

Lesson 5

A-level Statistics : Year 1 Partitioning Data

5.1 Mean & Standard deviation from Grouped Frequency Tables

Example

The time taken for ten runners to complete a race are summarised in the table;

| Time taken (minutes) | Number of runners |
|----------------------|-------------------|
| 20 - 24 | 3 |
| 25 - 29 | 5 |
| 30 - 34 | 2 |

Find the mean and standard deviation of the times taken.

Give your answers in minutes and seconds.

5.2 Exercise

Question 1

The time taken by 15 shoppers in a supermarket is presented in the table;

| Time in Shop (minutes) | Number of shoppers |
|---------------------------|-----------------------|
| 10 - 16 | 8 |
| 17 - 23 | 5 |
| 24 - 30 | 2 |

Find the mean and standard deviation of the shopping times.
Give your answers in minutes and seconds.

In examination questions you are often told Σfx and $\Sigma f x^2$

This means that *you don't need to work them out*, with considerable less work then being needed to answer the question.

The next question is such a question...

Question 2

The times it took a random sample of swimmers to complete a sponsored swim are summarised in the table;

| | | | | | |
|----------------|---------|---------|---------|---------|---------|
| Time (minutes) | 20 - 29 | 30 - 39 | 40 - 49 | 50 - 59 | 60 - 69 |
| Frequency | 5 | 10 | 36 | 20 | 9 |

The mid-point of each class was represented by x and its corresponding frequency by f giving

$$\Sigma fx = 3740 \quad \Sigma f x^2 = 183040$$

Determine the mean and the standard deviation of the swim times.
Give your answers in minutes and seconds.

Question 3

The table below summarises data relating to the Average Rated Lifetime (ARL) of a random sample of 200 light bulbs taken from the production line of the Light Up You Life Forever™ LED lightbulb company.

| Lifetime, x (to nearest 100 hours) | Number of light-bulbs, f | |
|---|-------------------------------|--|
| 700 - 719 | 10 | |
| 720 - 729 | 14 | |
| 730 - 739 | 16 | |
| 740 - 749 | 21 | |
| 750 - 754 | 35 | |
| 755 - 759 | 41 | |
| 760 - 764 | 38 | |
| 765 - 769 | 15 | |
| 770 - 779 | 7 | |
| 780 - 799 | 3 | |

- (i) By linear interpolation, estimate the median and quartiles of these lifetimes. Give your answers in 100 hours to 1 decimal place.

The formulae for the mean, μ , and the standard deviation, σ , are;

$$u = \frac{\Sigma fx}{\Sigma f} \quad \sigma = \sqrt{\frac{\Sigma f x^2}{\Sigma f} - \mu^2}$$

- (ii) Given that $\Sigma fx = 150\,232.5$ and $\Sigma f x^2 = 112\,899\,573.8$
- (b) determine the mean lifetime of a lightbulb.
- (c) determine the standard deviation of the data.

- (iii) One method of assessing the skewness of a distribution is to calculate

$$\frac{3 (\text{mean} - \text{median})}{\text{standard deviation}}$$

Evaluate this, to 1 decimal place, for the above distribution.

- (iv) Use the quartiles to assess skewness and state whether or not you feel the result is compatible with your answer to part (iii).

Question 4

SI examination question from January 2008, Q2

Cotinine is a chemical that is made by the body from nicotine which is found in cigarette smoke. A doctor tested the blood of 12 patients, who claimed to smoke a packet of cigarettes a day, for cotinine.

The results, in appropriate units, are shown below;

| Patient | A | B | C | D | E | F | G | H | I | J | K | L |
|---------------------|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Cotinine level, x | 160 | 390 | 169 | 175 | 125 | 420 | 171 | 250 | 210 | 258 | 186 | 243 |

[You may use $\Sigma x^2 = 724\,961$]

- (a) Find the mean and standard deviation of the level of cotinine in a patient's blood.

[4 marks]

- (b) Find the median, upper and lower quartiles of these data

[3 marks]

A doctor suspects that some of his patients have been smoking more than a packet of cigarettes per day. He decides to use $Q_3 + 1.5(Q_3 - Q_1)$ to determine if any of the cotinine results are far enough away from the upper quartile to be outliers.

- (c) Identify which patient(s) may have been smoking more than a packet of cigarettes a day. Show your working clearly.

[4 marks]

Research suggests that cotinine levels in the blood form a skewed distribution. One measure of skewness is found using

$$\frac{(Q_1 - 2Q_2 + Q_3)}{(Q_3 - Q_1)}$$

- (d) Evaluate this measure and describe the skewness of these data.

[3 marks]