

Exercise walk-through

2.7

Statistics – Numerical Measures of Variability

GAC Mathematics

1

The method

Method — Key knowledge: equal means can conceal different distributions

The **range** 全距 is maximum minus minimum. The **interquartile range (IQR)** 四分位距 is $Q_3 - Q_1$. For this sheet, order the data, split an even-sized list into equal halves, and use each half's median for its quartile. For an odd list, exclude the overall median before finding half-list medians. Other software may use different conventions. A box plot shows the median and quartiles; this sheet's whiskers run to minimum and maximum, not to an automatic outlier fence. **Standard deviation** 标准差 measures spread using squared deviations. Use population $\sigma = \sqrt{\sum(x - \mu)^2 / n}$ for a complete supplied population; sample estimate $s = \sqrt{\sum(x - \bar{x})^2 / (n - 1)}$ requires the stated sample convention. Standard deviation has the original units; variance has squared units.

Method — Worked example: read a five-number summary rather than every observation

Known five-number summary: minimum 2, $Q_1 = 4$, median 6, $Q_3 = 8$, maximum 12. The overall span and middle-half span are different.

$$R = x_{\max} - x_{\min} = 12 - 2 = 10$$

$$IQR = Q_3 - Q_1 = 8 - 4 = 4$$

The box does not identify every original value, the sample size or why the spread arose.

Exercises 2.1·2.2

2.1 Find the range of 3, 5, 8 and 12.

[2]

2.2 A five-number summary is minimum 4, first quartile 7, median 9, third quartile 13, maximum 20. Find its IQR and range.

[2]

Exercises 2.1·2.2

2.1 Find the range of 3, 5, 8 and 12.

[2]

2.1

Subtract the minimum from the maximum.

$$R = x_{max} - x_{min} = 12 - 3 = 9$$

Range 9.

2.2 A five-number summary is minimum 4, first quartile 7, median 9, third quartile 13, maximum 20. Find its IQR and range.

[2]

2.2

$$IQR = Q_3 - Q_1 = 13 - 7 = 6. \text{ Range } R = 20 - 4 = 16.$$

Exercise 2.3

2.3 Using the stated half-list convention, find the median, first quartile and third quartile of 2, 4, 6, 8, 10, 12, 14, 16. [3]

Exercise 2.3

2.3 Using the stated half-list convention, find the median, first quartile and third quartile of 2, 4, 6, 8, 10, 12, 14, 16. [3]

2.3

Median is the mean of the fourth and fifth values: $m = (8 + 10)/2 = 9$. Lower-half median $Q_1 = (4 + 6)/2 = 5$. Upper-half median $Q_3 = (12 + 14)/2 = 13$.

Exercise 2.4

2.4 The complete population is 2, 4, 6. Find its mean, population variance and population standard deviation, giving the latter to 3 significant figures. [3]

Exercise 2.4

2.4 The complete population is 2, 4, 6. Find its mean, population variance and population standard deviation, giving the latter to 3 significant figures. [3]

2.4

Mean $\mu = (2 + 4 + 6)/3 = 4$.

$$\sigma^2 = \frac{\sum (x - \mu)^2}{n} = \frac{(2 - 4)^2 + (4 - 4)^2 + (6 - 4)^2}{3} = \frac{8}{3}$$

Population variance.

$$\sigma = \sqrt{\sigma^2} = \sqrt{8/3} \approx 1.63$$

Population standard deviation.

Exercise 3.1

3.1 Two complete populations are A: 4, 5, 6 and B: 1, 5, 9. Find each mean and population standard deviation, to 3 significant figures. Compare their spread without claiming a cause. [6]

Exercise 3.1

3.1 Two complete populations are A: 4, 5, 6 and B: 1, 5, 9. Find each mean and population standard deviation, to 3 significant figures. Compare their spread without claiming a cause. [6]

3.1

Both means are 5. Use population division by n, not n minus 1.

$$\sigma_A = \sqrt{\frac{(4-5)^2 + (5-5)^2 + (6-5)^2}{3}} = \sqrt{2/3} \approx 0.816$$

A standard deviation.

$$\sigma_B = \sqrt{\frac{(1-5)^2 + (5-5)^2 + (9-5)^2}{3}} = \sqrt{32/3} \approx 3.27$$

B standard deviation. B is more spread around the shared mean. The summaries do not identify why the populations differ.

Exercise 3.2

3.2 The ordered data are 2, 4, 6, 8, 10, 12, 14, 16. (a) Draw a minimum-to-maximum box plot using the stated quartile convention. (b) Find the range and IQR. (c) Replace 16 by 100 and find the new range and IQR. Explain why these two spread measures respond differently and whether the plot alone can establish the cause of the extreme value. [8]

Exercise 3.2

3.2 The ordered data are 2, 4, 6, 8, 10, 12, 14, 16. (a) Draw a minimum-to-maximum box plot using the stated quartile convention. (b) Find the range and IQR. (c) Replace 16 by 100 and find the new range and IQR. Explain why these two spread measures respond differently and whether the plot alone can establish the cause of the extreme value. [8]

3.2

(a) Five-number summary is 2, 5, 9, 13, 16. Correct quartiles and median earn B1; a box from 5 to 13 with median 9 earns B1; whiskers to 2 and 16 on a labelled scale earn B1.

(b) Range $R = 16 - 2 = 14$. IQR $= 13 - 5 = 8$.

(c) New range $R' = 100 - 2 = 98$. Quartiles stay 5 and 13, so IQR remains 8. The range uses the changed extreme directly, while the middle-half endpoints stay fixed; neither the plot nor these summaries establishes the cause of 100. Do not delete that value without checking its origin.