

GCSE Higher Mathematics

Practice Test 6: 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 daily step counts:

Steps (thousands)	Frequency
2-4	8
5-7	15
8-10	21
11-13	26
14-16	18
17-19	12
20-22	7
23-25	3

Calculate:

- The total number of people surveyed
 - An estimate of the mean daily steps (in thousands)
 - The modal class
 - An estimate of the median daily steps
 - An estimate of the range
2. For the data set: 18, 21, 24, 27, 30, 33, 36, 39, 42, 48
- 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 T: mean = 84, standard deviation = 16, $n = 18$
 - Set U: mean = 76, standard deviation = 22, $n = 42$
- 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 reaction times (in milliseconds) of 54 participants are summarized:

$$\sum x = 14040, \sum x^2 = 3686400$$

- (a) Calculate the mean reaction time
- (b) Calculate the variance
- (c) Calculate the standard deviation
- (d) If each reaction time is reduced by 12ms, find the new mean and standard deviation

Section B: Histograms and Frequency Density

5. The histogram shows the distribution of daily energy usage:
- [Imagine a histogram with: 0-10 kWh (density 2.8), 10-20 kWh (density 4.5), 20-30 kWh (density 3.2), 30-45 kWh (density 1.8), 45-60 kWh (density 1.0)]*
- (a) Complete the frequency table
 - (b) Calculate the total number of households
 - (c) Estimate the mean energy usage
 - (d) Find the modal class
 - (e) What percentage of households use more than 30 kWh daily?
6. Draw a histogram for this data about exam scores:

Score (%)	Frequency
30-40	14
40-50	28
50-55	20
55-65	35
65-75	30
75-90	24

- (a) Calculate the frequency density for each class
 - (b) Draw the histogram
 - (c) Estimate the median exam score
 - (d) What fraction of students scored above 65%?
7. A histogram shows data with unequal class widths. The class 80-85 has frequency density 24 and the class 85-95 has frequency 140.
- (a) Find the frequency for the 80-85 class
 - (b) Find the frequency density for the 85-95 class
 - (c) If the total frequency is 600, suggest frequencies for other classes

Section C: Cumulative Frequency and Box Plots

8. The table shows the cumulative frequency of monthly internet usage:

Usage (GB)	Cumulative Frequency
≤ 20	12
≤ 40	28
≤ 60	48
≤ 80	70
≤ 100	86
≤ 120	96
≤ 140	102
≤ 160	105

- Draw the cumulative frequency curve
 - Find the median
 - Find the quartiles Q1 and Q3
 - Calculate the interquartile range
 - Draw a box plot
 - Estimate the 10th percentile
9. Two box plots show the distribution of house prices in two neighborhoods:
[Imagine box plots: Area A (min 180000, Q1 220000, median 260000, Q3 300000, max 380000), Area B (min 160000, Q1 200000, median 240000, Q3 280000, max 360000)]
 Compare the distributions by commenting on:
- Central tendency (medians)
 - Spread (ranges and IQRs)
 - Shape and outliers
 - Which area has more variable house prices?
10. The cumulative frequency curve for call durations (in minutes) passes through these points: (1, 0), (3, 18), (5, 36), (7, 54), (9, 68), (11, 78), (13, 84)
- Find the median call duration
 - Find the quartiles
 - What percentage have call durations between 4 and 8 minutes?
 - Draw the corresponding box plot

Section D: Scatter Graphs and Correlation

11. The table shows data for 10 online stores:

Advertising spend (£000s)	5	8	11	14	17	20	23	26	29	32
Monthly sales (£000s)	45	62	78	95	111	128	144	161	177	194

- Plot a scatter graph
- Describe the correlation
- Calculate the equation of the line of best fit
- Use your line to predict sales for £19,000 advertising spend

- (e) Estimate the advertising spend needed for £150,000 sales
 - (f) Calculate the correlation coefficient
12. The equation of a regression line is $y = 4.5x - 28$.
- (a) Interpret the gradient
 - (b) Interpret the y-intercept
 - (c) If $x = 22$, predict y
 - (d) If $y = 139$, estimate x
 - (e) State assumptions made when using this model
13. Classify these correlation coefficients and describe the relationships:
- (a) $r = 0.84$
 - (b) $r = -0.72$
 - (c) $r = 0.19$
 - (d) $r = -0.96$
 - (e) $r = 0.61$

Section E: Advanced Probability

14. A container holds 10 white balls, 8 black balls, and 7 gray balls. Two balls are drawn without replacement.
- (a) Draw a tree diagram
 - (b) Find $P(\text{both white})$
 - (c) Find $P(\text{both same color})$
 - (d) Find $P(\text{at least one black})$
 - (e) Find $P(\text{different colors})$
15. The probability that a printer jams on any print job is 0.04, independently of other jobs.
- (a) Find the probability it jams on exactly 1 out of 10 jobs
 - (b) Find the probability it jams on at least 2 out of 10 jobs
 - (c) Find the expected number of jams in 50 jobs
 - (d) In 100 print jobs, find $P(\text{more than 6 jams})$
16. A safety test has 18 multiple choice questions, each with 4 options. A trainee guesses randomly.
- (a) Find $P(\text{correct answer on one question})$
 - (b) Find $P(\text{exactly 6 correct answers})$
 - (c) Find $P(\text{at least 12 correct answers})$
 - (d) Find the expected number of correct answers
 - (e) Find the most likely number of correct answers
17. Events L and M are such that $P(L) = 0.38$, $P(M) = 0.64$, and $P(L \cap M) = 0.19$.
- (a) Find $P(L \cap M)$
 - (b) Find $P(L')$
 - (c) Find $P(L \cup M)$
 - (d) Find $P(M \cap L)$
 - (e) Are L and M independent? Justify your answer

Section F: Conditional Probability and Independence

18. A survey of 320 employees about flexible working gives:
- 192 work from home
 - 128 have flexible hours
 - 64 have both arrangements
- (a) Draw a Venn diagram
- (b) Find $P(\text{works from home} \mid \text{has flexible hours})$
- (c) Find $P(\text{has flexible hours} \mid \text{works from home})$
- (d) Find $P(\text{has exactly one flexible arrangement})$
- (e) Are the work arrangements independent? Explain
19. In an electronics factory, 80% of components are tested by Scanner 1 and 20% by Scanner 2. Scanner 1 detects 96% of faults, Scanner 2 detects 92% of faults.
- (a) Draw a tree diagram
- (b) Find the probability a fault is detected
- (c) If a fault is detected, find the probability it was found by Scanner 1
- (d) If a fault is missed, find the probability it was missed by Scanner 2
20. A game uses 18 action cards numbered 1-18 and 15 power cards numbered 1-15. A card is drawn at random.
- (a) Find $P(\text{action card and multiple of 4})$
- (b) Find $P(\text{power card} \mid \text{odd number})$
- (c) Find $P(\text{number} < 12)$
- (d) Are card type and number value independent?
21. A biometric scanner is 97% accurate for authorized users and 99% accurate for unauthorized users. 12% of scan attempts are unauthorized.
- (a) Find the probability of access being denied
- (b) If access is denied, find the probability the user was unauthorized
- (c) If access is granted, find the probability the user was authorized
- (d) Comment on the effectiveness of the biometric scanner

Section G: Hypothesis Testing and Sampling

22. A traffic light is suspected of having longer red phases than designed. Out of 28 cycles observed, 18 had red phases longer than expected.
- (a) State the null and alternative hypotheses (expected proportion is 50%)
- (b) Calculate the probability of getting 18 or more long red phases if designed correctly
- (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 52 smartphone screens has mean brightness 420 nits and standard deviation 68 nits.
- (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 98% confidence level?
24. A streaming service claims 85% content availability. In a sample of 180 requests, 147 find available content.
- (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

1. A bank records customer wait times over 240 visits. The data shows:
 - Mean = 8.5 minutes
 - Standard deviation = 2.8 minutes
 - Distribution is approximately normal
 - (a) Find $P(\text{wait time} > 12 \text{ minutes})$
 - (b) Find the wait time exceeded by only 5% of customers
 - (c) What percentage of customers wait between 6 and 10 minutes?
 - (d) If the bank serves 600 customers daily, estimate complaints (assuming complaints when wait > 11 minutes)
2. A production line samples 6 units every hour. Over 10 hours, the number of defective units found was: 0, 1, 0, 2, 1, 0, 1, 3, 0, 1.
 - (a) Calculate the mean and standard deviation
 - (b) Test whether the defect rate exceeds 9%
 - (c) Create a control chart with warning limits
 - (d) Comment on process stability
3. Compare these three bond investment options over 10 years:
 - Bond X: Mean return 4.2%, standard deviation 2.1%
 - Bond Y: Mean return 6.8%, standard deviation 4.2%
 - Bond Z: Mean return 3.1%, standard deviation 1.5%
 - (a) Calculate the coefficient of variation for each
 - (b) Which offers the best risk-adjusted return?
 - (c) Using normal distribution, find $P(\text{return} < 2\%)$ for each bond
 - (d) Recommend an option for a risk-averse investor
4. A driving school studies the relationship between lesson hours and test pass rate. The correlation is 0.67.
 - (a) What does this correlation suggest?
 - (b) If lesson hours have mean 28 and standard deviation 8, and pass rate has mean 0.75 and standard deviation 0.18, find the regression equation
 - (c) Predict the pass rate for someone taking 35 lesson hours

- (d) Calculate the coefficient of determination and interpret it
- 5. Design a statistical investigation to compare customer satisfaction between two service centers:
 - (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**