GCSE Higher Mathematics
Practice Test 5: 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 movie ratings:

Rating (out of 10)	Frequency
2.0-2.9	7
3.0-3.9	12
4.0-4.9	18
5.0-5.9	24
6.0-6.9	22
7.0-7.9	16
8.0-8.9	10
9.0-9.9	4

Calculate:

(a) The total number of movies rated
(b) An estimate of the mean rating
(c) The modal class
(d) An estimate of the median rating
(e) An estimate of the range

2. For the data set: 13, 16, 19, 22, 25, 28, 31, 34, 37, 43

(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 $1.5 \times$ IQR rule

3. Two data sets have the following properties:

- Set R: mean = 72, standard deviation = 13, n = 24
- Set S: mean = 68, standard deviation = 18, n = 36

(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 distances (in km) traveled by 42 delivery vehicles are summarized:

$$\sum x = 5040, \sum x^2 = 615600$$

(a) Calculate the mean distance

(b) Calculate the variance

(c) Calculate the standard deviation

(d) If each vehicle travels an additional 8 km, find the new mean and standard deviation

Section B: Histograms and Frequency Density

5. The histogram shows the distribution of website loading times:

[Imagine a histogram with: 0-1 sec (density 5.4), 1-2 sec (density 3.8), 2-3 sec (density 2.6), 3-5 sec (density 1.2), 5-8 sec (density 0.7)]

(a) Complete the frequency table

(b) Calculate the total number of websites tested

(c) Estimate the mean loading time

(d) Find the modal class

(e) What percentage of websites load in under 2 seconds?

6. Draw a histogram for this data about student absences per term:

Absences	Frequency
0-2	36
3-5	42
6-7	18
8-10	21
11-15	20
16-25	15

(a) Calculate the frequency density for each class

(b) Draw the histogram

(c) Estimate the median number of absences

(d) What fraction of students have more than 7 absences?

7. A histogram shows data with unequal class widths. The class 60-65 has frequency density 18 and the class 65-75 has frequency 120.

(a) Find the frequency for the 60-65 class

(b) Find the frequency density for the 65-75 class

(c) If the total frequency is 500, suggest frequencies for other classes

Section C: Cumulative Frequency and Box Plots

8. The table shows the cumulative frequency of rainfall amounts:

Rainfall (mm)	Cumulative Frequency
≤ 5	9
≤ 10	24
≤ 15	42
≤ 20	61
≤ 25	76
≤ 30	87
≤ 35	94
≤ 40	98

   (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 weekly earnings for two professions:

   [Imagine box plots: Profession A (min 420, Q1 520, median 620, Q3 720, max 920), Profession B (min 380, Q1 480, median 580, Q3 680, max 880)]

   Compare the distributions by commenting on:

   (a) Central tendency (medians)

   (b) Spread (ranges and IQRs)

   (c) Shape and outliers

   (d) Which profession has more variable earnings?

10. The cumulative frequency curve for processing times (in milliseconds) passes through these points: (50, 0), (75, 16), (100, 34), (125, 52), (150, 67), (175, 78), (200, 84)

    (a) Find the median processing time

    (b) Find the quartiles

    (c) What percentage have processing times between 90ms and 160ms?

    (d) Draw the corresponding box plot

Section D: Scatter Graphs and Correlation

11. The table shows data for 10 restaurants:

Tables	12	18	24	30	36	42	48	54	60	66
Weekly revenue (£000s)	8.5	12.2	15.8	19.5	23.1	26.8	30.4	34.1	37.7	41.4

    (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 revenue for a 45-table restaurant

(e) Estimate the number of tables for £22,000 weekly revenue

(f) Calculate the correlation coefficient

12. The equation of a regression line is $y = -0.8x + 150$.

(a) Interpret the gradient

(b) Interpret the y-intercept

(c) If $x = 75$, predict y

(d) If $y = 102$, estimate x

(e) State assumptions made when using this model

13. Classify these correlation coefficients and describe the relationships:

(a) $r = 0.91$

(b) $r = -0.64$

(c) $r = 0.27$

(d) $r = -0.89$

(e) $r = 0.78$

Section E: Advanced Probability

14. A bag contains 9 silver coins, 7 gold coins, and 6 bronze coins. Two coins are drawn without replacement.

(a) Draw a tree diagram

(b) Find P(both silver)

(c) Find P(both same metal)

(d) Find P(at least one gold)

(e) Find P(different metals)

15. The probability that a server crashes on any day is 0.06, independently of other days.

(a) Find the probability it crashes on exactly 2 out of 8 days

(b) Find the probability it crashes on at least 3 out of 8 days

(c) Find the expected number of crashes in 15 days

(d) In a 60-day period, find P(more than 5 crashes)

16. A security quiz has 25 yes/no questions. An employee guesses randomly.

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

(b) Find P(exactly 15 correct answers)

(c) Find P(at least 20 correct answers)

(d) Find the expected number of correct answers

(e) Find the most likely number of correct answers

17. Events J and K are such that P(J) = 0.42, P(K) = 0.58, and P(J K) = 0.21.

(a) Find P(J K)

(b) Find P(J')

(c) Find P(J — K)

(d) Find P(K — J)

(e) Are J and K independent? Justify your answer

Section F: Conditional Probability and Independence

18. A survey of 250 drivers about vehicle types gives:

    - 140 drive petrol cars
    - 90 drive electric cars
    - 45 drive both types

    (a) Draw a Venn diagram
    (b) Find P(drives petrol — drives electric)
    (c) Find P(drives electric — drives petrol)
    (d) Find P(drives exactly one type)
    (e) Are the vehicle types independent? Explain

19. In a textile mill, 75% of fabric is woven by Loom A and 25% by Loom B. Loom A produces 1.5% defective fabric, Loom B produces 4% defective fabric.

    (a) Draw a tree diagram
    (b) Find the probability fabric is defective
    (c) If fabric is defective, find the probability it came from Loom A
    (d) If fabric is not defective, find the probability it came from Loom B

20. A library has 16 hardcover books numbered 1-16 and 12 paperback books numbered 1-12. A book is selected at random.

    (a) Find P(hardcover and perfect square number)
    (b) Find P(paperback — number greater than 10)
    (c) Find P(number is prime)
    (d) Are book format and number type independent?

21. A quality test is 91% accurate for approved items and 96% accurate for rejected items. 8% of items are rejected.

    (a) Find the probability of a failed test result
    (b) If an item fails the test, find the probability it should be rejected
    (c) If an item passes the test, find the probability it should be approved
    (d) Comment on the effectiveness of the quality test

Section G: Hypothesis Testing and Sampling

22. A vending machine is suspected of giving incorrect change. Out of 35 transactions tested, 8 give incorrect change.

    (a) State the null and alternative hypotheses (expected rate is 15%)
    (b) Calculate the probability of getting 8 or more incorrect changes if the rate is 15%
    (c) At the 5% significance level, is there evidence the machine has problems?
    (d) What would be a Type I error in this context?

23. A sample of 38 car tires has mean tread depth 6.8mm and standard deviation 1.4mm.

    (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 larger standard deviation?

24. A hotel claims 92% guest satisfaction. In a sample of 150 guests, 134 report satisfaction.

(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 library records book checkout durations over 200 transactions. The data shows:

- Mean = 14 days
- Standard deviation = 5 days
- Distribution is approximately normal

(a) Find P(checkout duration ¿ 22 days)

(b) Find the duration exceeded by only 12% of checkouts

(c) What percentage of books are returned between 8 and 18 days?

(d) If the library processes 500 checkouts weekly, estimate overdue books (assuming overdue after 21 days)

26. A packaging line samples 8 items every 45 minutes. Over 12 sampling periods, the number of damaged items found was: 0, 1, 2, 0, 1, 3, 1, 0, 2, 1, 0, 1.

(a) Calculate the mean and standard deviation

(b) Test whether the damage rate exceeds 10%

(c) Create a control chart with warning limits

(d) Comment on process stability

27. Compare these three savings account options over 5 years:

- Account A: Mean return 3.5%, standard deviation 1.2%
- Account B: Mean return 5.8%, standard deviation 2.8%
- Account C: Mean return 2.1%, standard deviation 0.8%

(a) Calculate the coefficient of variation for each

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

(c) Using normal distribution, find P(negative return) for each account

(d) Recommend an option for a very conservative saver

28. A fitness center studies the relationship between workout frequency and weight loss. The correlation is 0.73.

(a) What does this correlation suggest?

(b) If workout frequency has mean 4.2 sessions/week and standard deviation 1.8, and weight loss has mean 2.8 kg/month and standard deviation 1.5, find the regression equation

(c) Predict weight loss for someone working out 6 times per week

(d) Calculate the coefficient of determination and interpret it

29. Design a statistical investigation to test whether a new medication reduces recovery time:

    (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