Department of Mathematics
Year 13 Scheme of Work – Statistics

$3 - 2 = 1$ → $\sin^2\theta + \cos^2\theta = 1$ → $\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty} e^{-\frac{1}{2}t^2} dt = 1$

Table of Contents

Chapter/Objectives	Teaching Points

REGRESSION, CORRELATION AND HYPOTHESIS TESTING

Chapter 1 – Regression, Correlation and Hypothesis Testing (Part A – Change of Variable) By the end of the sub-unit, students should: • be able to change the variable in a regression line; • be able to estimate values from regression line.	Start the revision of topics from year one by recapping regression. This needs to be extended to working with changing variables (coding) within regression lines. This relies on logarithms from the pure content and students should be able to work with equations of the form $y = ax^n$ and $y = kb^x$. Students will need to know how to put these into linear form and be able to estimate a and n or k and b. An understanding of reliability when extrapolating will also need to be recapped.
Chapter 1 – Regression, Correlation and Hypothesis Testing (Part B - Statistical hypothesis testing for zero correlation) By the end of the sub-unit, students should: • understand correlation coefficients; • be able to calculate the PMCC (calculator only); • be able to interpret a correlation coefficient; • be able to conduct a hypothesis test for a correlation coefficient.	Recap scatter diagrams and the terminology used in year one to describe correlation. Students should understand that measures of correlation can be calculated to identify the strength of correlation. They need to understand that one of these, the product moment correlation coefficient (PMCC) is denoted by r, and that $\lvert r\rvert \leq 1$. If $r = \pm 1$ then the data points lie on a perfect straight line on a graph. Students are expected to be able to calculate r using their calculators, but are not required to know or use the formula. They should be able to interpret their value for the PMCC in the context of the question. Students are required to perform hypotheses tests for correlation coefficients. The hypotheses need to be stated in terms of ρ where ρ represents the population correlation coefficient. All tests should have the null hypothesis $H_0: \rho = 0$. Tables of critical values or a p-value will be given to students.

PROBABILITY

Chapter 2 – Probability (Part A Conditional Probability) By the end of the sub-unit, students should: • understand and be able to use probability formulae using set notation; • be able to use tree diagrams, Venn diagrams and two-way tables; • understand and be able to use the conditional probability formula $\mathrm{P}(A\mid B) = \frac{\mathrm{P}(A\cap B)}{\mathrm{P}(B)}$.	Begin by recapping the use of tree diagrams and Venn diagrams, focusing on the use of set notation for probabilities. Introduce the use of two-way tables to find probabilities and use worded questions which are solved most efficiently by forming a two-way table. Students need to be familiar with and be able to use $\mathrm{P}(A') = 1 - \mathrm{P}(A)$, the addition rule: $\mathrm{P}(A \cup B) = \mathrm{P}(A) + \mathrm{P}(B) - \mathrm{P}(A \cap B)$ and the conditional probability formula $\mathrm{P}(A \cap B) = \mathrm{P}(A)\mathrm{P}(B\mid A)$. Use worded questions where students have to form the set notation as well as questions where the information is already given using set notation. Ensure the teaching of this section is combined with questions to recap the properties of mutually exclusive and independent events. Make sure these are now answered using set notation too.

Chapter/Objectives	Teaching Points

PROBABILITY

Chapter 2 – Probability (Part B Questioning Assumptions in Probability) By the end of the sub-unit, students should: • be able to model with probability; • be able to critique assumptions made and the likely effect of more realistic assumptions.	Students should know that probability can be used to predict how likely experiments are to have given outcomes. They should be able to determine all of the outcomes of an experiment (and know that these are called the sample space) and be able to determine the probability of each outcome of a given sample space. Students should also have an awareness of wider modelling where outcomes cannot be determined. Students should be able to question and critique any assumptions made in any given scenario. For example, assumptions about independence a reasonable assumption or whether a coin or dice is fair or biased. They should be able to look at the effect of these assumptions and have an awareness of assumptions that may be more realistic.

NORMAL DISTRIBUTION

Chapter 3 – Normal Distribution (Part A Understanding and Use Normal Distribution) By the end of the sub-unit, students should: • understand the properties of the Normal distribution; • be able to find probabilities using the Normal distribution; • know the position of the points of inflection of a Normal distribution.	The Normal distribution needs to be linked to histograms and the mean. A good way to introduce the topic is to look at heights on a histogram and show how it can be smoothed into the Normal distribution curve, stating this is due to the Normal being a continuous distribution. Discuss all the properties of the Normal distribution, making sure students are confident with the symmetry of the distribution, that mean = mode = median and the asymptotic nature of the bell-shaped curve. Cover the proportions of data within 1, 2 and 3 standard deviations of the mean and remind students that the area under the curve is 1. Students are expected to know that the points of inflection on the Normal curve are at $x = \mu \pm \sigma$ (they are not expected to be able to derive this). As with notation for the binomial distribution, students should understand the notation $X \sim \mathrm{N}(\mu, \sigma^2)$ for the Normal distribution. Students are expected to find the probabilities from Normal distributions using their calculators. However, students do need to know the standardisation formula $Z = \frac{x-\mu}{\sigma}$ and be able to transform X values to Z values. Be clear that the denominator is the standard deviation rather than the variance which may be given. Students should be encouraged to draw diagrams to represent the distribution and use this to check (at least for > or < 0.5) the probability they find using their calculator. Diagrams will also help students when working backwords from a probability to find a Z value, a diagram will indicate whether the Z value is positive or negative. Again, students are expected to use their calculator to find these values. Questions may involve the use of linear simultaneous equations to find for example both the mean and standard deviation of the Normal distribution. You should recap the probability of independent events as this can be incorporated into questions involving the Normal distribution.

Chapter/Objectives	Teaching Points

NORMAL DISTRIBUTION

Chapter 3 – Normal Distribution (Part B Using Normal Distribution as an Approximation) By the end of the sub-unit, students should: • be able to find the mean and variance of a binomial distribution; • understand and be able to apply a continuity correction; • be able to use the Normal distribution as an approximation to the binomial distribution.	Begin by recapping the binomial distribution and making clear that the Normal distribution is continuous and the binomial distribution is discrete. Students need to understand that the binomial distribution can be approximated by the Normal distribution when n is large and p is close to 0.5. Look at the parameters needed for the Normal distribution (μ and σ^2) and cover how the mean and variance are approximated from the binomial distribution ($\mu = np$ and $\sigma^2 = np(1-p)$). Students should be confident with the notation that $X \sim \mathrm{B}(n, p)$ is approximated by $Y \sim \mathrm{N}(np, np(1-p))$. Encourage them to write both distributions when answering questions involving an approximation. When calculating probabilities for a binomial distribution which has been approximated by the Normal distribution it is important to remember that a discrete distribution has become a continuous distribution and the continuity correction needs to be introduced. It is useful here to look back at the bar chart diagrams you used in year one. To help students understand the continuity correction label the edges of say the 8 bar with the boundaries 7.5 and 8.5 etc. If for the binomial distribution the probability $\mathrm{P}(X \leq 8)$ is required then shade the whole of the 8 bar and below; this indicates that the corresponding Normal probability is $\mathrm{P}(X < 8)$. Using the binomial distribution, for $\mathrm{P}(X < 8)$ the 8 bar won't be shaded but every bar below it will. This indicates that using the Normal distribution the probability will be $\mathrm{P}(Y < 7.5)$. The same principle works for probabilities of the form $\mathrm{P}(X > a)$ and $\mathrm{P}(X \geq a)$. Make sure students are clear that for Normal probabilities < and ≤ are interchangeable as it is a continuous distribution. Once students have mastered using the Normal distribution as an approximation to the binomial distribution make sure you give them the opportunity to solve questions where they have to explain which distribution can be used before solving the problem, and whether an approximation is necessary or not. Ensure they are competent in explaining why they have chosen the distribution or approximation, clearly stating the relevant properties of their chosen distribution. They should also be able to describe why they have discounted the use of a distribution or approximation.
Chapter 3 – Normal Distribution (Part C Hypothesis Testing for the Mean) By the end of the sub-unit, students should be able to: • be able to conduct a statistical hypothesis test for the mean of the Normal distribution; • be able to interpret the results in context.	Remind students of the properties of the Normal distribution and the parameters it uses. Questions could involve a known, given or assumed variance and students should be aware of this. Hypothesis tests need to be carried out for the mean of the Normal distribution. For $X \sim \mathrm{N}(\mu, \sigma^2)$, students need to understand that for a sample, $\overline{X} \sim \mathrm{N}\left(\mu, \frac{\sigma^2}{n}\right)$. Refer back to the formula used to translate X into Z and make sure students know they can test μ using $\frac{\overline{X}-\mu}{\frac{\sigma}{\sqrt{n}}} \sim \mathrm{N}(0, 1^2)$. This is the third type of hypothesis test that students are expected to be able to carry out so the importance of using the correct parameter in the hypotheses should be emphasised here. Hypotheses for the Normal distribution should be stated in terms of μ. As in all cases conclusions need to be written clearly and in the context of the question