UNIT 3: The Normal distribution

Return to overview

SPECIFICATION REFERENCES

4.2 Understand and use the Normal distribution as a model; find probabilities using the Normal distribution
Link to histograms, mean, standard deviation, points of inflection and the binomial distribution

4.3 Select an appropriate probability distribution for a context, with appropriate reasoning, including recognising when the binomial or the Normal model may not be appropriate

5.3 Conduct a statistical hypothesis test for the mean of the Normal distribution with known, given or assumed variance and interpret the results in context

PRIOR KNOWLEDGE

GCSE (9–1) in Mathematics at Higher Tier

A19 Solve two simultaneous equations in two variables (linear/linear or linear/quadratic) algebraically

AS Mathematics – Statistics content

3.1 Probability calculations, independent events (See Unit 3 of the SoW)

AS Mathematics – Statistics content Unit 4

4.1 Properties of the binomial distribution (See Unit 4 of the SoW)

3.1 Probability is the area under a curve (See Unit 4 of the SoW)

AS Mathematics – Statistics content Unit 5

5.1 Use appropriate language of statistical hypothesis testing (See Unit 5a of the SoW)

5.2 Be able to apply a hypothesis test to the binomial distribution (See Unit 5b of the SoW)

KEYWORDS

Binomial, discrete distribution, discrete random variable, uniform, cumulative probabilities Normal, mean, variance, continuous distribution, histogram, inflection, appropriate probability distribution.

3a. Understand and use the Normal distribution (4.2)

Teaching time
5 hours

OBJECTIVES

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.

TEACHING POINTS

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 backwards 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.

OPPORTUNITIES FOR PROBLEM SOLVING/MODELLING

Use plenty of reverse problem examples worded in a variety of ways. Ensure students can find Z values for quartiles and percentiles.

COMMON MISCONCEPTIONS/EXAMINER REPORT QUOTES

Main errors are due to confusion between probabilities and Z values, particularly when it comes to notation, and not using the full four decimal place accuracy in calculations.

An emphasis on using diagrams alongside the calculations should help address some of the difficulties.

NOTES

Knowledge of the probability density function is not required, neither are any derivations of mean, variance or the cumulative distribution function.

3b. Use the Normal distribution as an approximation to the binomial distribution; Selecting the appropriate distribution (4.2) (4.3)	Teaching time 5 hours

OBJECTIVES

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.

TEACHING POINTS

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 $\mathrm{P}(X \le 8)$ 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 $\mathrm{P}(X \ge a)$. Make sure students are clear that for Normal probabilities $<$ and $\le$ 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.

OPPORTUNITIES FOR PROBLEM SOLVING/MODELLING

Use an example when n is large and p is close to 0.5 and look at a variety of cumulative probabilities in both the binomial and the Normal distributions to show students how good the approximation is to the binomial distribution.

COMMON MISCONCEPTIONS/ EXAMINER REPORT QUOTES

Correctly applying continuity corrections can prove difficult with students either not applying one or otherwise adding 0.5 rather than subtracting or vice versa.

3c. Statistical hypothesis testing for the mean of the Normal distribution (5.3)	Teaching time 6 hours

OBJECTIVES

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.

TEACHING POINTS

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, $\bar{X} \sim \mathrm{N}\left(\mu, \frac{\sigma^2}{n}\right)$.

Refer back to the formula used to translate X into Z and make sure students know they can test μ using $\frac{\bar{X} - \mu}{\frac{\sigma}{\sqrt{n}}} \sim \mathrm{N}(0, 1^2)$.

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.

OPPORTUNITIES FOR PROBLEM SOLVING/MODELLING

Now all types of hypothesis testing have been covered students should be given mixed problems of all types from years one and two in order to practise distinguishing between tests. Ensure all hypotheses are written in terms of the correct parameter.

COMMON MISCONCEPTIONS/ EXAMINER REPORT QUOTES

Common errors in exam situations include: not expressing hypotheses precisely enough; using an incorrect parameter or not using a parameter at all; incorrectly applying the continuity correction; and not giving a conclusion or answer to the question using the given context.

NOTES

Knowledge of the central limit theorem is not required, neither are proofs of the sample formulae used.