

Statistics GCSE**Paper 2**

Edexcel Higher - 2026

Higher Tier

Variant 2

1ST0/2H

Instructions

- Write all answers in the spaces provided.
- Answer all questions.
- You must show all your working.
- There may not be enough space to show all your working out.

Information

- This is a practise paper to aid your revision for your exams.
- This site, and all that work on it, have no affiliation or relationship with any exam board.
- This site is not endorsed by any company or charity, unless we state otherwise.

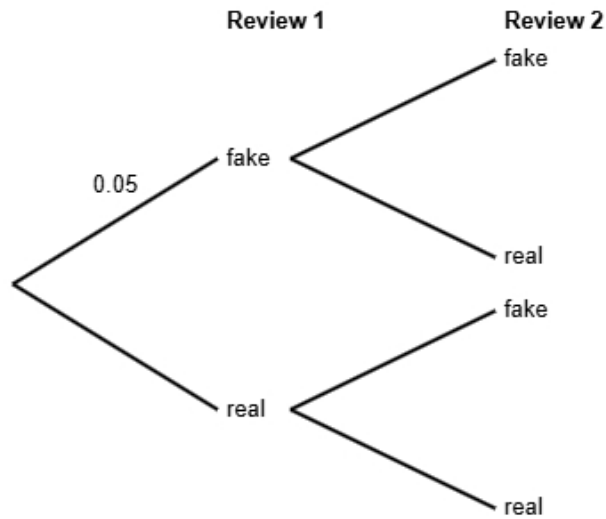
Copyright

- Sharing this PDF is strictly forbidden unless you have the author's permission.
- This paper is only authorised to be used by the person who has bought it.
- You may not store this online or any shared area such as an intranet.
- Please contact us if you have any queries.

Advice

- You can get support for all these questions at our website: www.statsgcse.com
- This paper and more are available on our site with questions that change subtly after each attempt.
- Good luck!

- 1 Research suggests that 5% of online product reviews are fake.
All other reviews are genuine.
Emma is reading two reviews for a product.
She does not know if each review is fake or real.



- (a) Complete the probability tree diagram.

(2 marks)

- (b) Find the probability that both of Emma's reviews are real.

(2 marks)

- (c) Emma states that the probability that exactly one reviews is fake is less than 10%
Find out whether or not Emma is correct.

(3 marks)

2 The table shows information about apartments for rent in Manchester.

| number of rooms | number of apartments |
|-----------------|----------------------|
| 1 | 225 |
| 2 | 180 |
| 3 | 495 |
| 4 | 120 |
| 5 or more | 180 |
| Total | 1200 |

A researcher wants to investigate the price of these apartments and takes a stratified sample of 80 apartments according to the number of rooms.

(a) The researcher says the mode of the number of rooms for these apartments is 3.

Explain how the researcher knows this.

(1 mark)

(b) Work out the number of apartments in the sample for each number of rooms.

| number of rooms | number of apartments in the sample |
|-----------------|------------------------------------|
| 1 | |
| 2 | |
| 3 | |
| 4 | |
| 5 or more | |

(3 marks)

(c) Describe how the 80 apartments in the sample should be selected.

(3 marks)

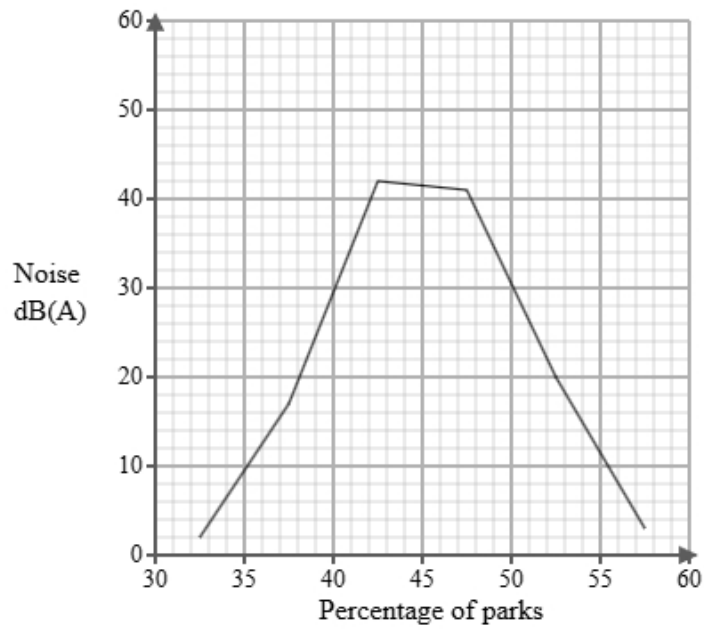
- 3 Luca works for a public health institute. He has been tasked with investigating noise levels in urban parks. Below is a section of the spreadsheet he used to record his findings.

| Noise dB(A) | Percentage of parks |
|------------------|---------------------|
| $30 < n \leq 35$ | 8 |
| $35 < n \leq 40$ | six |
| $40 < n \leq 45$ | 8 |
| $45 < n \leq 50$ | 111 |
| $50 < n \leq 55$ | 57 |
| $55 < n \leq 60$ | 10 |
| Total | 100 |

Luca cleans the data to create the table below.

| Noise dB(A) | Percentage of parks |
|------------------|---------------------|
| $30 < n \leq 35$ | 8 |
| $35 < n \leq 40$ | 6 |
| $40 < n \leq 45$ | 8 |
| $45 < n \leq 50$ | 11 |
| $50 < n \leq 55$ | 57 |
| $55 < n \leq 60$ | 10 |
| Total | 100 |

A frequency polygon has been drawn for noise levels in residential streets.



- i) On the same graph, draw the frequency polygon for noise levels in urban parks.
- ii) Using the two frequency polygons, compare the skew of the distributions and explain what your comparison means in context.

(4 marks)

- 4 The table shows information about the consumer price index (CPI) and average cost of a book (£) in England for Jan 1995, Jan 2005 and Jan 2015.

| | Jan 1995 | Jan 2005 | Jan 2015 |
|----------------------------|-----------------|-----------------|-----------------|
| consumer price index | 100 | 116 | 149 |
| average cost of a book (£) | 4.57 | 5.15 | 7.35 |

Describe how the increase in average cost of a book (£) compares with the CPI over the ten years to Jan 2005 and over the twenty years to Jan 2015.

(5 marks)

5 Tom is investigating how the distance travelled in km, x , affects the resale price (£), y for two types of bicycles, type A and type B.

He found ten bicycles of each type and recorded their distance travelled and resale price and plotted each on scatter diagrams.

He then drew a line of best fit on each diagram and found the gradient and y-intercept of each line.

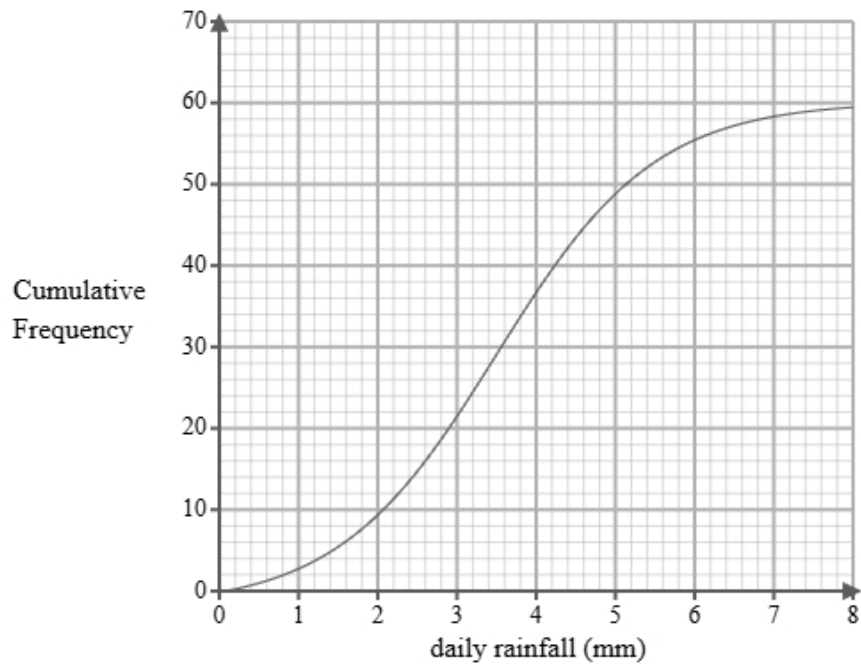
Here are the results:

| Type | Gradient of line of best fit | y-intercept of line of best fit |
|------|------------------------------|---------------------------------|
| A | -15 | 600 |
| B | -10 | 750 |

Interpret and compare these results in context.

(5 marks)

- 6 A researcher measures the daily rainfall, in millimetres, recorded over 60 days in a small town. A cumulative frequency diagram is drawn from the data.



Complete the table below from the cumulative frequency diagram.

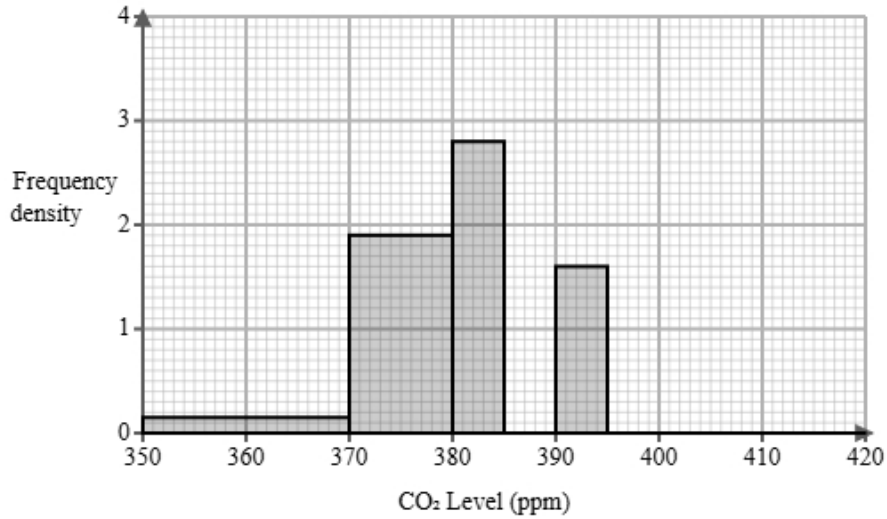
| Lower quartile | Median | Upper quartile |
|----------------|--------|----------------|
| | | |

(2 marks)

7 The concentration of carbon dioxide in the air is recorded in parts per million (ppm).

A scientist is analysing the CO₂ levels in the atmosphere of 60 urban areas during a 12-hour night period with minimal traffic.

The partially completed histogram and grouped frequency table provide details about these CO₂ concentrations.



| CO ₂ Level c (ppm) | Frequency |
|---------------------------------|-----------|
| $350 < c \leq 370$ | 3 |
| $370 < c \leq 380$ | 19 |
| $380 < c \leq 385$ | |
| $385 < c \leq 390$ | 6 |
| $390 < c \leq 395$ | |
| $395 < c \leq 420$ | 10 |

(a) Complete the table using the information from the histogram.

(2 marks)

(b) Complete the histogram using the information from the table.

(2 marks)

(c) The scientist finds the following summary statistics for the data.

$$\sum c = 23091$$

$$\sum c^2 = 8891399$$

$$n = 60$$

Explain whether or not there may be any outliers in the scientist's data by calculating the limits for outliers using the mean and standard deviation.

You must round all values to 2 decimal places.

(5 marks)

(d) A different scientist is analysing the CO₂ levels in the atmosphere of industrial areas during a 12-hour day period.

They find the following summary statistics for the data.

mean = 389.45

median = 412

standard deviation = 11.5

Calculate and interpret the skew for the industrial areas.

You must round your answer to 2 decimal places.

(3 marks)

(e) Find the class interval that contains the 35th percentile.

(1 mark)

- 8 Leo is researching the final league position of basketball teams in a local league and the mean heights of all the players in each team.

The table below shows the data collected.

| Team | Mean Height (cm) | Height Rank | Final Position | d | d ² |
|------------|------------------|-------------|----------------|----|----------------|
| Arundel | 171 | 3 | 5 | -2 | 4 |
| Brighton | 178 | 6 | 7 | -1 | 1 |
| Cambridge | 179 | 7 | 6 | 1 | 1 |
| Derby | 176 | 5 | 3 | 2 | 4 |
| Enfield | 168 | 1 | 2 | | |
| Farnham | 169 | 2 | 1 | | |
| Gloucester | 180 | 8 | 8 | | |
| Hastings | 173 | 4 | 4 | | |

- (a) Leo would like to see if there is an association between the final position and the mean value.

Suggest a diagram that Leo could draw.

(1 mark)

(b) i) Calculate Spearman's rank correlation coefficient from the data in the table and leave your answer to 2 decimal places.

ii) Interpret your answer to **part i**, referring to the effects of any anomalous data.

(5 marks)

(c) Leo used Spearman's rank correlation coefficient to analyse the data.

Amelia suggests that Leo could have used Pearson's product moment correlation coefficient.

Discuss whether using Pearson's product moment correlation coefficient is appropriate for this data.

(3 marks)

- 9 Luca works for a public health institute. He has been tasked with investigating noise levels in urban parks. Below is a section of the spreadsheet he used to record his findings.

| Noise dB(A) | Percentage of parks |
|------------------|---------------------|
| $30 < n \leq 35$ | 8 |
| $35 < n \leq 40$ | six |
| $40 < n \leq 45$ | 8 |
| $45 < n \leq 50$ | 111 |
| $50 < n \leq 55$ | 57 |
| $55 < n \leq 60$ | 10 |
| Total | 100 |

Luca cleans the data to create the table below.

| Noise dB(A) | Percentage of parks |
|------------------|---------------------|
| $30 < n \leq 35$ | 8 |
| $35 < n \leq 40$ | 6 |
| $40 < n \leq 45$ | 8 |
| $45 < n \leq 50$ | 11 |
| $50 < n \leq 55$ | 57 |
| $55 < n \leq 60$ | 10 |
| Total | 100 |

- (a) Give a reason Luca cleaned the data.

(1 mark)

(b) Use linear interpolation to work out an estimate of the median noise level.
Round your answer to one decimal place.

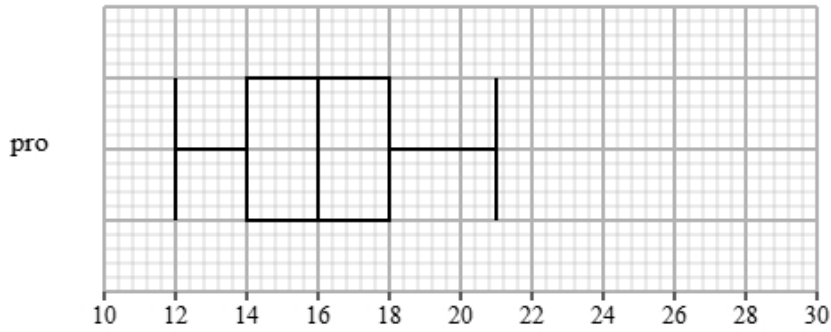
(3 marks)

_____ dB(A)

10 Liam recorded the completion times for pro and beginner runners in a 5K race.

Both groups ran the same course.

The box plot presents data on the completion times for the pro runners.



The table gives information about the completion times for the beginner runners.

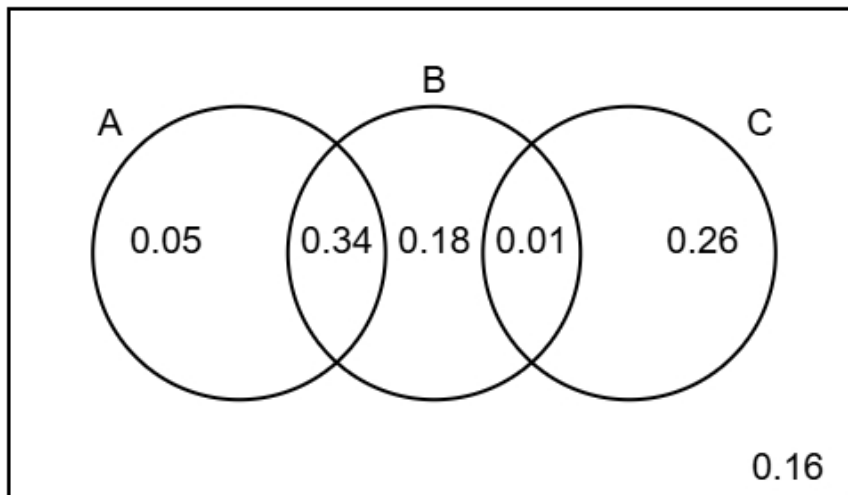
| Least tall | Lower quartile | Median | Upper quartile | Most tall |
|------------|----------------|--------|----------------|-----------|
| 17 | 19 | 20 | 24 | 30 |

Compare the two distributions of completion times.

Give three comparisons and interpret one of these comparisons.

(4 marks)

11 The Venn diagram illustrates the probabilities associated with events A, B, and C.



(a) Identify the **two** events that are mutually exclusive, giving a reason for your answer.

(2 marks)

(b) Find $P(B)$

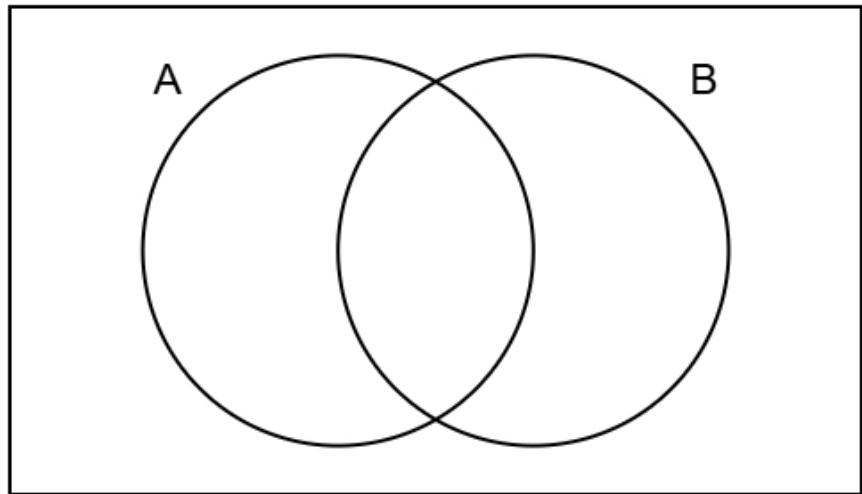
(1 mark)

(c) Find $P(A \text{ or } C)$

(2 marks)

(d) Complete the Venn diagram to show **only** the probabilities for events A and B.

(2 marks)



12 The figures below show the amount, in millions, of tourists who visited Japan between 2012 and 2018

24 26 28 27 29 31 33

The table gives a summary of the amount, in millions, of tourists who visited USA between 2012 and 2018

| Mean | Standard Deviation | Largest Amount |
|------|--------------------|----------------|
| 29 | 2 | 32 |

Compare the amount of tourists in Japan and USA between 2012 and 2018

You may use:

$$24^2 + 26^2 + 28^2 + 27^2 + 29^2 + 31^2 + 33^2 = 5656$$

(5 marks)

13 The reaction times of a group of adult drivers have a mean of 1.4 seconds and a standard deviation of 0.3 seconds.

(a) Mark is adult driver with a standardised score of 0.

Find Mark's reaction time.

(1 mark)

_____ seconds

(b) Priya and Sarah are both adult drivers in the group.

Priya's standardised score for reaction time is 0.8 seconds.

Sarah's standardised score for reaction time is -1.2 seconds.

Priya had a slower reaction time than Sarah.

How much slower is Priya?

(3 marks)

_____ seconds

