

The Big 50 Revision Guidelines for S1

If you can understand all of these you'll do very well...

1. Know what is meant by a statistical model and the “Modelling cycle” of continuous refinement
2. Understand how inferences from a suitably representative sample can be used to study the wider population
3. Name at least three sampling techniques and describe the relative advantages and disadvantages of each in terms of practicality, economy and reliability
4. Know how to categorise data types according to labels such as discrete, continuous, qualitative and quantitative and be able to give real-life examples of each
5. Understand the distinction between Upper/Lower Class Boundaries and Upper/Lower Class Limits, and why the Upper Class Boundary is used in cumulative frequency graphs
6. Understand that graphical representation can result in loss of data, especially through over-simplification
7. Know how to construct a cumulative frequency step polygon
8. Know how to construct a single or double sided Stem and Leaf diagram
9. Know a variety of techniques for the display of data and understand that according to context some are more effective than others
10. Understand the necessary calculations of frequency density and area required for the construction and interpretation of variable-width histograms
11. Understand the distinction between Measures of Central Tendency (Location) and Measures of Dispersion (Spread) and give examples of each
12. Know how to interpret the Mode, Median and Mean of a set of data, and how to estimate these from a grouped frequency chart, including interpolation as necessary

13. Understand what is meant by the Coding of data, and the implications on the calculation of the mean of the original data set
14. Appreciate the conventions used for the determination of Upper and Lower Quartiles from a small set of discrete data ($n < 50$) when $n/4$ is or is not an integer
15. Understand and use the formula to find the r^{th} of s quantiles within a given class, given the lower class boundary b , the total frequency n , the cumulative frequency f up to b , the class frequency c and the class width w
16. Appreciate the assumptions made when interpolating or extrapolating data
17. Know how to find the quartiles, both from a cumulative frequency graph and from a grouped frequency chart, and to use them to describe the skewness of a data set
18. Know how to calculate a measure of skewness using the mean, median and s.d.
19. Know how to calculate the variance and standard deviation of a set of data, and how to do this efficiently on a scientific calculator
20. Understand the notation used to describe overlapping and disjoint sets (Venn Diagram notation) and the graphical interpretation of Union, Intersection and Complement
21. Understand the simplifications made for probability work at KS3 and GCSE, and how these are extended into more general contexts at KS5
22. Know the meaning of Dependent, Independent, Mutually Exclusive and Conditional in the context of probability
23. Understand how to apply the formula to calculate the probability of one event given the probability of another
24. Understand and use the notation $P(A)$, $P(A')$, $P(A \cup B)$, $P(A \cap B)$ and $P(A | B)$
25. Given $P(C)$, $P(S | C)$ and $P(S | C')$, know how to calculate (for example) $P(S \cap C')$, $P(S)$ and $P(C | S)$, and be able to give both a graphical illustration and a real-life context for these calculations

26. Understand why $P(A | B) = P(B | A) = 0$ for mutually exclusive events A and B .
27. Understand why $P(A | C) = P(A)$ and $P(C | A) = P(C)$ implies that events A and C are independent
28. Understand and use the Addition and Multiplication Rules for probability
29. Know how to use probability trees and Venn Diagrams to solve probability problems
30. Know how to use the factorial function $n!$ to calculate number of arrangements, and to use the formula $\binom{n}{r}$ or ${}^n C_r$ to deal with repeated items
31. Understand and use a Probability Distribution Function $P(X = x)$ for a discrete random variable and appreciate why for all such PDFs, $\sum P(X = x) = 1$
32. Appreciate the distinction between unbiased $(n - 1)$ and biased (n) estimators for the calculation of the population variance from the sample data
33. Understand and use the Expectation of a random variable, and of a function of a random variable, where in general $E(g(x)) = \sum g(x)P(X = x)$
34. Appreciate that $E(X^2) \neq [E(X)]^2$ and be able to explain the difference
35. Know how to write the Mean and Variance of a random variable X in terms of $E(X)$, $E(X^2)$ and $E^2(X) = [E(X)]^2$
36. Know how to calculate the Expectation and Variance of a linear function of a random variable: $E(aX + b)$ and $Var(aX + b)$
37. Understand the concept of Correlation between the two variables in a bivariate data set as a measure of the quality of a “best fit” line
38. Understand the concept of Linear Regression as a means of fitting a straight line through a set of data points, especially by the Method of Least Squares

39. Understand the distinction between the regression line y on x and the regression line x on y , and know when each is appropriate
40. Know how to calculate S_{xx} , S_{xy} and S_{yy} given n , $\sum x$, $\sum y$, $\sum xy$, $\sum x^2$ and $\sum y^2$
41. Know how to calculate and interpret (Pearson's) Product Moment Correlation Coefficient PMCC, especially with the assistance of a scientific calculator
42. Know how to calculate the equation of a regression line of A on B, especially with the assistance of a scientific calculator
43. Understand why the Regression Line must go through the Mean Point (\bar{x}, \bar{y})
44. Understand the concept of a Normal Distribution $X \sim N(\mu, \sigma^2)$ and the Standard Normal Distribution $Z \sim N(0,1)$, and the standardising process from $X \longrightarrow Z$
45. Understand and use tables or scientific calculator to determine approximate areas under the Normal curve between any two points
46. Know the relationship between areas under the Normal curve and associated probability calculations for a Normally-distributed continuous random variable
47. Know the approximate percentages of a Normal distribution contained within ± 1 , ± 2 and ± 3 standard deviations of the mean
48. Know how to use the symmetry of the Normal Distribution to calculate related probabilities
49. Know how to interpret phrases such as "at least", "between", "no more than" etc. in the context of probability questions
50. Know how to interpret tabulated or calculated values in the context of the original problem.