

GCSE Higher Mathematics

Practice Test 7: 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 weekly shopping costs:

Cost (£)	Frequency
20-29	9
30-39	16
40-49	23
50-59	28
60-69	20
70-79	14
80-89	8
90-99	5

Calculate:

- The total number of shoppers surveyed
 - An estimate of the mean weekly cost
 - The modal class
 - An estimate of the median weekly cost
 - An estimate of the range
2. For the data set: 26, 29, 32, 35, 38, 41, 44, 47, 50, 56
- Calculate the mean
 - Find the median and quartiles (Q1 and Q3)
 - Calculate the interquartile range
 - Calculate the standard deviation
 - Identify any outliers using the $1.5 \times \text{IQR}$ rule
3. Two data sets have the following properties:
- Set V: mean = 94, standard deviation = 18, $n = 22$
 - Set W: mean = 86, standard deviation = 24, $n = 38$
- 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 fuel consumption (liters per 100km) of 45 vehicles is summarized:

$$\sum x = 360, \sum x^2 = 2916$$

- (a) Calculate the mean fuel consumption
- (b) Calculate the variance
- (c) Calculate the standard deviation
- (d) If each vehicle's consumption increases by 1.5 L/100km, find the new mean and standard deviation

Section B: Histograms and Frequency Density

5. The histogram shows the distribution of video streaming quality (Mbps):
[Imagine a histogram with: 0-5 Mbps (density 3.4), 5-10 Mbps (density 5.2), 10-15 Mbps (density 4.6), 15-25 Mbps (density 2.4), 25-40 Mbps (density 1.2)]
- (a) Complete the frequency table
 - (b) Calculate the total number of users tested
 - (c) Estimate the mean streaming quality
 - (d) Find the modal class
 - (e) What percentage of users get quality below 10 Mbps?
6. Draw a histogram for this data about employee overtime hours:

Overtime (hours)	Frequency
0-5	32
5-10	48
10-12	22
12-16	28
16-20	20
20-30	16

- (a) Calculate the frequency density for each class
 - (b) Draw the histogram
 - (c) Estimate the median overtime hours
 - (d) What fraction of employees work more than 12 hours overtime?
7. A histogram shows data with unequal class widths. The class 100-105 has frequency density 28 and the class 105-115 has frequency 160.
- (a) Find the frequency for the 100-105 class
 - (b) Find the frequency density for the 105-115 class
 - (c) If the total frequency is 800, suggest frequencies for other classes

Section C: Cumulative Frequency and Box Plots

8. The table shows the cumulative frequency of train journey times:

Time (minutes)	Cumulative Frequency
≤ 30	11
≤ 45	26
≤ 60	44
≤ 75	64
≤ 90	78
≤ 105	88
≤ 120	94
≤ 135	97

- Draw the cumulative frequency curve
- Find the median
- Find the quartiles Q1 and Q3
- Calculate the interquartile range
- Draw a box plot
- Estimate the 85th percentile

9. Two box plots show the distribution of test scores for two classes:

[Imagine box plots: Class A (min 42, Q1 58, median 68, Q3 76, max 92), Class B (min 38, Q1 54, median 64, Q3 74, max 88)]

Compare the distributions by commenting on:

- Central tendency (medians)
 - Spread (ranges and IQRs)
 - Shape and outliers
 - Which class has more variable test scores?
10. The cumulative frequency curve for assembly times (in seconds) passes through these points: (120, 0), (140, 15), (160, 32), (180, 49), (200, 63), (220, 74), (240, 80)
- Find the median assembly time
 - Find the quartiles
 - What percentage have assembly times between 150 and 210 seconds?
 - Draw the corresponding box plot

Section D: Scatter Graphs and Correlation

11. The table shows data for 10 coffee shops:

Distance from station (m)	50	120	190	260	330	400	470	540	610	680
Daily customers	185	168	151	134	117	100	83	66	49	32

- Plot a scatter graph
- Describe the correlation
- Calculate the equation of the line of best fit
- Use your line to predict customers for a shop 350m from station

- (e) Estimate the distance for a shop with 90 daily customers
 - (f) Calculate the correlation coefficient
12. The equation of a regression line is $y = 0.6x + 18$.
- (a) Interpret the gradient
 - (b) Interpret the y-intercept
 - (c) If $x = 85$, predict y
 - (d) If $y = 96$, estimate x
 - (e) State assumptions made when using this model
13. Classify these correlation coefficients and describe the relationships:
- (a) $r = 0.88$
 - (b) $r = -0.69$
 - (c) $r = 0.24$
 - (d) $r = -0.91$
 - (e) $r = 0.73$

Section E: Advanced Probability

14. A lottery uses 12 red balls, 9 blue balls, and 8 yellow balls. Two balls are drawn without replacement.
- (a) Draw a tree diagram
 - (b) Find $P(\text{both red})$
 - (c) Find $P(\text{both same color})$
 - (d) Find $P(\text{at least one blue})$
 - (e) Find $P(\text{different colors})$
15. The probability that a sensor fails on any day is 0.09, independently of other days.
- (a) Find the probability it fails on exactly 2 out of 6 days
 - (b) Find the probability it fails on at least 3 out of 6 days
 - (c) Find the expected number of failures in 20 days
 - (d) In a 40-day period, find $P(\text{more than 5 failures})$
16. A certification exam has 16 true/false questions. A candidate guesses randomly.
- (a) Find $P(\text{correct answer on one question})$
 - (b) Find $P(\text{exactly 10 correct answers})$
 - (c) Find $P(\text{at least 13 correct answers})$
 - (d) Find the expected number of correct answers
 - (e) Find the most likely number of correct answers
17. Events N and P are such that $P(N) = 0.52$, $P(P) = 0.48$, and $P(N \cap P) = 0.26$.
- (a) Find $P(N \cap P)$
 - (b) Find $P(N')$
 - (c) Find $P(N \cup P)$
 - (d) Find $P(P \cap N)$
 - (e) Are N and P independent? Justify your answer

Section F: Conditional Probability and Independence

18. A survey of 280 gamers about platform preferences gives:
- 168 play on console
 - 112 play on PC
 - 56 play on both platforms
- (a) Draw a Venn diagram
- (b) Find $P(\text{plays console} \mid \text{plays PC})$
- (c) Find $P(\text{plays PC} \mid \text{plays console})$
- (d) Find $P(\text{plays on exactly one platform})$
- (e) Are the platform preferences independent? Explain
19. In a pharmaceutical lab, 85% of samples are analyzed by Method A and 15% by Method B. Method A has 98% accuracy, Method B has 94% accuracy.
- (a) Draw a tree diagram
- (b) Find the probability a result is accurate
- (c) If a result is accurate, find the probability it used Method A
- (d) If a result is inaccurate, find the probability it used Method B
20. A card game uses 20 spell cards numbered 1-20 and 16 creature cards numbered 1-16. A card is drawn at random.
- (a) Find $P(\text{spell card and divisible by 5})$
- (b) Find $P(\text{creature card} \mid \text{prime number})$
- (c) Find $P(\text{number} < 15)$
- (d) Are card type and number divisibility independent?
21. A fraud detection system is 93% accurate for legitimate transactions and 89% accurate for fraudulent transactions. 6% of transactions are fraudulent.
- (a) Find the probability a transaction is flagged as fraudulent
- (b) If a transaction is flagged, find the probability it is actually fraudulent
- (c) If a transaction passes, find the probability it is legitimate
- (d) Comment on the effectiveness of the fraud detection system

Section G: Hypothesis Testing and Sampling

22. A production machine is suspected of producing oversized items. Out of 32 items tested, 11 are oversized.
- (a) State the null and alternative hypotheses (expected rate is 25%)
- (b) Calculate the probability of getting 11 or more oversized items if rate is 25%
- (c) At the 5% significance level, is there evidence of a problem?
- (d) What would be a Type I error in this context?
23. A sample of 48 tablet computers has mean battery life 9.2 hours and standard deviation 1.8 hours.
- (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 smaller confidence level?
24. A social media platform claims 78% user engagement. In a sample of 160 posts, 118 receive engagement.
- (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 hospital records patient discharge times over 200 cases. The data shows:
- Mean = 6.8 days
 - Standard deviation = 2.2 days
 - Distribution is approximately normal
- (a) Find $P(\text{discharge time} < 10 \text{ days})$
 - (b) Find the discharge time exceeded by only 20% of patients
 - (c) What percentage of patients are discharged between 4 and 8 days?
 - (d) If the hospital admits 150 patients weekly, estimate long-stay cases (assuming long-stay is > 9 days)
26. A quality control team samples 14 products every 2 hours. Over 8 sampling periods, the number of faulty products found was: 1, 0, 2, 1, 3, 0, 2, 1.
- (a) Calculate the mean and standard deviation
 - (b) Test whether the fault rate exceeds 7%
 - (c) Create a control chart with warning limits
 - (d) Comment on process stability
27. Compare these three pension fund options over 25 years:
- Fund P: Mean return 5.5%, standard deviation 3.8%
 - Fund Q: Mean return 8.2%, standard deviation 6.4%
 - Fund R: Mean return 3.9%, standard deviation 2.1%
- (a) Calculate the coefficient of variation for each
 - (b) Which offers the best risk-adjusted return?
 - (c) Using normal distribution, find $P(\text{return} < 1\%)$ for each fund
 - (d) Recommend an option for someone near retirement
28. A tutoring center studies the relationship between tutoring sessions and grade improvement. The correlation is 0.71.
- (a) What does this correlation suggest?
 - (b) If tutoring sessions have mean 12 and standard deviation 4.5, and grade improvement has mean 18 points and standard deviation 8, find the regression equation

- (c) Predict grade improvement for a student attending 16 tutoring sessions
 - (d) Calculate the coefficient of determination and interpret it
29. Design a statistical investigation to test whether a new packaging design increases sales:
- (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**