13.5, standard deviation 1.36
b 4.0 °C c 5 days
- 11 a Mean 3.42, standard deviation 1.61
b Mean 9.84 knots, standard deviation 3.22 knots
- 12 a Mean 15.8 cm, standard deviation 2.06 cm
b The mean wingspan will decrease.
c Mean 57 cm, standard deviation 3 cm

Challenge

Mean 3.145 cm, standard deviation 1.39 cm