# GCSE Higher Mathematics
# Practice Test 9: Statistics

**Instructions:**
Answer all questions. Show your working clearly.
Calculators may be used unless stated otherwise.
Time allowed: 90 minutes

## Section A: Advanced Averages and Spread

1. The table shows the distribution of podcast episode lengths:

| Length (minutes) | Frequency |
|---|---|
| 10-19 | 8 |
| 20-29 | 14 |
| 30-39 | 22 |
| 40-49 | 27 |
| 50-59 | 19 |
| 60-69 | 13 |
| 70-79 | 9 |
| 80-89 | 5 |

   Calculate:

   (a) The total number of episodes analyzed
   (b) An estimate of the mean episode length
   (c) The modal class
   (d) An estimate of the median episode length
   (e) An estimate of the range

2. For the data set: 21, 24, 27, 30, 33, 36, 39, 42, 45, 51

   (a) Calculate the mean
   (b) Find the median and quartiles (Q1 and Q3)
   (c) Calculate the interquartile range
   (d) Calculate the standard deviation
   (e) Identify any outliers using the 1.5 × IQR rule

3. Two data sets have the following properties:

   - Set A: mean = 78, standard deviation = 14, n = 32
   - Set B: mean = 82, standard deviation = 19, n = 28

   (a) Calculate the combined mean

(b) Calculate the combined standard deviation

(c) Which set is more consistent? Explain.

(d) Calculate the coefficient of variation for each set

4. The ages (in years) of 56 conference attendees are summarized:

$$\sum x = 2464, \sum x^2 = 111616$$

(a) Calculate the mean age

(b) Calculate the variance

(c) Calculate the standard deviation

(d) If everyone ages by 2 years, find the new mean and standard deviation

## Section B: Histograms and Frequency Density

5. The histogram shows the distribution of home broadband speeds (Mbps):

*[Imagine a histogram with: 0-25 Mbps (density 2.4), 25-50 Mbps (density 3.8), 50-75 Mbps (density 4.2), 75-125 Mbps (density 2.6), 125-200 Mbps (density 1.2)]*

(a) Complete the frequency table

(b) Calculate the total number of homes tested

(c) Estimate the mean broadband speed

(d) Find the modal class

(e) What percentage of homes have speeds below 50 Mbps?

6. Draw a histogram for this data about weekly exercise hours:

| Exercise hours | Frequency |
|---|---|
| 0-3 | 28 |
| 3-6 | 42 |
| 6-8 | 24 |
| 8-12 | 32 |
| 12-16 | 24 |
| 16-25 | 18 |

(a) Calculate the frequency density for each class

(b) Draw the histogram

(c) Estimate the median exercise hours

(d) What fraction of people exercise more than 8 hours weekly?

7. A histogram shows data with unequal class widths. The class 200-205 has frequency density 36 and the class 205-215 has frequency 240.

(a) Find the frequency for the 200-205 class

(b) Find the frequency density for the 205-215 class

(c) If the total frequency is 1200, suggest frequencies for other classes

## Section C: Cumulative Frequency and Box Plots

8. The table shows the cumulative frequency of daily water consumption:

| Water (liters) | Cumulative Frequency |
|---|---|
| $\leq 1.0$ | 13 |
| $\leq 1.5$ | 29 |
| $\leq 2.0$ | 51 |
| $\leq 2.5$ | 73 |
| $\leq 3.0$ | 89 |
| $\leq 3.5$ | 101 |
| $\leq 4.0$ | 107 |
| $\leq 4.5$ | 110 |

   (a) Draw the cumulative frequency curve

   (b) Find the median

   (c) Find the quartiles Q1 and Q3

   (d) Calculate the interquartile range

   (e) Draw a box plot

   (f) Estimate the 95th percentile

9. Two box plots show the distribution of project completion times for two teams:

   *[Imagine box plots: Team A (min 14, Q1 22, median 30, Q3 38, max 54), Team B (min 16, Q1 24, median 32, Q3 40, max 56)]*

   Compare the distributions by commenting on:

   (a) Central tendency (medians)

   (b) Spread (ranges and IQRs)

   (c) Shape and outliers

   (d) Which team has more variable completion times?

10. The cumulative frequency curve for repair times (in hours) passes through these points: (2, 0), (4, 16), (6, 34), (8, 52), (10, 67), (12, 78), (14, 84)

    (a) Find the median repair time

    (b) Find the quartiles

    (c) What percentage have repair times between 5 and 9 hours?

    (d) Draw the corresponding box plot

## Section D: Scatter Graphs and Correlation

11. The table shows data for 10 tutoring centers:

| Tutors employed | 4 | 7 | 10 | 13 | 16 | 19 | 22 | 25 | 28 | 31 |
|---|---|---|---|---|---|---|---|---|---|---|
| Weekly students | 48 | 78 | 108 | 138 | 168 | 198 | 228 | 258 | 288 | 318 |

    (a) Plot a scatter graph

    (b) Describe the correlation

    (c) Calculate the equation of the line of best fit

    (d) Use your line to predict students for a center with 20 tutors

(e) Estimate the number of tutors needed for 240 weekly students

(f) Calculate the correlation coefficient

12. The equation of a regression line is $y = -1.8x + 210$.

(a) Interpret the gradient

(b) Interpret the y-intercept

(c) If $x = 65$, predict $y$

(d) If $y = 138$, estimate $x$

(e) State assumptions made when using this model

13. Classify these correlation coefficients and describe the relationships:

(a) $r = 0.87$

(b) $r = -0.63$

(c) $r = 0.29$

(d) $r = -0.92$

(e) $r = 0.76$

## Section E: Advanced Probability

14. A bag contains 13 purple tokens, 9 orange tokens, and 8 green tokens. Two tokens are drawn without replacement.

(a) Draw a tree diagram

(b) Find P(both purple)

(c) Find P(both same color)

(d) Find P(at least one orange)

(e) Find P(different colors)

15. The probability that a smart bulb malfunctions on any day is 0.05, independently of other days.

(a) Find the probability it malfunctions on exactly 2 out of 9 days

(b) Find the probability it malfunctions on at least 3 out of 9 days

(c) Find the expected number of malfunctions in 40 days

(d) In 80 days, find P(more than 5 malfunctions)

16. A licensing exam has 24 multiple choice questions, each with 5 options. A candidate guesses randomly.

(a) Find P(correct answer on one question)

(b) Find P(exactly 6 correct answers)

(c) Find P(at least 18 correct answers)

(d) Find the expected number of correct answers

(e) Find the most likely number of correct answers

17. Events S and T are such that P(S) = 0.34, P(T) = 0.56, and P(S  T) = 0.17.

(a) Find P(S  T)

(b) Find P(S')

(c) Find P(S — T)

(d) Find P(T — S)

(e) Are S and T independent? Justify your answer

## Section F: Conditional Probability and Independence

18. A survey of 400 movie watchers about streaming preferences gives:

    - 240 watch action movies
    - 160 watch comedies
    - 80 watch both genres

    (a) Draw a Venn diagram
    (b) Find P(watches action — watches comedy)
    (c) Find P(watches comedy — watches action)
    (d) Find P(watches exactly one genre)
    (e) Are the movie preferences independent? Explain

19. In a textile factory, 65% of products are inspected by Robot A and 35% by Robot B. Robot A detects 97% of defects, Robot B detects 93% of defects.

    (a) Draw a tree diagram
    (b) Find the probability a defect is detected
    (c) If a defect is detected, find the probability it was found by Robot A
    (d) If a defect is missed, find the probability it was missed by Robot B

20. A card collection has 26 rare cards numbered 1-26 and 20 common cards numbered 1-20. A card is selected at random.

    (a) Find P(rare card and divisible by 7)
    (b) Find P(common card — perfect square)
    (c) Find P(number ¿ 18)
    (d) Are card rarity and number characteristics independent?

21. A plagiarism detector is 89% accurate for original work and 95% accurate for plagiarized work. 7% of submissions are plagiarized.

    (a) Find the probability a submission is flagged as plagiarized
    (b) If a submission is flagged, find the probability it is actually plagiarized
    (c) If a submission passes, find the probability it is original
    (d) Comment on the effectiveness of the plagiarism detector

## Section G: Hypothesis Testing and Sampling

22. A lottery machine is suspected of bias towards certain numbers. Out of 36 draws, number 7 appears 8 times.

    (a) State the null and alternative hypotheses (expected frequency for fair machine is 1/10)
    (b) Calculate the probability of getting 8 or more occurrences if fair
    (c) At the 5% significance level, is there evidence of bias?
    (d) What would be a Type I error in this context?

23. A sample of 58 electric vehicle batteries has mean range 285 km and standard deviation 42 km.

    (a) Calculate a 95% confidence interval for the population mean
    (b) Interpret your confidence interval

(c) What assumptions are made?

(d) How would the interval change with a 99% confidence level?

24. A video platform claims 82% viewer retention. In a sample of 175 videos, 138 maintain viewer retention.

(a) Test at 5% level whether the claim is justified

(b) Calculate the critical value

(c) State your conclusion

(d) What is the p-value for this test?

## Section H: Problem Solving and Integration

25. A pharmacy records prescription fulfillment times over 300 orders. The data shows:

- Mean = 12.5 minutes
- Standard deviation = 3.2 minutes
- Distribution is approximately normal

(a) Find P(fulfillment time ¿ 18 minutes)

(b) Find the time exceeded by only 25% of orders

(c) What percentage of orders are fulfilled between 8 and 16 minutes?

(d) If the pharmacy processes 450 orders daily, estimate urgent orders (assuming urgent when ¿ 15 minutes)

26. A quality assurance team samples 18 products every 4 hours. Over 7 sampling periods, the number of non-conforming products found was: 1, 3, 0, 2, 1, 4, 2.

(a) Calculate the mean and standard deviation

(b) Test whether the non-conformance rate exceeds 8%

(c) Create a control chart with warning limits

(d) Comment on process stability

27. Compare these three cryptocurrency investment options over 2 years:

- Crypto A: Mean return 18%, standard deviation 24%
- Crypto B: Mean return 25%, standard deviation 35%
- Crypto C: Mean return 12%, standard deviation 16%

(a) Calculate the coefficient of variation for each

(b) Which offers the best risk-adjusted return?

(c) Using normal distribution, find P(loss ¿ 10%) for each option

(d) Recommend an option for a high-risk investor

28. A language learning app studies the relationship between daily usage and proficiency improvement. The correlation is 0.69.

(a) What does this correlation suggest?

(b) If daily usage has mean 45 minutes and standard deviation 18, and proficiency improvement has mean 22 points and standard deviation 9, find the regression equation

(c) Predict proficiency improvement for someone using the app 60 minutes daily

(d) Calculate the coefficient of determination and interpret it

29. Design a statistical investigation to test whether remote work affects team collaboration effectiveness:

(a) State hypotheses

(b) Describe the sampling method

(c) Identify variables and potential confounding factors

(d) Outline the analysis plan

(e) Discuss limitations and assumptions

**Answer Space**

Use this space for your working and answers.

**END OF TEST**

Total marks: 100

**For more resources and practice materials, visit:**
**stepupmaths.co.uk**