

INTERNATIONAL ADVANCED LEVEL

MATHEMATICS/ FURTHER MATHEMATICS/ PURE MATHEMATICS

SCHEME OF WORK STATISTICS 2

Pearson Edexcel International Advanced Subsidiary in Mathematics (XMA01)

Pearson Edexcel International Advanced Subsidiary in Further Mathematics (XFM01)

Pearson Edexcel International Advanced Subsidiary in Pure Mathematics (XPM01)

Pearson Edexcel International Advanced Level in Mathematics (YMA01)

Pearson Edexcel International Advanced Level in Further Mathematics (YFM01)

Pearson Edexcel International Advanced Level in Pure Mathematics (YPM01)

First teaching September 2018

First examination from January 2019

First certification from August 2019 (International Advanced Subsidiary) and August 2020
(International Advanced Level)



Statistics 2

Unit	Title	Estimated hours
1	The binomial and Poisson distributions	
<u>a</u>	The binomial distribution	5
<u>b</u>	The Poisson distribution	6
<u>c</u>	Mean and variance of the binomial and Poisson distributions	5
<u>d</u>	Poisson distribution as an approximation to the binomial distribution	5
2	Continuous random variables	
<u>a</u>	Continuous random variables, the probability density function and the cumulative distribution function	9
<u>b</u>	Summary statistics for continuous random variables	6
3	Continuous distributions	
<u>a</u>	The continuous uniform distribution	3
<u>b</u>	Using the Normal distribution as an approximation to the binomial and Poisson distributions; Selecting the appropriate distribution	7
4	Hypothesis tests	
<u>a</u>	Introduction to sampling terminology; Advantages and disadvantages of sampling	3
<u>b</u>	Language of hypothesis testing; Significance levels; Critical regions	2
<u>c</u>	Carry out hypothesis tests involving the binomial distribution	5
<u>d</u>	Hypothesis test for the mean of a Poisson distribution	4
		60 hours

UNIT 1: The binomial and Poisson distributions

[Return to overview](#)

SPECIFICATION REFERENCES

- 1.1 The binomial and Poisson distributions
- 1.2 The mean and variance of the binomial and Poisson distributions
- 1.3 The use of the Poisson distribution as an approximation to the binomial distribution

PRIOR KNOWLEDGE

International GCSE/GCSE (9–1) in Mathematics at Higher Tier

- Order positive and negative integers, decimals and fractions; use the symbols $=$, $\neq$, $<$, $>$, $\leq$, and $\geq$

IAL Mathematics – Statistics 1 content

- Simple discrete probability distributions (See S1 SoW Unit 4)
- Mean and variance (See S1 SoW Unit 1)

KEYWORDS

Binomial distribution, probability, discrete distribution, discrete random variable, cumulative probabilities, sample space, probability distribution, probability function, mean, expectation, expected value, standard deviation, variance, Poisson distribution, exponential series, factorial, parameter, independent, singly.

1a. The binomial distribution (1.1)

Teaching time

5 hours

OBJECTIVES

By the end of the sub-unit, students should:

- understand and be able to use the binomial distribution;
- be able to calculate probabilities using the binomial distribution.

TEACHING POINTS

Students will be expected to model real-world situations by using simple discrete probability distributions. Students will be expected to comment critically on how appropriate a given probability model may be for a situation.

The notation $X \sim B(n, p)$ may be used, so you should ensure students are familiar with this from the outset. Make sure the properties of the binomial are clear for all students, so they know that a fixed number of trials is needed, there are only two possible outcomes per trial, the outcome of each trial is independent and the probability of success remains constant.

Once the binomial distribution has been introduced link back to thinking about probability being the area under a curve. Use a bar chart for discrete binomial distributions and show how this would smooth into a curve if it were a continuous distribution.

Students need to calculate probabilities using the binomial distribution for both individual and cumulative probabilities, either by applying the binomial formula or by reference to the tables. Students could also be shown how to use the calculator functions to find cumulative probabilities.

The bar chart model mentioned earlier helps students distinguish between for example $P(X < 2)$ and $P(X \leq 2)$, also to understand $P(X \geq 6) = 1 - P(X \leq 5)$. Explain this is due to the binomial being a discrete distribution. This is essential when manipulating before using tables or the calculator to find probabilities. Encourage students to shade the bars required, or use a number line and tick the values required, to help with this understanding.

Emphasise the importance of reading questions carefully. The probability of success can be worded negatively in the question for example ‘the probability of people failing their driving test first time is 0.6’.

OPPORTUNITIES FOR PROBLEM SOLVING/MODELLING

Look at a wide variety of real-world scenarios and model using a number of different distributions to ensure students are fluent in their comments on the appropriateness of a particular distribution.

COMMON MISCONCEPTIONS/EXAMINER REPORT QUOTES

The most common difficulty is with manipulating inequalities: ‘A significant number of students were unable to cope with the expression $P(5 \leq X < 11)$. There were students who translated this expression into the more convenient form $P(5 \leq X \leq 10)$ and then in turn transformed this into an equivalent form that can be applied to the table of cumulative probabilities: $P(X \leq 10) - P(X \leq 4)$. However, there were also many instances of incorrect versions such as: $P(X < 11) - P(X \geq 5)$, $P(X \leq 10) + P(X \geq 5)$, $P(X \leq 10) - (1 - P(X \geq 5))$ and $P(X \leq 11)$ – either $P(X \leq 5)$ or $P(X \leq 4)$.’

In a similar vein, students have a tendency to write, for example, $P(X > 2)$ as $1 - P(X \leq 1)$ instead of $1 - P(X \leq 2)$.

A level Further Mathematics: Statistics 2

NOTES

It would be good for understanding to see what the random variable looks like in pdf form and table form although this is not explicitly in the specification.

1b. The Poisson distribution (1.1)

Teaching time

6 hours

OBJECTIVES

By the end of the sub-unit, students should:

- be able to use the Poisson distribution to model a real-world situation;
- know the conditions for a Poisson distribution;
- be able to comment critically on the appropriateness of using the Poisson distribution as a model;
- know how to use the formula, tables and calculators to find probabilities, including cumulative probabilities;
- know the additive property of Poisson distributions.

TEACHING POINTS

The Poisson distribution is similar to the binomial distribution in many ways. The formula

$P(X = x) = e^{-\lambda} \frac{\lambda^x}{x!}$ is given in the formulae booklet. Encourage students to use this to find single probabilities, or probabilities for values of λ which don't appear in the tables, as well as practising using their calculators to find single and cumulative probabilities.

Look out for the word (or implication of) 'rate' in a question as this suggests a Poisson distribution, as does 'an average number of occurrences'.

Questions involving the Poisson distribution should be stated using the notation $X \sim \text{Po}(\lambda)$.

Explain the conditions for a Poisson distribution. If X is the number of events occurring in a fixed interval of space or time, then the events occur:

- independently of each other
- singly in space or time
- at a constant rate.

Students will be expected to use the Poisson distribution to model a real-world situation and to comment critically on the appropriateness. Conditions for a Poisson distribution must always relate to the context of the question.

The additive property of Poisson distributions means that if a rate of occurrence is given for one interval (e.g. telephone calls arrive at a rate of 1.5 per minute), then the value of λ can also be calculated for a different interval appropriate to the question, e.g. for a 10-minute period, $\lambda = 10 \times 1.5 = 15$.

It is important that conclusions are always written in the context of the original question.

OPPORTUNITIES FOR REASONING/PROBLEM SOLVING

The Poisson distribution can be related to the exponential series $e^x = 1 + \frac{x^1}{1!} + \frac{x^2}{2!} + \frac{x^3}{3!} \dots$. Setting $x = \lambda$ and dividing by e^λ leads to an infinite series equal to 1. The terms of this series could then be used as probabilities to define a probability distribution.

A common question is to use a probability calculated from a Poisson distribution as the subsequent probability in a binomial distribution. For example:

Accidents occur randomly at a road junction at a rate of 18 every year. Show that the probability of at least 1 accident in a randomly selected month is 0.777. Find the probability that there is at least one accident in exactly 4 of the next 6 months.

Questions with two inequalities could be asked. For example:

A level Further Mathematics: Statistics 2

A shop sells radios at a rate of 2.5 per week. Deliveries of these radios come every four weeks. The manager wishes to make sure that the probability of the shop running out of radios during a four-week period is less than 0.01. Find the smallest number of radios the manager should have in stock immediately after the delivery.

To solve this problem, we need to solve the inequality $P(Y > s) < 0.01$ where s is the number of radios the manager should have in stock and Y is the number of radios sold in a four-week period.

You could extend to examples using the additive property of Poisson distributions such as:

A company has 2 machines P and Q. On average there are 0.8 breakdowns per month on machine P and 6 breakdowns per year on machine Q.

To find the probability of there being a total of 2 breakdowns on these two machines in a given month, we first need to make sure both variables are for the same unit of time: Q has 0.8 per month, and P has 0.5 breakdowns *per month*, so we would use $\lambda = 0.8 + 0.5 = 1.3$

COMMON MISCONCEPTIONS/EXAMINER REPORT QUOTES

The conditions for the Poisson distribution must be put into the context of the question. Students may learn when a Poisson distribution is suitable, but not be able to use the context given in the question. The conditions must be given precisely, it is not enough to simply mention ‘rate’ to get the mark for ‘constant rate’.

Students must be clear on how to write ‘at least’, ‘at most’, ‘more than’ and ‘less than’ using inequalities. It is a fairly common mistake, for example, for students to work out $P(X < 10)$ or $P(X > 10)$ when they should instead be calculating ‘at least 10’.

Inequalities can also cause problems when ‘flipping’ a probability: ‘any errors that arose were generally from writing $P(Y \geq 10) = 1 - P(Y \leq 10)$ or $1 - P(Y < 9)$ ’.

Similarly, ‘many students did not understand that to work out the probability that “the next patient arrives” (requiring $P(X \geq 1)$) is the same as $1 - P(X = 0)$. Many calculated $P(X = 1)$.’

Students need to be able to use the formula for values of λ that don’t occur in the tables: ‘A curious error was to approximate 1.8 by 2 to allow the question to be solved by use of tables. This scored zero marks.’

NOTES

No proofs are required for the additive property of Poisson distributions.

The Poisson distribution has applications in modern life for modelling radioactive decay, predicting the arrival of buses, describing the spread of trees in a forest and predicting the number of telephone calls arriving at an exchange.

1c. Mean and variance of the binomial and Poisson distributions (1.2)**Teaching time**

5 hours

OBJECTIVES

By the end of the sub-unit, students should:

- know how to find the mean and variance of the binomial distribution;
- know how to find the mean and variance of the Poisson distribution;
- be able to solve problems involving the mean and variance of the binomial and Poisson distributions.

TEACHING POINTS

The formulae for mean and variance of both binomial and Poisson distributions are given in the formulae booklet:

- If $X \sim B(n, p)$, then $E(X) = np$ and $\text{Var}(X) = np(1 - p)$
- If $Y \sim \text{Po}(\lambda)$, then $E(Y) = \lambda$ and $\text{Var}(Y) = \lambda$

Questions could include finding $P(\mu - \sigma < X < \mu + \sigma)$ or finding a sample size for a binomial distribution when the mean, standard deviation or variance is given.

OPPORTUNITIES FOR REASONING/PROBLEM SOLVING

The property that $E(X) = \text{Var}(X) = \lambda$ can sometimes be used to justify why a Poisson distribution is a suitable model for a particular situation.

NOTES

Students are not required to derive formulae for mean and variance of binomial or Poisson distributions.

1d. Poisson distribution as an approximation to the binomial distribution (1.3)**Teaching time**
5 hours**OBJECTIVES**

By the end of the sub-unit, students should:

- be able to use the Poisson distribution as an approximation to the binomial distribution;
- know when it is appropriate to use the Poisson distribution as an approximation to the binomial distribution.

TEACHING POINTS

Evaluating binomial probabilities when n is large can be difficult and in such cases, it may be simpler to use an approximation. The binomial distribution can be approximated by a Poisson distribution when n is large and p is small. This is quite vague, and while there is no rule for how ‘large’ n needs to be, or how ‘small’ p needs to be, a general rule is that we need $np \leq 10$. The approximation works best the larger the n and the smaller the p .

COMMON MISCONCEPTIONS/EXAMINER REPORT QUOTES

It is important that students can correctly switch from the number of ‘successes’ to the number of ‘failures’ for a binomial distribution. For example, as an examiner commented: ‘Most students were able to see that the exact distribution was $B(100, 0.975)$ for the number of times the cadet hits the target. Few realised that they should define $B(100, 0.025)$ as the model for the number of times the cadet misses the target and then follow this with an approximation to the Poisson obtaining $Po(2.5)$. Even when they used $Po(2.5)$, many were unable to change “the cadet hits at least 95” to “the cadet misses at most 5”.’

NOTES

Although formal derivations for when the Poisson distribution can be approximated to the binomial are not required, mathematical software can be used to demonstrate that a binomial distribution approximates to a Poisson distribution when p is small and n is large.

UNIT 2: Continuous random variables

[Return to Overview](#)

SPECIFICATION REFERENCES

- 2.1 The concept of a continuous random variable
- 2.2 The probability density function and the cumulative distribution function for a continuous random variable
- 2.3 Relationship between density and distribution functions
- 2.4 Mean and variance of continuous random variables
- 2.5 Mode, median and quartiles of continuous random variables

PRIOR KNOWLEDGE

IAL Mathematics – Statistics 1 content

- Histograms (See S1 SoW Unit 1)
- Mean, mode, median, variance (See S1 SoW Unit 1)
- Probability (See S1 SoW Unit 2)

IAS Mathematics – Pure 1 and 2 content

- Differentiation (See P1 SoW Unit 4 and P2 SoW Unit 7)
- Integration (See P1 SoW Unit 5 and P2 SoW Unit 8)

KEYWORDS

Discrete, continuous, continuous random variable, probability density function, histogram, relative frequency, frequency density, cumulative distribution function, mean, expected value, standard deviation, variance, mode, median, quartile, percentile, positive skew, negative skew.

2a. Continuous random variables, the probability density function and the cumulative distribution function (2.1) (2.2) (2.3)
Teaching time
9 hours

OBJECTIVES

By the end of the sub-unit, students should:

- understand the concept of a continuous random variable;
- be able to find and use the probability density function (p.d.f.) for a continuous random variable;
- be able to find and use the cumulative distribution function (c.d.f.) for a continuous random variable;
- know the relationship between probability density and cumulative distribution functions and convert between the two.

TEACHING POINTS

Students will be expected to link the concept of continuous random variables with their knowledge of histograms and frequency polygons. The cumulative distribution function (c.d.f.) gives the probability that the value is less than or equal to x . Since the area under a curve gives probabilities, the c.d.f is found by integrating the p.d.f.

Care must be taken at the end of intervals, and to include all the intervals before it, particularly if the p.d.f. is given as more than two 'pieces'.

Remember that $\int x^n dx = \frac{x^{n+1}}{n+1} + c$. If this has not been covered yet, it is easily inferred from 'reversing' differentiation.

Notation for the c.d.f. is $F(x)$.

The c.d.f will always have one more 'piece' than the p.d.f., to account for $F(x) = 1$. $F(x)$ must be defined over the whole range $(-\infty, \infty)$.

For continuous random variables, $P(X \geq m) = P(X > m)$ and $P(X = m) = 0$. Note that the probability of a single value happening in a continuous distribution is always zero.

If it is possible to sketch the curve for the p.d.f., this will help in finding the c.d.f. and determining the skewness of the distribution. If asked for a sketch, remember to label the axes and label the important values. These are the boundaries of the given range of x -values and their corresponding y -values. Students must also make it clear on the graph when $f(x) = 0$.

The total area under the curve in the required interval is 1, and this property can often be used to solve for an unknown value.

The link between the p.d.f. and the c.d.f. should be clear. If X is a continuous random variable with c.d.f. $F(x)$ and p.d.f. $f(x)$:

$$f(x) = \frac{dF(x)}{dx} \text{ and } F(x) = \int_{-\infty}^x f(t)dt.$$

Where there are two or more 'pieces' it is often easier to define the c.d.f. by finding the indefinite integral for each 'piece' and calculating a constant of integration by using, for example, $F(\text{lower limit}) = 0$ and $F(\text{upper limit}) = 1$.

A level Further Mathematics: Statistics 2

OPPORTUNITIES FOR REASONING/PROBLEM SOLVING

Students should be able to answer questions such as:

The duration, t , of phone calls, in minutes, is modelled by the function

$$f(t) = \begin{cases} \frac{1}{k}(t - 2), & 2 \leq t \leq 12 \\ 0, & \text{otherwise} \end{cases}.$$

Bill makes a call. What is the probability the call lasts between 3 and 5 minutes?

Questions involving conditional probabilities could also be considered, such as ‘What is the probability the call lasts at least 5 minutes given that Bill has already spent 3 minutes on the call?’

COMMON MISCONCEPTIONS/EXAMINER REPORT QUOTES

Particular care is needed when finding the c.d.f. from a p.d.f. that is expressed piecewise, as highlighted by one examiner: ‘persistence and attention to detail were required, particularly regarding algebra and arithmetic. However, the real problem for many students was conceptual. Almost all students could integrate all three parts of $f(x)$. However, a substantial number apparently did not address the central notion of cumulative probability: they failed to add $\frac{1}{9}$ for the second part and $\frac{7}{9}$ for the third part. The students who used the + c method were able to use $F(0) = 0$ and $F(6) = 1$ to gain the 2nd and 4th line of the c.d.f. however they did not know what to use to calculate the c to gain the 3rd line; many simply used $F(0) = 0$ for a second time.’

Students need to be clear on the difference between discrete and continuous distributions: ‘A few tried to treat this as a discrete distribution substituting every integer from 1 to 9 into the formula and adding them to equal 1.’

Care must be taken in defining the c.d.f. fully, including the last ‘piece’, stating that $F(x) = 1$ for the remaining values: ‘Some students failed to completely define $F(x)$, as given in the question, and so lost the second mark.’

NOTES

The formulae used in defining $f(x)$ will be restricted to simple polynomials.

2b. Summary statistics for continuous random variables (2.4) (2.5)

Teaching time
6 hours
OBJECTIVES

By the end of the sub-unit, students should:

- know how to find the mean and variance of continuous random variables;
- know how to find the expectation of a linear function of a continuous random variable;
- know how to find the mode, median and percentiles of continuous random variables;
- be able to use calculations to assess the suitability of models;

TEACHING POINTS

Finding the mean for a continuous random variable is similar to finding the mean for a discrete random variable: $\mu = E(X) = \sum xp(x)$ and replacing $\sum$ by $\int_{-\infty}^{\infty}$ and $p(x)$ by $f(x)dx$ gives

$\mu = E(X) = \int_{-\infty}^{\infty} xf(x)dx$ for continuous random variables.

Similarly, the variance is still found using $\text{Var}(X) = E(X^2) - (E(X))^2$ with $E(X^2) = \int_{-\infty}^{\infty} x^2 f(x)dx$.

If the graph of the p.d.f. is symmetrical, $E(X)$ will occur on the line of symmetry. Examples where there are two or more ‘pieces’ should also be considered.

The mode of a p.d.f. is the value of x where $f(x)$ is ‘highest’; a sketch is often helpful. Sometimes the mode occurs at a stationary point, but not always.

To find the median, we use the c.d.f. $F(x)$ and solve for m in the equation $F(m) = 0.5$. Values for other quartiles as well as percentiles are found similarly, with 0.5 being replaced with the appropriate value, e.g. the lower quartile, Q_1 , is given by $F(Q_1) = 0.25$, and the upper quartile, Q_3 , is given by $F(Q_3) = 0.75$.

Students could be reminded that, for example, if the median is less than the mean, the distribution has a positive skew and vice versa.

The formulae for expectation and variance of a function of a continuous random variable are the same as for discrete random variables:

- $E(aX) = aE(X)$
- $E(X + b) = E(X) + b$
- $E(aX + b) = aE(X) + b$
- $\text{Var}(aX) = a^2\text{Var}(X)$
- $\text{Var}(X + b) = \text{Var}(X)$
- $\text{Var}(aX + b) = a^2\text{Var}(X)$.

OPPORTUNITIES FOR REASONING/PROBLEM SOLVING

Questions may require students to use the calculations of mean, variance, etc. to assess the suitability of models or determine whether a given p.d.f. is valid using the conditions: $\int_{-\infty}^{\infty} f(x)dx = 1$ and $f(x) \geq 0 \quad \forall x$.

COMMON MISCONCEPTIONS/EXAMINER REPORT QUOTES

A fairly common mistake is for students to forget the ‘ x ’ in the formula for finding the mean and integrate the p.d.f. rather than $xf(x)$.

A level Further Mathematics: Statistics 2

When the p.d.f. is expressed piecewise, care needs to be taken when calculating $E(X)$. Common errors include: using only one of the 'parts' of the p.d.f.; adding all the 'parts' before integrating; and integrating all the 'parts' successfully, but then dividing by the number of parts.

When finding the median, some students may try to evaluate $F(0.5)$ rather than solving the equation $F(m) = 0.5$

When the p.d.f. is given piecewise, students need to check which 'piece' of the p.d.f. or c.d.f. contains their median as highlighted by an examiner: 'A significant number of students used the wrong part of their version of $F(x)$. Students would be advised to do a little bit of preliminary work, which consists of:

$F(1) = \frac{1}{9} < \frac{1}{2}$; $F(4) = \frac{7}{9} > \frac{1}{2}$ so $1 < m < 4$, so use the third line of $F(x)$.'

Sketching the p.d.f. can be very helpful, particularly when finding the mode and students must be clear that the mode is the maximum point on their function: 'The majority of students either forgot to justify their answer was the mode or they did not know that to justify their answer was the mode they needed to prove it was a maximum point.'

Just as with discrete distributions, the formula for $\text{Var}(X)$ can also be written as $\text{Var}(X) = E(X^2) - (E(X))^2$ and students may be asked to calculate $E(X^2)$ when given $\text{Var}(X)$ and $E(X)$. This has caused difficulty on occasion: '... a minority of students were unable to distinguish between $E(X^2)$ and $\text{Var}(X)$.'

NOTES

The formulae used in defining $f(x)$ will be restricted to simple polynomials.

UNIT 3: Continuous distributions

[Return to overview](#)

SPECIFICATION REFERENCES

- 3.1 The continuous uniform (rectangular) distribution
- 3.2 Use of the Normal distribution as an approximation to the binomial distribution and the Poisson distribution, with the application of the continuity correction

PRIOR KNOWLEDGE

- Properties of the binomial and Poisson distributions (See Unit 1)
- Probability is the area under a curve (See Unit 2)

IAL Mathematics – Statistics 1 content

- Probability calculations, independent events (See S1 SoW Unit 2)
- Normal distribution (See S1 Unit 5)

KEYWORDS

Continuous uniform distribution, cumulative probabilities, Normal, mean, variance, continuous distribution, histogram, appropriate probability distribution.

3a. The continuous uniform distribution (3.1)

Teaching time

3 hours

OBJECTIVES

By the end of the sub-unit, students should:

- understand and be able to use the continuous uniform (rectangular) distribution;
- be able to derive the mean and variance of a continuous uniform distribution;
- be able to find the cumulative distribution function.

TEACHING POINTS

The continuous uniform (rectangular) distribution is denoted by $X \sim U[a, b]$.

Probabilities can be found by working out the area of a rectangle. Sketch the p.d.f. first.

Students may be asked to derive the formulae for the mean and variance of a continuous uniform distribution.

To derive $E(X)$, use the definition $E(X) = \int_{-\infty}^{\infty} xf(x)dx$. This gives

$$\begin{aligned}
 E(X) &= \frac{1}{b-a} \int_a^b x dx \\
 &= \frac{1}{b-a} \left[\frac{x^2}{2} \right]_a^b \\
 &= \frac{1}{b-a} \left(\frac{b^2 - a^2}{2} \right) \\
 &= \frac{1}{2} \frac{1}{(b-a)} (b-a)(b+a) \\
 &= \frac{1}{2} (a+b).
 \end{aligned}$$

Similarly, to derive $\text{Var}(X)$, use $\text{Var}(X) = E(X^2) - (E(X))^2$ and $E(X^2) = \int_{-\infty}^{\infty} x^2 f(x) dx$, so

$$\begin{aligned}
 E(X^2) &= \frac{1}{b-a} \int_a^b x^2 dx \\
 &= \frac{1}{b-a} \left[\frac{x^3}{3} \right]_a^b \\
 &= \frac{1}{b-a} \left(\frac{b^3 - a^3}{3} \right) \\
 &= \frac{1}{3} \frac{1}{(b-a)} (b-a)(b^2 + ab + a^2) \\
 &= \frac{1}{3} (b^2 + ab + a^2).
 \end{aligned}$$

$$\begin{aligned}
 \text{Then } \text{Var}(X) &= \frac{1}{3} (b^2 + ab + a^2) - \left\{ \frac{1}{2} (a+b) \right\}^2 \\
 &= \frac{1}{12} (4b^2 + 4ab + 4a^2 - 3a^2 - 6ab - 3b^2) \\
 &= \frac{1}{12} (b^2 - 2ab + a^2) \\
 &= \frac{1}{12} (b-a)^2.
 \end{aligned}$$

Students should also be able to derive the formula for $F(x)$ from first principles:

$$\begin{aligned}
 \text{If } a \leq x \leq b, F(x) &= \int_a^x \frac{1}{b-a} dt \\
 &= \left[\frac{t}{b-a} \right]_a^x \\
 &= \frac{x-a}{b-a}.
 \end{aligned}$$

A level Further Mathematics: Statistics 2

OPPORTUNITIES FOR REASONING/PROBLEM SOLVING

The uniform distribution is often used as a model for errors made by rounding up or down when recording measurements. If values are measured to the nearest unit, then the distribution of rounding errors, R , will follow $R \sim U[-0.5, 0.5]$.

Other questions involve, for example, a piece of string cut into two pieces.

COMMON MISCONCEPTIONS/EXAMINER REPORT QUOTES

The formulae for mean and variance are given in the formula booklet and can be used: ‘A few students integrated [to find the variance] in contrast to the majority who used the formula $\text{Var} = \frac{1}{12}(b - a)^2$ from the formula book.’

3b. Using the Normal distribution as an approximation to the binomial and Poisson distributions; Selecting the appropriate distribution (3.2)

Teaching time
7 hours

OBJECTIVES

By the end of the sub-unit, students should:

- understand and be able to apply a continuity correction;
- be able to use the Normal distribution as an approximation to the binomial distribution;
- be able to use the Normal distribution as an approximation to the Poisson distribution;
- be able to select an appropriate 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. Mathematical software can be used to visualise this and to demonstrate that the distributions are very similar as p approaches 0.5 and as n increases. 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 $X \sim B(n, p)$ is approximated by $Y \sim N(np, np(1 - p))$. 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, or number lines, you used previously.

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 $P(X \leq 8)$ is required then shade the whole of the 8 bar and below; this indicates that the corresponding Normal probability is $P(X < 8.5)$.

Using the binomial distribution, for $P(X < 8)$ the 8 bar won't be shaded but every bar below it will. This indicates that using the Normal distribution the probability will be $P(Y < 7.5)$. The same principle works for probabilities of the form $P(X > a)$ and $P(X \geq a)$. Make sure students are clear that for Normal probabilities $<$ and $\leq$ 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.

Recap the Poisson distribution and again make clear that the Normal distribution is continuous and the Poisson distribution is discrete. Mathematical software can be used to demonstrate that the Poisson distribution approaches a normal distribution as the value of the parameter λ increases. Although there is no rule for how 'large' λ needs to be, a general rule is that we need $\lambda > 10$ to be able to use the normal distribution as an approximation to a Poisson distribution.

Again look at the parameters needed for the Normal distribution (μ and σ^2) and discuss how the mean and variance are approximated from the Poisson distribution $\mu = \lambda$ and $\sigma^2 = \lambda$. Students should be confident with

A level Further Mathematics: Statistics 2

the notation that $X \sim \text{Po}(\lambda)$ is approximated by $Y \sim N(\lambda, \lambda)$. Encourage them to write both distributions when answering questions involving an approximation.

OPPORTUNITIES FOR PROBLEM SOLVING/MODELLING

Use an example when n is large and p is close to 0.5 or where λ is large and look at a variety of cumulative probabilities in both the binomial or Poisson and the Normal distributions to show students how good the approximation is to the binomial or Poisson distribution. Use of the calculator functions for the three distributions would be particularly useful here.

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.

UNIT 4: Hypothesis tests

[Return to overview](#)

SPECIFICATION REFERENCES

- 4.1 Population, census and sample. Sampling unit, sampling frame
- 4.2 Concepts of a statistic and its sampling distribution
- 4.3 Concept and interpretation of a hypothesis test. Null and alternative hypotheses
- 4.4 Critical region
- 4.5 One-tailed and two-tailed tests
- 4.6 Hypothesis tests for the parameter p of a binomial distribution and for the mean of a Poisson distribution

PRIOR KNOWLEDGE

- An understanding of how to calculate binomial and Poisson probabilities (See Unit 1)
- International GCSE/GCSE (9–1) in Mathematics at Higher Tier
- Infer properties of populations or distributions from a sample, while knowing the limitations of sampling (GCSE only)
 - Apply statistics to describe a population

KEYWORDS

Population, census, sample, sampling unit, sampling frame, simple random sampling, sampling distribution, hypotheses, significance level, one-tailed test, two-tailed test, test statistic, null hypothesis, alternative hypothesis, critical value, critical region, acceptance region, p-value, binomial model, accept, reject, parameter, inference, significant, Poisson distribution.

4a. Introduction to sampling terminology; Advantages and disadvantages of sampling (4.1) (4.2)**Teaching time**

3 hour

OBJECTIVES

By the end of the sub-unit, students should:

- understand and be able to use the terms ‘population’ and ‘sample’;
- understand and be able to use the terms ‘statistic’ and ‘sampling distribution’;
- be able to describe advantages and disadvantages of sampling compared to census.

TEACHING POINTS

Real-world data can be used to look at a population and discuss reasons for sampling from it.

Students will be expected to be able to comment on the advantages and disadvantages associated with a census and a sample.

Discuss in context the meanings of populations and samples, sampling units and sampling frames. Ensure students are given the opportunity, and are able, to give full and thorough answers within the context of the question.

Discuss the meaning of a statistic and generate sampling distributions for simple statistics, such as the mean or median value of the samples of size 2 or 3.

One question might be: ‘A bag contains a large number of coins. It contains only 1p and 2p coins in the ratio 1 : 3. A random sample of size 3 is taken from the bag. List all the possible samples and find the sampling distribution of the mean value of the samples.’

OPPORTUNITIES FOR PROBLEM SOLVING/MODELLING

There is an opportunity here to introduce students to real-world data and to provide an opportunity to practise defining the statistical terms within a context.

COMMON MISCONCEPTIONS/EXAMINER REPORT QUOTES

Some students confuse sample sizes and population sizes, but the recurring problem is not giving answers in context. Students should also make sure they are clear about the definitions of the terminology for sampling and given plenty of opportunity to interpret these in a given context.

4b. Language of hypothesis testing; Significance levels; Critical regions
(4.3) (4.4) (4.5)**Teaching time**
2 hours**OBJECTIVES**

By the end of the sub-unit, students should:

- understand and be able to apply the language of statistical hypothesis testing, developed through a binomial model.

TEACHING POINTS

The concept of a hypothesis could be introduced initially by posing some hypotheses yourself. You may wish to use real-world data or present a simple situation such as four children playing a game of chance where one child appears to be winning more than 25% of the games, and so you might ask: ‘is this child cheating?’

Following this, introduce the null and alternative hypotheses and their respective notation H_0 and H_1 . Discuss how to move from statements like the one above to using the language of the binomial distribution in terms of looking at p , the probability of success. Also discuss the difference between a one-tailed and a two-tailed test. The idea of a significance level and critical regions should be introduced. Make sure that students understand that $X \sim B(n, p)$, the number of successes, is the test statistic.

The focus of this sub-unit is the language used in terms of hypothesis testing, but a scenario must be set. Another example to use might be ‘the number of 6s thrown in 50 throws of a dice’. Students could carry out this experiment and you could use their results to form a variety of tests which would cover all of the terminology without actually carrying out the tests. Make sure you save these examples to be tested in the next sub-unit.

Definitions of hypothesis test, significance level and critical region should be learned:

- **hypothesis test:** a statement made about the value of a population parameter that we test by collecting evidence in the form of a sample
- **significance level:** the probability of incorrectly rejecting the null hypothesis
- **critical region:** the range of values of the test statistic that would lead to rejecting the null hypothesis.

All of the terms from the keywords section should be thoroughly discussed and understood before attempting to carry out a hypothesis test.

OPPORTUNITIES FOR PROBLEM SOLVING/MODELLING

Use a wide variety of scenarios from the real world and invite students to offer their own scenarios; discuss their suitability for hypothesis testing.

COMMON MISCONCEPTIONS/EXAMINER REPORT QUOTES

Emphasise the importance of stating hypotheses clearly using the correct notation.

NOTES

The expected value of the binomial distribution being np needs to be appreciated for a two-tailed test.

**4c. Carry out hypothesis tests involving the binomial distribution
(4.4) (4.5) (4.6)****Teaching time**
5 hours**OBJECTIVES**

By the end of the sub-unit, students should:

- be able to conduct a statistical hypothesis test for the proportion in the binomial distribution and interpret the results in context;
- be able to find critical regions for a hypothesis test involving a binomial distribution;
- understand that a sample is being used to make an inference about the population;
- appreciate that the significance level is the probability of incorrectly rejecting the null hypothesis.

TEACHING POINTS

Once all the terminology that has been discussed in the previous sub-unit is fully understood, you can go back to the examples you used and conduct the hypothesis tests.

If the alternative hypothesis is two-tailed (i.e. contains $\neq$), then the significance level needs to be halved.

Where the critical region is required, questions asking for the value in each tail to be as close as possible to α are answered differently to questions where this is not asked. If this is *not* asked for, the probability must be strictly within the significance level, rather than as close to it as possible either side.

The ‘actual significance level of a test’, the *actual* probability of being in the critical region, should also be covered.

Carry out the tests both by finding the critical value or critical region to compare with your test statistic and by finding the probability (p-value) of the test statistic and comparing it with the significance level. Ensure students are competent with both methods. Make sure hypotheses are always written clearly in terms of p , the probability of success.

Spend time making sure that students can write clear and concise conclusions in the given context of the questions.

When using a sample of data, ensure students understand, what it infers about the population itself.

OPPORTUNITIES FOR PROBLEM SOLVING/MODELLING

Again, use a wide variety of scenarios from the real world and make sure all conclusions are written in these contexts.

COMMON MISCONCEPTIONS/ EXAMINER REPORT QUOTES

The most common error in these sorts of questions include not writing a clear conclusion in the context of the question. Students either omit the context or sometimes fail to give any conclusion to their calculations. Correct notation is important when describing the critical region: ‘There were still a few students using incorrect notation for critical regions: $P(X \leq 1)$, for example, is not a critical region: it is a probability.’

4d. Hypothesis test for the mean of a Poisson distribution (4.4) (4.5) (4.6)

Teaching time
4 hours

OBJECTIVES

By the end of the sub-unit, students should:

- understand and be able to apply the language of statistical hypothesis testing in the context of a Poisson distribution;
- be able to carry out hypothesis tests to test for the mean of a Poisson distribution;
- be able to find critical regions for a hypothesis test involving a Poisson distribution.

TEACHING POINTS

Testing for the mean of a Poisson distribution is a simple extension of the hypothesis test for the proportion of the binomial distribution. The null hypothesis will be that $\lambda = [\text{some value}]$, and the alternative hypothesis is that λ is $<$ or $>$ or $\neq$ [the value]. The test statistic is $X \sim \text{Po}(\lambda)$, the number of events in a given time or space.

Again, carry out the tests both by finding the critical value or critical region to compare with the value of your test statistic, and by finding the probability (p-value) of the test statistic and comparing it with the significance level. Ensure students are competent with both methods. Make sure hypotheses are always written clearly in terms of λ .

Spend time making sure that students can write clear and concise conclusions in the given context of the questions.

Type I errors are not part of the specification, but it is important that students understand what the significance level of a test actually means. Discuss carefully that rejecting the null hypothesis may actually be incorrect and the significance level is the probability of this. Also cover ‘the actual significance level of a test’ with students when finding the critical regions.

OPPORTUNITIES FOR REASONING/PROBLEM SOLVING

Applications include testing if a new drug is effective, or whether introducing a new speed limit on a certain stretch of road reduces accidents.

COMMON MISCONCEPTIONS/EXAMINER REPORT QUOTES

Common errors that are seen in exams include:

- not being able to find a critical region;
- confusing critical values and critical regions;
- confusing the probability of being in the critical region with the critical region;
- not clearly stating hypotheses;
- not giving a conclusion within the context of the question.