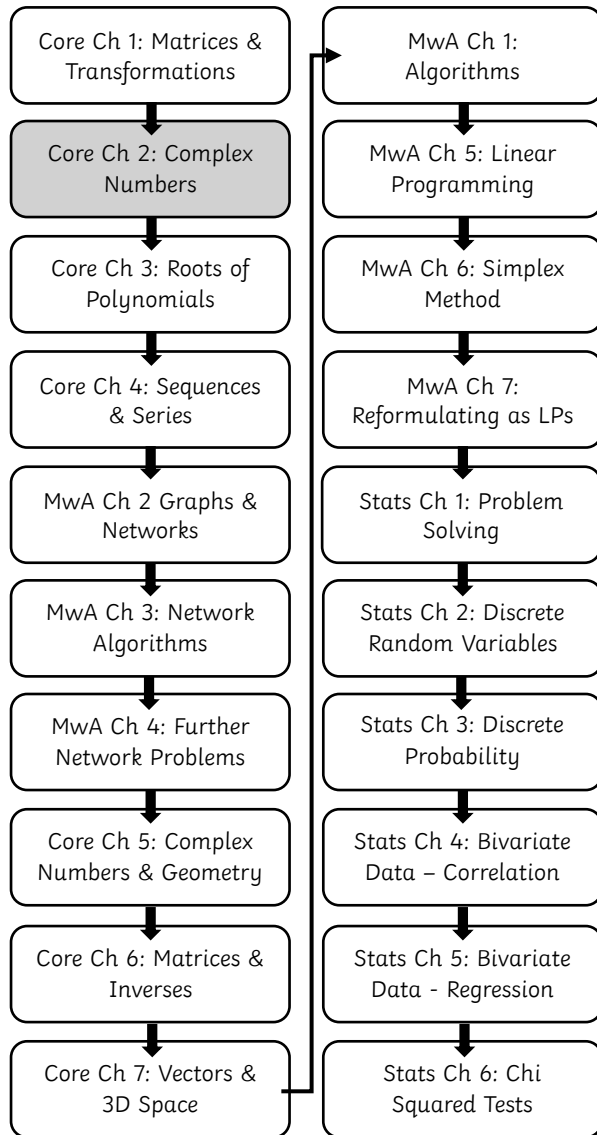


Core 1 Chapter 1: Matrices and Transformations	
Know a matrix with $r$ rows and $c$ columns has order $r \times c$	
Know that equal matrices have the same order and have matching corresponding elements	
Know a matrix with the same number of rows and columns is a square matrix	
Know the identity matrix is of the form $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$	
Know a zero matrix is of the form $\begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$	
Add and subtract matrices of the same order	
Multiply a matrix by a scalar	
Know two matrices are conformable for multiplication if their orders are of the form $p \times q$ and $q \times r$ respectively, and be able to multiply conformable matrices	
Use a calculator to carry out matrix calculations	
Know that matrix addition and multiplication are associative because $\mathbf{A} + (\mathbf{B} + \mathbf{C}) = (\mathbf{A} + \mathbf{B}) + \mathbf{C}$ and $\mathbf{A}(\mathbf{BC}) = (\mathbf{AB})\mathbf{C}$	
Know that matrix addition is commutative because $\mathbf{A} + \mathbf{B} = \mathbf{B} + \mathbf{A}$ , but matrix multiplication is not because, in general, $\mathbf{AB} \neq \mathbf{BA}$	
Find the matrix associated with a linear transformation in two dimensions: <ul style="list-style-type: none"> <li>• Reflections in the coordinate axes and <math>y = \pm x</math></li> <li>• Rotations about the origin</li> <li>• Enlargements, centre the origin</li> <li>• Stretches parallel to the coordinate axes</li> <li>• Shears with the coordinate axes as fixed lines</li> </ul>	
Find the matrix associated with a linear transformation in three dimensions: <ul style="list-style-type: none"> <li>• Reflection in <math>x = 0</math>, <math>y = 0</math> or <math>z = 0</math></li> <li>• Rotations through multiples of <math>90^\circ</math> about the <math>x</math>, <math>y</math> or <math>z</math> axes</li> </ul>	
Know the composite of the transformation represented by $\mathbf{M}$ followed by the transformation represented by $\mathbf{N}$ is represented by the matrix product $\mathbf{NM}$	
Find the invariant points for a linear transformation using $\mathbf{M} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x \\ y \end{pmatrix}$	
Find the invariant lines for a linear transformation, where the image of every point on AB is also on AB	

Specification: OCR Further Mathematics B (MEI) H645

Specification	Ref.	Learning outcomes	Notes	Notation	Exclusions	
<b>CORE PURE: MATRICES AND TRANSFORMATIONS (a)</b>						
Matrix addition and multiplication	Pm1	Be able to add, subtract and multiply conformable matrices, and to multiply a matrix by a scalar.	With and without a calculator for matrices up to $3 \times 3$ .	$\mathbf{M} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ .		
	m2	Understand and use the zero and identity matrices, understand what is meant by equal matrices.		$\mathbf{0}$ (zero) $\mathbf{I}$ (identity).		
	m3	Know that matrix multiplication is associative but not commutative.				
Linear transformations and their associated matrices	m4	Be able to find the matrix associated with a linear transformation and vice-versa.	2-D transformations include the following. <ul style="list-style-type: none"> <li>• Reflection in the <math>x</math> and <math>y</math> axes and in <math>y = \pm x</math>.</li> <li>• Rotation centre the origin through an angle <math>\theta</math> (counter clockwise positive)</li> <li>• Enlargement centre the origin</li> <li>• Stretch parallel to <math>x</math> or <math>y</math> axis</li> <li>• Shear <math>x</math> or <math>y</math> axis fixed, shear factor<sup>1</sup></li> </ul> 3-D transformations will be confined to reflection in one of $x = 0$ , $y = 0$ , $z = 0$ or rotation of multiples of $90^\circ$ about $x$ , $y$ or $z$ axis <sup>2</sup> . Learners should know that any linear transformation may be represented by a matrix.	Matrices will be shown in bold type, transformations in non-bold type. The image of the column vector $\mathbf{r}$ under the transformation associated with matrix $\mathbf{M}$ is $\mathbf{Mr}$ .		
	<sup>1</sup> A shear may be defined by giving the fixed line and the image of a point. (The fixed line of a shear is a line of invariant points.) The shear factor is the distance moved by a point divided by its perpendicular distance from the fixed line. Learners should know this, but the shear factor should not be used to define a shear as there are different conventions about the sign of a shear factor. <sup>2</sup> Positive angles counter clockwise when looking towards the origin from the positive side of the axis of rotation.					
	m5	Understand successive transformations in 2-D and the connection with matrix multiplication.	Describe a transformation as a combination of two of those above.			More than 2 dimensions.
*	Understand the language of vectors in two dimensions and three dimensions.	Scalar, vector, modulus, magnitude, direction, position vector, unit vector, cartesian components, equal vectors, parallel vectors.		$\mathbf{i}, \mathbf{j}, \mathbf{k}, \hat{\mathbf{r}}, \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix}$		

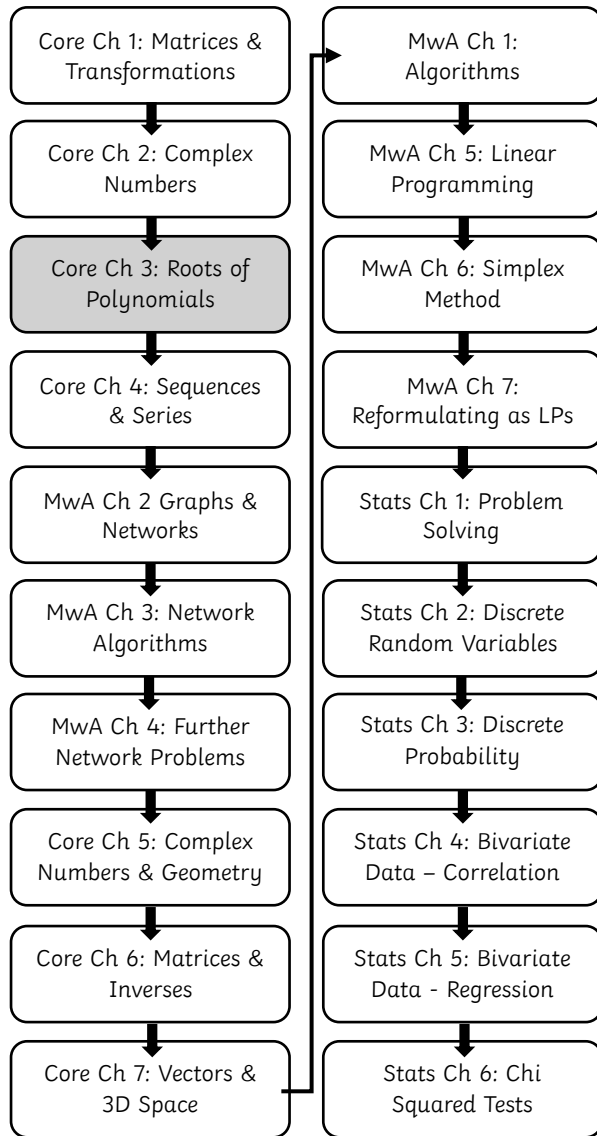
Specification	Ref.	Learning outcomes	Notes	Notation	Exclusions
<b>CORE PURE: MATRICES AND TRANSFORMATIONS (a)</b>					
Invariance	Pm6	Know the meaning of, and be able to find, invariant points and invariant lines for a linear transformation.			More than 2 dimensions.



Core 1 Chapter 2: Introduction to Complex Numbers	
Understand how complex numbers extend the number system	
Know that complex numbers have the form $z = x + yi$ with $i^2 = -1$ where $x$ is called the real part $Re(z)$ and $y$ is called the imaginary part $Im(z)$	
Solve quadratic equations with complex roots	
Know that the complex conjugate of $z = x + yi$ is $z^* = x - yi$	
Add and subtract complex numbers by adding or subtracting their real and imaginary parts separately	
Multiply complex numbers by expanding the brackets then simplify using $i^2 = -1$	
Solve problems involving complex numbers by equating real and imaginary parts	
Divide complex numbers by writing as a fraction, multiplying the top and bottom by the conjugate of the denominator, then simplifying the answer	
Know that complex numbers are equal only if both the real parts and the imaginary parts are equal	
Represent a complex number $x + yi$ on an Argand diagram as the point $(x, y)$	
Represent addition and subtraction of two complex numbers on an Argand diagram	

Specification	Ref.	Learning outcomes	Notes	Notation	Exclusions
<b>CORE PURE: COMPLEX NUMBERS (a)</b>					
Language of complex numbers	Pj1	Understand the language of complex numbers.	Real part, imaginary part, complex conjugate, modulus, argument, real axis, imaginary axis.	$z = x + yi$ $z^* = x - yi$ $\text{Re}(z) = x, \text{Im}(z) = y$	
Complex numbers and polynomial equations with real coefficients	j2	Be able to solve any quadratic equation with real coefficients.		$i^2 = -1$	
	j3	Know that the complex roots of polynomial equations with real coefficients occur in conjugate pairs. Be able to solve cubic or quartic equations with real coefficients.	Use of the factor theorem once a real root has been determined. Sufficient information will be given to deduce at least one complex root or quadratic factor for quartics.		Equations with degree $> 4$ .
Arithmetic of complex numbers	j4	Be able to add, subtract, multiply and divide complex numbers given in the form $x + yi$ , $x$ and $y$ real.	Division using complex conjugates.		
	j5	Understand that a complex number is zero if and only if both the real and imaginary parts are zero.			

Specification	Ref.	Learning outcomes	Notes	Notation	Exclusions
<b>CORE PURE: COMPLEX NUMBERS (a)</b>					
The Argand diagram	j9	Be able to represent and interpret complex numbers and their conjugates on an Argand diagram.			
	j10	Be able to represent the sum, difference, product and quotient of two complex numbers on an Argand diagram.			



<b>Core 1 Chapter 3: Roots of Polynomials</b>	
Know the relationships between the roots and coefficients of quadratic equations	
Know the relationships between the roots and coefficients of cubic equations	
Know the relationships between the roots and coefficients of quartic equations	
Form new equations whose roots are related to the roots of a given equation by a linear transformation	
Understand that complex roots of polynomial equations with real coefficients occur in conjugate pairs	
Solve cubic and quartic equations with complex roots	
Know the shorthand notation and summary formulae for roots of a polynomial of degree $n$ with roots $\alpha, \beta, \gamma, \delta$	

**Formulas to Learn**

For a polynomial with coefficients  $a, b, c \dots$  and roots  $\alpha, \beta, \gamma$  etc

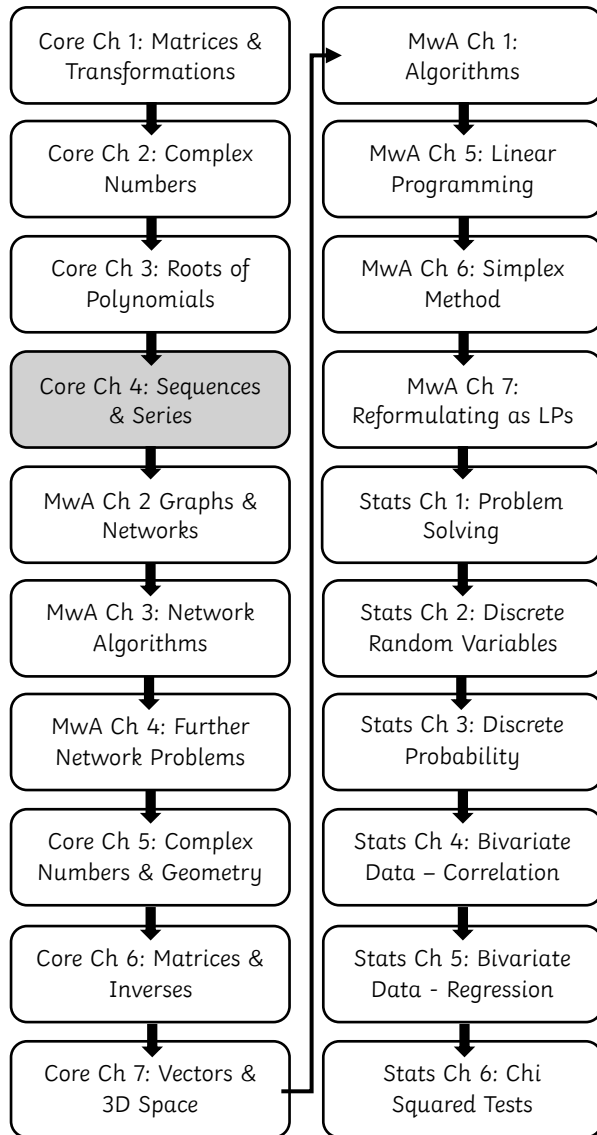
$$\sum \alpha = -\frac{b}{a}$$

$$\sum \alpha\beta = \frac{c}{a}$$

$$\sum \alpha\beta\gamma = -\frac{d}{a}$$

$$\sum \alpha\beta\gamma\delta = \frac{e}{a}$$

Specification	Ref.	Learning outcomes	Notes	Notation	Exclusions
<b>CORE PURE: ALGEBRA (a)</b>					
Relations between the roots and coefficients of polynomial equations	Pa1	Understand and use the relationships between the roots and coefficients of quadratic, cubic and quartic equations.		Roots $\alpha, \beta, \gamma, \delta$ .	Equations of degree $\geq 5$ .
	a2	Be able to form a new equation whose roots are related to the roots of a given equation by a linear transformation.	For a cubic or quartic equation.		Non-linear transformations of roots.



Core 1 Chapter 4: Sequences and Series	
Know what is meant by a sequence and a series	
Find the sum of a series using standard formulae for $\sum r$ , $\sum r^2$ , $\sum r^3$	
Find the sum of a series using the method of differences	
Use proof by induction to prove given results for the sum of a series	
Use proof by induction to prove given results for the nth term of a sequence	
Use proof by induction to prove given results for the nth power of a matrix	

**Formula Booklet Extract**

**Series**

$$\sum_{r=1}^n r^2 = \frac{1}{6}n(n+1)(2n+1) \quad \sum_{r=1}^n r^3 = \frac{1}{4}n^2(n+1)^2$$

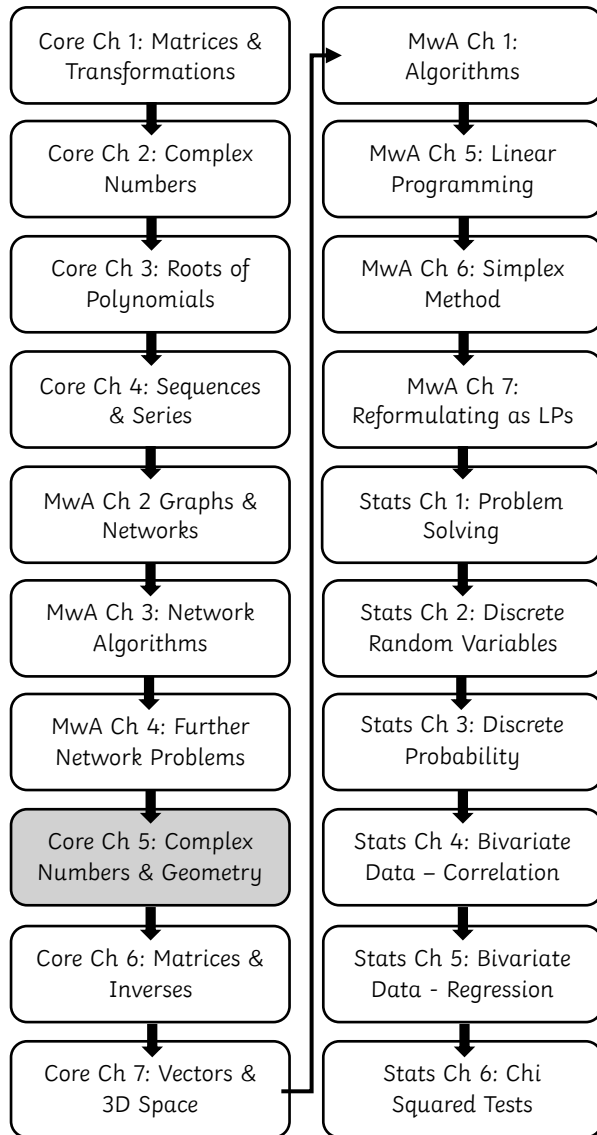
**Formulas to Learn**

$$\sum_{r=1}^n r = \frac{1}{2}n(n+1)$$

Specification: OCR Further Mathematics B (MEI) H645

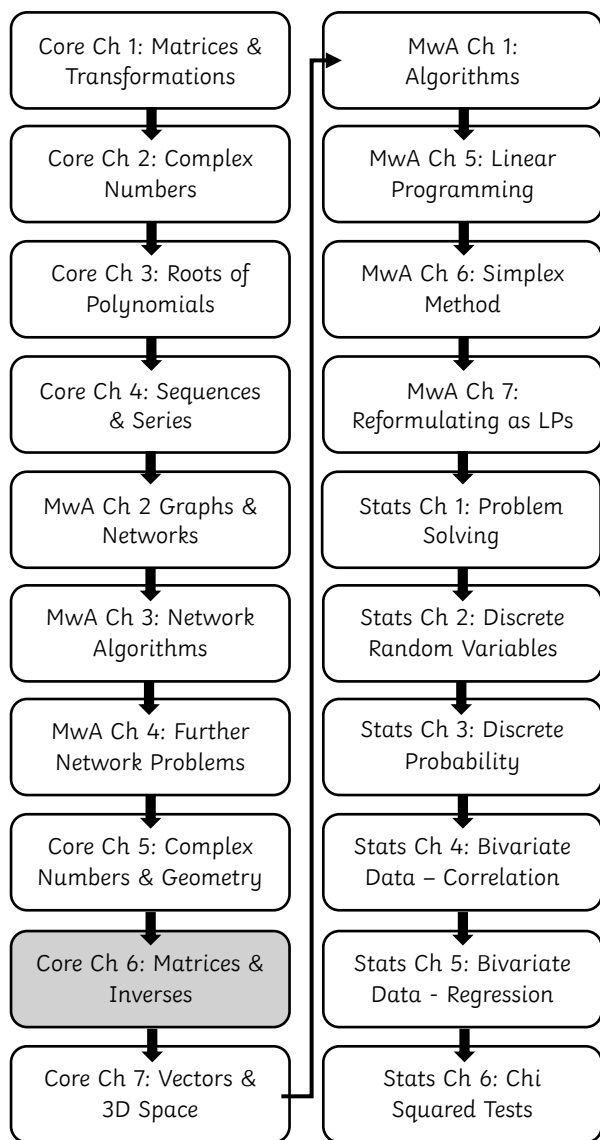
Specification	Ref.	Learning outcomes	Notes	Notation	Exclusions
<b>CORE PURE: SERIES (a)</b>					
Summation of series	Ps1	Be able to use standard formulae for $\Sigma r$ , $\Sigma r^2$ and $\Sigma r^3$ and the method of differences to sum series.	Formulae for $\Sigma r^2$ and $\Sigma r^3$ will be given but proof could be required, e.g. by induction.	$\sum_{r=1}^n r^2$	

Specification	Ref.	Learning outcomes	Notes	Notation	Exclusions
<b>CORE PURE: PROOF (a)</b>					
Proof	*	Be able to prove mathematical results by deduction and exhaustion, and disprove false conjectures by counter example.	Includes proofs of results used in this specification, where appropriate.		
Induction	Pp4	Be able to construct and present a proof using mathematical induction for given results for a formula for the $n$ th term of a sequence, the sum of a series or the $n$ th power of a matrix.	The result to be proved will be given. E.g. for the sequence given by $u_1 = 0$ , $u_{n+1} = u_n + 2n$ prove that $u_n = n^2 - n$ .	$u_n, \sum_{r=1}^n r^2$	



<b>Core 1 Chapter 5: Complex Numbers and Geometry</b>	
Find the modulus of a complex number $z = x + yi$ using $ z  = \sqrt{x^2 + y^2}$	
Find the principal argument of a complex number using radians, where $-\pi < \theta < \pi$	
Express a complex number in modulus-argument form $z = r(\cos \theta + i \sin \theta)$ where $r =  z $ and $\theta = \arg z$ - this can also be written $(r, \theta)$	
Multiply complex numbers in modulus-argument form using the rules $ z_1 z_2  =  z_1   z_2 $ and $\arg(z_1 z_2) = \arg(z_1) + \arg(z_2)$	
Divide complex numbers in modulus-argument form using the rules $\left  \frac{z_1}{z_2} \right  = \frac{ z_1 }{ z_2 }$ and $\arg\left(\frac{z_1}{z_2}\right) = \arg(z_1) - \arg(z_2)$	
Represent multiplication and division of two complex numbers on an Argand diagram	
Know that the distance between the points $z_1$ and $z_2$ on an Argand diagram is $ z_1 - z_2 $	
Represent and interpret sets of complex numbers as loci on an Argand diagram: <ul style="list-style-type: none"> <li>• <math> z - a  = r</math> is a circle with centre <math>a</math> and radius <math>r</math></li> <li>• <math>\arg(z - a) = \theta</math> is a half-line starting at <math>z = a</math> at an angle <math>\theta</math> to the real axis</li> <li>• <math> z - a  =  z - b </math> is the perpendicular bisector of the points <math>a</math> and <math>b</math></li> </ul>	
Represent and interpret regions defined by inequalities based on the above	

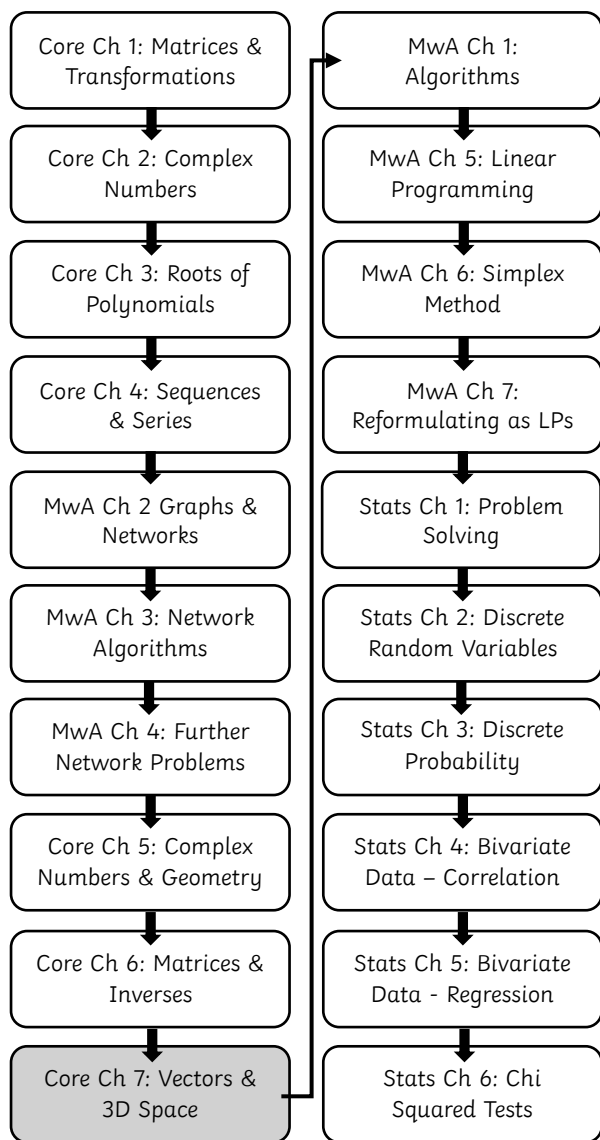
Specification	Ref.	Learning outcomes	Notes	Notation	Exclusions
<b>CORE PURE: COMPLEX NUMBERS (a)</b>					
Modulus-argument form	j6	Be able to use radians in the context of complex numbers.	Use exact values of trigonometric functions for multiples of $\frac{\pi}{4}$ and $\frac{\pi}{6}$ .		
	j7	Be able to represent a complex number in modulus-argument form. Be able to convert between the forms $z = x + yi$ and $z = r(\cos \theta + i \sin \theta)$ where $r$ is the modulus and $\theta$ is the argument of the complex number.	$zz^* =  z ^2$	$ z $ is the modulus of $z$ . $\arg z$ for principal argument, where $-\pi < \arg z \leq \pi$ . Radian measure.	
	j8	Be able to multiply and divide complex numbers in modulus-argument form.	$z_1 z_2 = r_1 r_2 [\cos(\theta_1 + \theta_2) + i \sin(\theta_1 + \theta_2)]$ $\frac{z_1}{z_2} = \frac{r_1}{r_2} [\cos(\theta_1 - \theta_2) + i \sin(\theta_1 - \theta_2)]$ The identities for $\sin(\theta \pm \phi)$ and $\cos(\theta \pm \phi)$ may be assumed in the derivation of these results.		
The Argand diagram	j9	Be able to represent and interpret complex numbers and their conjugates on an Argand diagram.			
	j10	Be able to represent the sum, difference, product and quotient of two complex numbers on an Argand diagram.			
	j11	Be able to represent and interpret sets of complex numbers as loci on an Argand diagram.	Circles of the form $ z - a  = r$ . Half lines of the form $\arg(z - a) = \theta$ . Lines of the form $ z - a  =  z - b $ . Regions defined by inequalities based on the above e.g. $ z - a  > r$ . Intersections and unions of these.	For regions defined by inequalities learners must state clearly which regions are included and whether the boundaries are included. No particular shading convention is expected.	$ z - a  = k z - b $ for $k \neq 1$ .



<b>Core 1 Chapter 6: Matrices and their Inverses</b>	
Find the determinant of a $2 \times 2$ matrix $M = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ using $\det M = ad - bc$	
Know that the determinant of a $2 \times 2$ matrix represents the area scale factor of the corresponding transformation	
Know that the determinant of a $3 \times 3$ matrix represents the volume scale factor of the corresponding transformation	
Know that a negative determinant indicates the orientation of the shape has reversed under a transformation and a positive determinant indicates the orientation is preserved.	
Understand that a zero determinant indicates all points are mapped to a line in 2D, or a plane in 3D	
Know that a singular matrix has a zero determinant and no inverse	
Know that a non-singular matrix has non-zero determinant and the inverse can be found	
Know that if $A$ is non-singular then $AA^{-1} = A^{-1}A = I$	
Know the inverse of a non-singular $2 \times 2$ matrix $M = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ is $M^{-1} = \frac{1}{ad-bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$	
Use a calculator to find the determinant and inverse of a $3 \times 3$ matrix	
Use the product rule for inverse matrices $(MN)^{-1} = N^{-1}M^{-1}$	
Use matrices to solve a pair of linear simultaneous equations in two unknowns	
Use matrices to solve three linear simultaneous equations in three unknowns	
Know that if $\det M = 0$ there is no unique solution to the equations, instead having no solution or an infinite number of solutions	

Specification	Ref.	Learning outcomes	Notes	Notation	Exclusions
<b>CORE PURE: MATRICES AND TRANSFORMATIONS (a)</b>					
Determinant of a matrix	m7	Be able to calculate the determinant of a $2 \times 2$ matrix and a $3 \times 3$ matrix. Know the meaning of the terms singular and non-singular as applied to matrices.	With a calculator for $3 \times 3$ matrices. A singular square matrix is non-invertible and therefore has determinant zero.	$\begin{vmatrix} a & b \\ c & d \end{vmatrix}$ or $\det \mathbf{M}$ or $ \mathbf{M} $ .	
	m8	Know that the magnitude of the determinant of a $2 \times 2$ matrix gives the area scale factor of the associated transformation, and understand the significance of a zero determinant. Interpret the sign of a determinant in terms of orientation of the image.	E.g. Quadrilateral ABCD is labelled clockwise and transformed in 2-D; a negative determinant for the transformation matrix means that the labelling on the image A'B'C'D' is anticlockwise.		Proof.
	m9	Know that the magnitude of the determinant of a $3 \times 3$ matrix gives the volume scale factor of the associated transformation, and understand the significance of a zero determinant. Interpret the sign of a determinant in terms of orientation of the image.	The sign of the determinant determines whether the associated transformation preserves or reverses orientation ('handedness').  E.g. If a triangle ABC is labelled clockwise when seen from point S, then for a negative determinant, the triangle A'B'C' is anti-clockwise when seen from S'.		Proof
	m10	Know that $\det(\mathbf{MN}) = \det \mathbf{M} \times \det \mathbf{N}$ and the corresponding result for scale factors of transformations.	Scale factors in 2-D only.		Algebraic proof.

Specification	Ref.	Learning outcomes	Notes	Notation	Exclusions
Inverses of square matrices	m11	Understand what is meant by an inverse matrix.	Square matrices of any order.	$\mathbf{M}^{-1}$	
	m12	Be able to calculate the inverse of a non-singular $2 \times 2$ matrix or $3 \times 3$ matrix.	With a calculator for $3 \times 3$ matrices. $\det(\mathbf{A}^{-1}) = \frac{1}{\det \mathbf{A}}$		
	m13	Be able to use the inverse of a non-singular $2 \times 2$ or $3 \times 3$ matrix. Relate the inverse matrix to the corresponding inverse transformation.	E.g. to solve a matrix equation and interpret in terms of transformations: find the pre-image of a transformation.		
	m14	Understand and use the product rule for inverse matrices.	$(\mathbf{AB})^{-1} = \mathbf{B}^{-1} \mathbf{A}^{-1}$		



Core 1 Chapter 7: Vectors and 3D Space	
Find the scalar product of two vectors using $a_1b_1 + a_2b_2$ or $a_1b_1 + a_2b_2 + a_3b_3$	
Use the scalar product to find the angle between two vectors using $\cos \theta = \frac{a \cdot b}{ a  b }$	
Know that two vectors are perpendicular if and only if their scalar product is zero	
Know that the plane with equation $n_1x + n_2y + n_3z + d = 0$ has a normal vector $\mathbf{n} = \begin{pmatrix} n_1 \\ n_2 \\ n_3 \end{pmatrix}$ and goes through a point with position vector $\mathbf{a}$ where $d = -\mathbf{a} \cdot \mathbf{n}$	
Find the equation of a plane in vector form using $(\mathbf{r} - \mathbf{a}) \cdot \mathbf{n} = 0$ or $\mathbf{r} \cdot \mathbf{n} = \mathbf{a} \cdot \mathbf{n}$	
Find the equation of a plane in Cartesian form using $n_1x + n_2y + n_3z + d = 0$	
Find the angle between two planes by finding the angle between their normals	
Know that three distinct planes in three dimensions will be arranged in one of five ways: <ul style="list-style-type: none"> <li>• meet at a unique point of intersection</li> <li>• three parallel planes</li> <li>• two parallel planes that are cut by a third to form two parallel lines</li> <li>• a sheaf of planes that intersect in a common line</li> <li>• a prism of planes in which each pair of planes meet in a straight line but there are no common points of intersection between the three planes</li> </ul>	
Understand how solving three linear simultaneous equations in three unknowns relates to finding the point of intersection of three planes in three dimensions	

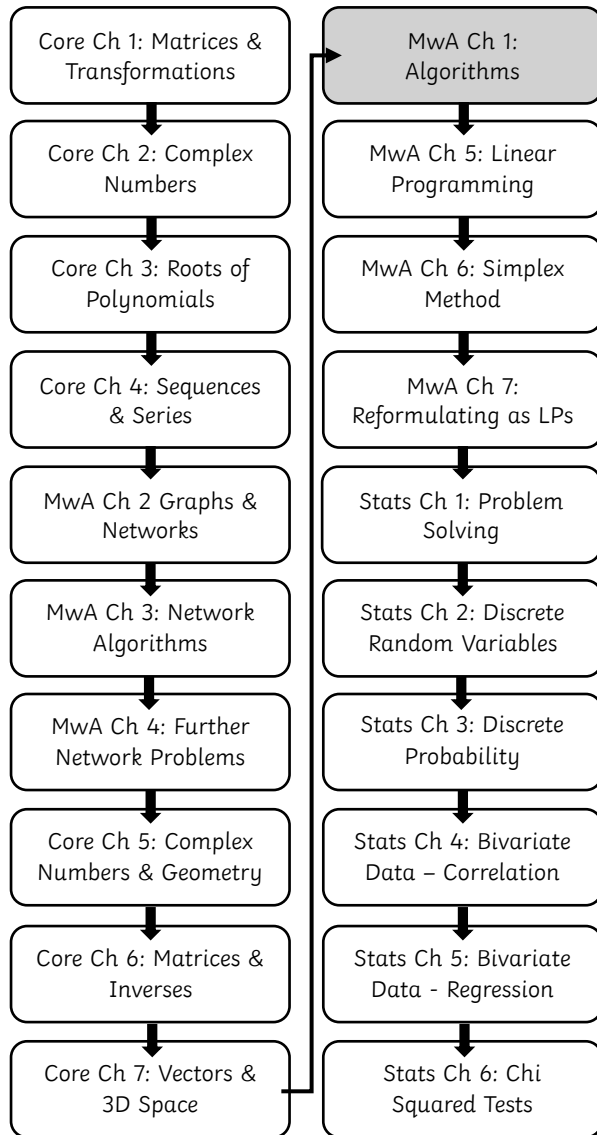
**Formula Sheet Extract**

Cartesian equation of a plane is  
 $n_1x + n_2y + n_3z + d = 0$

Cartesian equation of a line in 3-D is

$$\frac{x - a_1}{d_1} = \frac{y - a_2}{d_2} = \frac{z - a_3}{d_3}$$

Specification	Ref.	Learning outcomes	Notes	Notation	Exclusions
<b>CORE PURE: VECTORS AND 3-D SPACE (a)</b>					
Scalar products and the equations of planes	Pv1	Know how to calculate the scalar product of two vectors, and be able to use the two forms of the scalar product to find the angle between two vectors.	Including test for perpendicular vectors.	$\mathbf{a} \cdot \mathbf{b} = a_1 b_1 + a_2 b_2 + a_3 b_3 =  \mathbf{a}   \mathbf{b}  \cos \theta$	Proof of equivalence of two forms in general case.
	v2	Be able to form and use the vector and cartesian equations of a plane. Convert between vector and cartesian forms for the equation of a plane.	Plane: $(\mathbf{r} - \mathbf{a}) \cdot \mathbf{n} = 0$ $n_1 x + n_2 y + n_3 z + d = 0$ where $d = -\mathbf{a} \cdot \mathbf{n}$ .		The form $\mathbf{r} = \mathbf{a} + \lambda \mathbf{b} + \mu \mathbf{c}$
	v3	Know that a vector which is perpendicular to a plane is perpendicular to any vector in the plane.	If a vector is perpendicular to two non-parallel vectors in a plane, it is perpendicular to the plane.		
Intersection of planes	v4	Know the different ways in which three distinct planes can be arranged in 3-D space.	If two planes are parallel the third can be parallel or cut the other two in parallel lines; if no pair is parallel the planes can intersect in a point, form a sheaf or form a prismatic intersection.	A sheaf is where three planes share a common line. A prismatic intersection is where each pair of planes meets in a line; the three lines are parallel.	
	v5	Be able to solve three linear simultaneous equations in three variables by use of the inverse of the corresponding matrix. Interpret the solution or failure of solution geometrically in terms of the arrangement of three planes. Be able to find the intersection of three planes when they meet in a point.	Inverse obtained using a calculator. If the corresponding matrix is singular, learners should know the possible arrangements of the planes; they will be given extra information or guidance if required to distinguish between these arrangements.		Finding equation of lines of intersection of two planes.
	v6	Know that the angle between two planes can be found by considering the angle between their normals.	The angle between two non-perpendicular planes is the acute angle between them.		

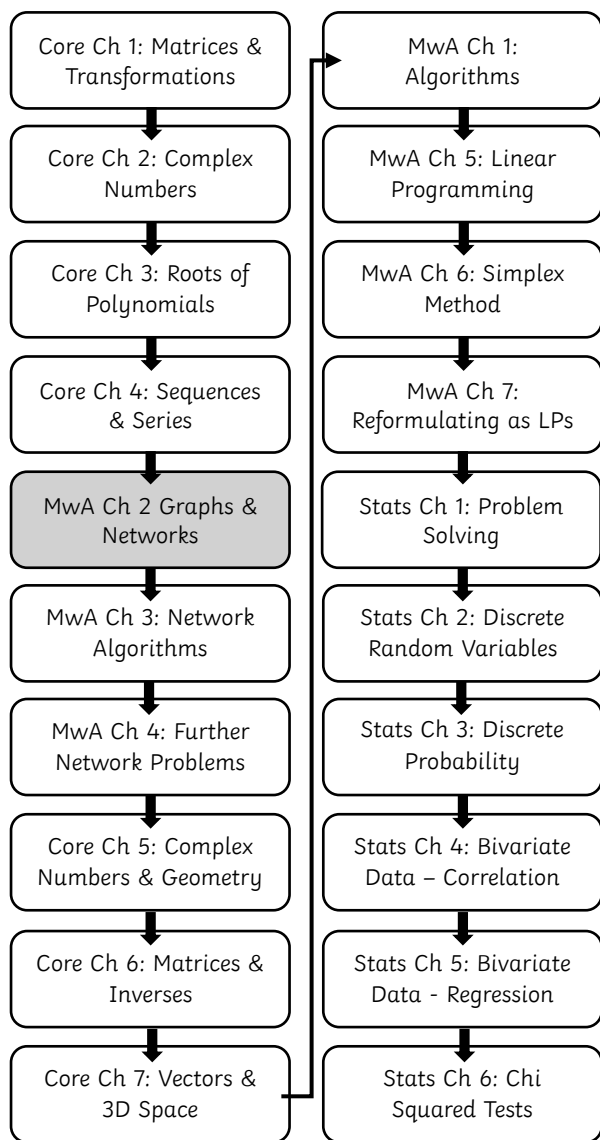


<b>Algorithms Chapter 1: Algorithms</b>		
Understand that an algorithm is a finite sequence of operations for carrying out a procedure or solving a problem – it must be unambiguous, deterministic and finite		
Know the terms finite, initial state (starting values), input, output and variable in the context of algorithms		
Be able to interpret and apply algorithms presented in a variety of formats		
Be able to repair, develop and adapt given algorithms		
Understand the use of loops, passes, decisions and the idea of an iterative process		
Understand and be able to use the basic idea of algorithmic complexity and be able to analyse the complexity of given algorithms		
Know that complexity can be used, among other things, to compare algorithms		
Know that for an algorithm with linear complexity, $O(n)$ , doubling the size will increase the approximate run-time by a scale factor of 2		
Know that for an algorithm with quadratic complexity, $O(n^2)$ , doubling the size will increase the approximate run-time by a scale factor of 4		
Know that for an algorithm with cubic complexity, $O(n^3)$ , doubling the size will increase the approximate run-time by a scale factor of 8		
Understand that algorithms can sometimes be proved correct or incorrect		
Know that heuristic algorithms usually provide a good/efficient solution, but not necessarily the optimal solution		
Know and be able to use the quick sort algorithm		
Be able to apply other sorting algorithms which are specified		
Be able to count the number of comparisons and/or swaps needed in particular applications of sorting algorithms, and relate this to complexity		
Be able to reason about a given sorting algorithm		
Know and be able to use first-fit and first-fit decreasing packing algorithms and full bin strategies		
Be able to count the number of comparisons needed in particular applications of packing algorithms, and relate this to complexity		

Specification: OCR Further Mathematics B (MEI) H645

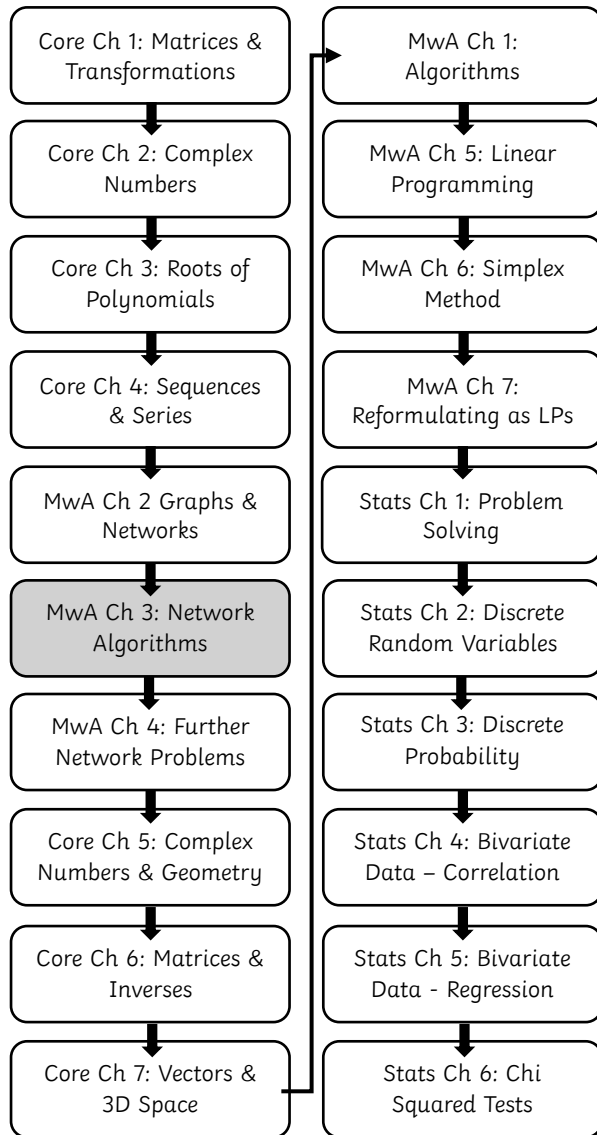
Specification	Ref.	Learning outcomes	Notes	Notation	Exclusions
<b>MODELLING WITH ALGORITHMS: ALGORITHMS</b>					
Algorithms	A1	Understand that an algorithm is a finite sequence of operations for carrying out a procedure or solving a problem. Understand that an algorithm can be the basis for a computer program.	Initial state; input; output; variable. 'Finite' means that the procedure terminates.		Algorithms with a random element.
	A2	Be able to interpret and apply algorithms presented in a variety of formats.	Formats include flowcharts; written English; pseudocode. E.g. in pseudocode, Let $i = i + 1$ means that the number in location $i$ is replaced by its current value plus 1. Questions will not be set requiring unduly repetitive calculations.	Loop, pass. 'if ... then...' 'Go to step ...' Iterative process.	Any particular version of pseudocode or programming language.
	A3	Be able to repair, develop and adapt given algorithms.			
	A4	Understand and be able to use the basic ideas of algorithmic complexity and be able to analyse the complexity of given algorithms. Know that complexity can be used, among other things, to compare algorithms.	Worst case; size of problem; effect on solution time of multiplying the size of a large problem by a given factor and/or repeatedly applying an algorithm.	Order notation e.g. $O(n^2)$ for quadratic complexity.	Analysis leading to non-polynomial complexity.
	A5	Understand that algorithms can sometimes be proved correct or incorrect.	Proof by exhaustion and disproof by counter-example.		

Specification	Ref.	Learning outcomes	Notes	Notation	Exclusions
<b>MODELLING WITH ALGORITHMS: ALGORITHMS</b>					
Algorithms (cont)	A6	Understand and know the importance of heuristics.	A heuristic (sometimes called a heuristic algorithm) is a method which finds a solution efficiently, with no guarantee that it is optimal. It is important when classic methods are inefficient or fail.	E.g. packing algorithms. E.g. find a solution to a linear problem which requires an integer solution by exploring around the solution to the corresponding LP.	
Sorting algorithms	A7	Know and be able to use the quick sort algorithm. Be able to apply other sorting algorithms which are specified.		Pivot values. Pass. Ascending, descending.	
	A8	Be able to count the number of comparisons and/or swaps needed in particular applications of sorting algorithms, and relate this to complexity.	Quick sort algorithm has (worst case) complexity $O(n^2)$ .		Average complexity.
	A9	Be able to reason about a given sorting algorithm.	E.g. explain why it will always work.		
Packing algorithms	A10	Know and be able to use first fit and first fit decreasing packing algorithms and full bin strategies.	Know that these are not guaranteed to be optimal.	Bin.	
	A11	Be able to count the number of comparisons needed in particular applications of packing algorithms, and relate this to complexity.	First fit and first fit decreasing algorithms have (worst case) complexity $O(n^2)$ .		



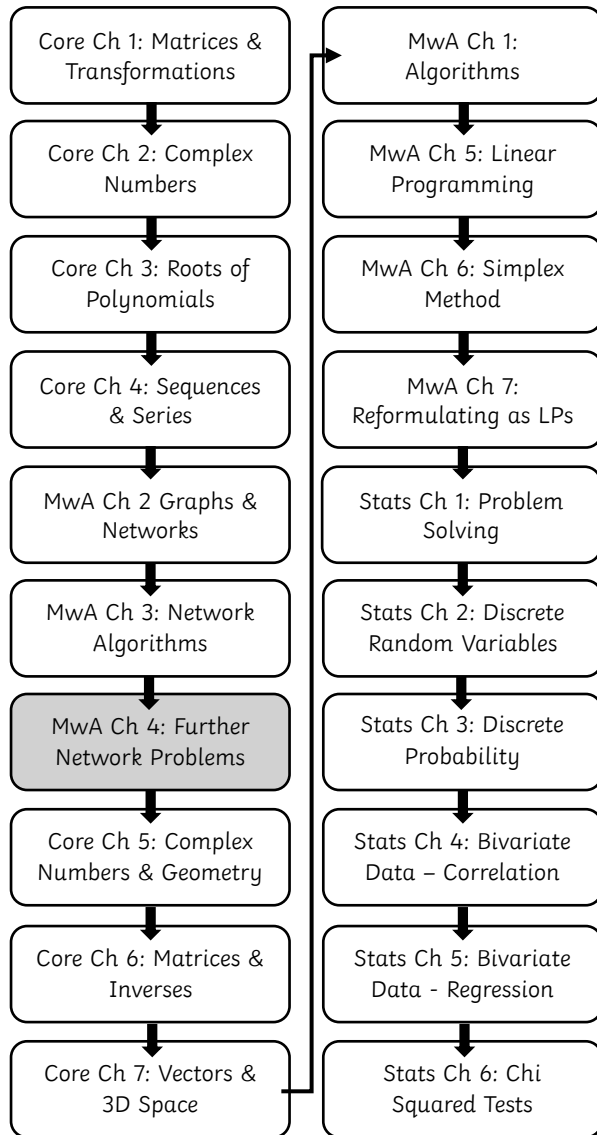
<b>Algorithms Chapter 2: Graphs and Networks</b>	
Understand and be able to use graphs and associated language (node/vertex, arc/edge, simple, connected, complete, bipartite, tree, digraph)	
Know that a graph is made up of vertices/nodes and edges/arcs which connect vertices and represent a relationship between the vertices	
Know that a loop is an edge that goes from one vertex back to itself again	
Know that the degree/order of a vertex is the number of edges going into it (so a loop contributes two)	
Know that a connected graph has every vertex joined directly or indirectly to each of the other vertices	
Know that a simple graph has no loops and only one edge connecting each pair of vertices (no repeated edges)	
Know that a tree is a simple, connected graph with the minimum number of edges: a tree with $n$ vertices has $n - 1$ edges	
Know that a complete graph is one in which every pair of vertices is connected by an edge so a complete graph on $n$ vertices has $\frac{1}{2}n(n - 1)$ edges	
Know that a bipartite graph has vertices in two sets, with edges only allowed to join a vertex from one set to a vertex in the other set	
Know that a digraph is a directed graph, in which at least one edge has a direction (usually indicated by an arrow)	
Know that an incidence matrix is a matrix used to represent a graph by showing the number of edges connecting each pair of vertices	
Know that two graphs are isomorphic if they have the same incidence matrix, showing the same connections between vertices	
Know that a network is a graph with weighted arcs, which can also be written in a table	
Use incidence matrices to represent graphs	
Model problems by using graphs	
Model problems by using networks	

Specification	Ref.	Learning outcomes	Notes	Notation	Exclusions
<b>MODELLING WITH ALGORITHMS: NETWORKS</b>					
Networks and graphs	AN1	Understand and be able to use graphs and associated language.	Node/vertex; arc/edge; tree; order of a node; simple, complete, connected and bipartite graphs; trees; digraphs.	Incidence matrix.	
	N2	Be able to model problems by using graphs.	E.g. river crossing problems. E.g. matching problems.		
	N3	Understand that a network is a graph with weighted arcs.	Directed and undirected networks.		
	N4	Be able to model problems by using networks.	E.g. shortest path, maximum flow. E.g. allocation and transportation problems.		



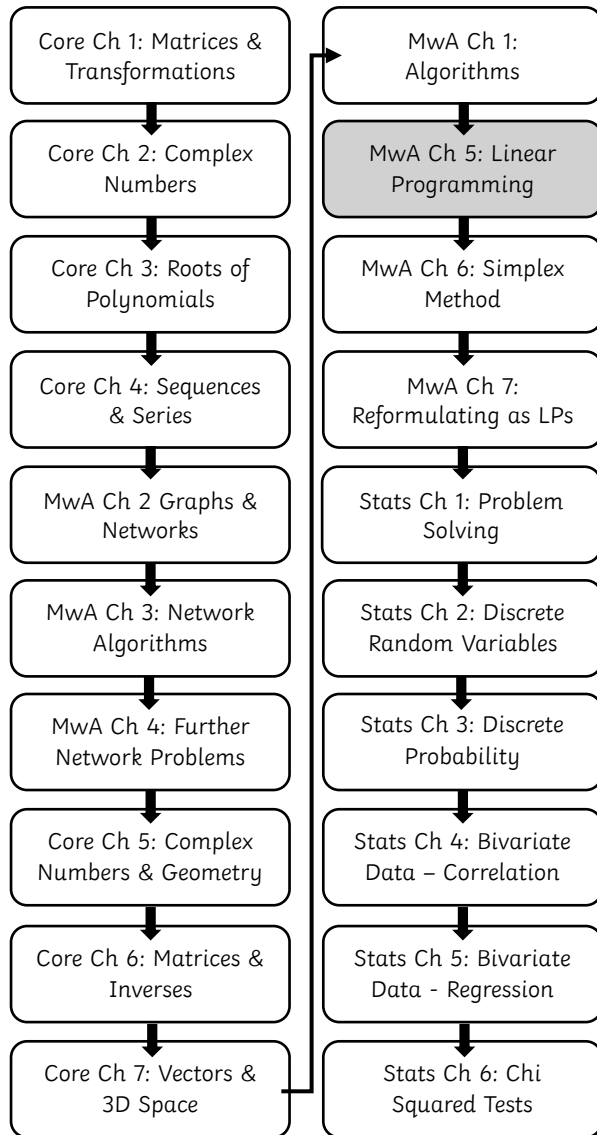
<b>Algorithms Chapter 3: Network Algorithms</b>	
Solve minimum connector problems using Kruskal's algorithm, where arcs are added in increasing order of weight	
Solve minimum connector problems from a graph using Prim's algorithm, where arcs are added in order of weight, to those already connected	
Solve minimum connector problems from a table using Prim's algorithm	
Model shortest path problems and solve using Dijkstra's algorithm	
Know that solutions to minimum connector problems are trees, and state the total weight of the tree	
Know and use the fact that Kruskal's, Prim's and Dijkstra's algorithms have quadratic complexity (in relation to the number of arcs)	

Specification	Ref.	Learning outcomes	Notes	Notation	Exclusions
<b>MODELLING WITH ALGORITHMS: NETWORKS</b>					
Kruskal's, Prim's and Dijkstra's algorithms	N5	Be able to solve minimum connector problems using Kruskal's and Prim's algorithms.	Kruskal's algorithm in graphical form only. Prim's algorithm in graphical or tabular form.	Minimum spanning tree.	
	N6	Model shortest path problems and solve using Dijkstra's algorithm.			
	N7	Know and use the fact that Kruskal's, Prim's and Dijkstra's algorithms have quadratic complexity.			



<b>Algorithms Chapter 4: Further Network Problems</b>	
Model precedence problems with an activity-on-arc network	
Understand dummy activities may be needed to establish a precedence or ensure arcs can be uniquely labelled using the events at their ends	
Use dummy activities, showing them with a directed, dashed line with zero duration	
Carry out a forward pass to establish the earliest time for each event	
Carry out a backward pass to establish the latest time for each event	
Know that the minimum time to complete a project is the minimum completion time and is the length of the longest path	
Know that any activity on the longest path is called a critical activity and the path is the critical path – any delay on a critical activity will delay the whole project	
Interpret outcomes from critical path analysis, including implications for criticality	
Analyse float (total, independent and interfering), resourcing and scheduling	
Draw or interpret a cascade chart to model a problem, using a timescale to show the scheduling of activities at their earliest start times	
Draw or interpret a resource histogram, showing how many workers are needed at each time	
Draw a schedule to show which activity each worker is doing at each time	
Model a transmission system as a network flow problem	
Use a single supersource to supply multiple sources or a single supersink to drain all the sinks	
Specify a cut, which is a partition of the vertices into two sets, one containing the source and one containing the sink	
State the capacity of a cut, the maximum possible flow across a cut, by adding up the values of all the cut arcs that can flow from S to T (arcs from T to S contribute 0)	
Know and use maximum flow/minimum cut theorem, which says that the maximum flow is equal to the minimum cut	

Specification	Ref.	Learning outcomes	Notes	Notation	Exclusions
<b>MODELLING WITH ALGORITHMS: NETWORKS</b>					
Network flows	N10	Be able to use a network to model a transmission system.	Single and super sources and sinks. Flow in = flow out for other nodes.	Source: S. Sink: T.	
	N11	Be able to specify a cut and calculate its capacity.	<i>Either</i> split the vertices into two sets, one containing S and the other T, <i>or</i> specify the arcs that are cut.		
	N12	Understand and use the maximum flow/minimum cut theorem.	If an established flow is equal to the capacity of an identified cut, then the flow is maximal and the cut is a minimum cut. Exhaustive testing of cuts will not be assessed.		Flow augmentation. Labelling algorithm.
Solving network problems using technology	N13	Understand that network algorithms can be explored, understood and tested in cases in which the algorithm can be run by hand, but for practical problems the algorithm needs to be formulated in a way suitable for computing power to be applied.	Formulations will be restricted to LPs. Questions may be set about the time taken by computer software to implement an algorithm when its complexity is known.		

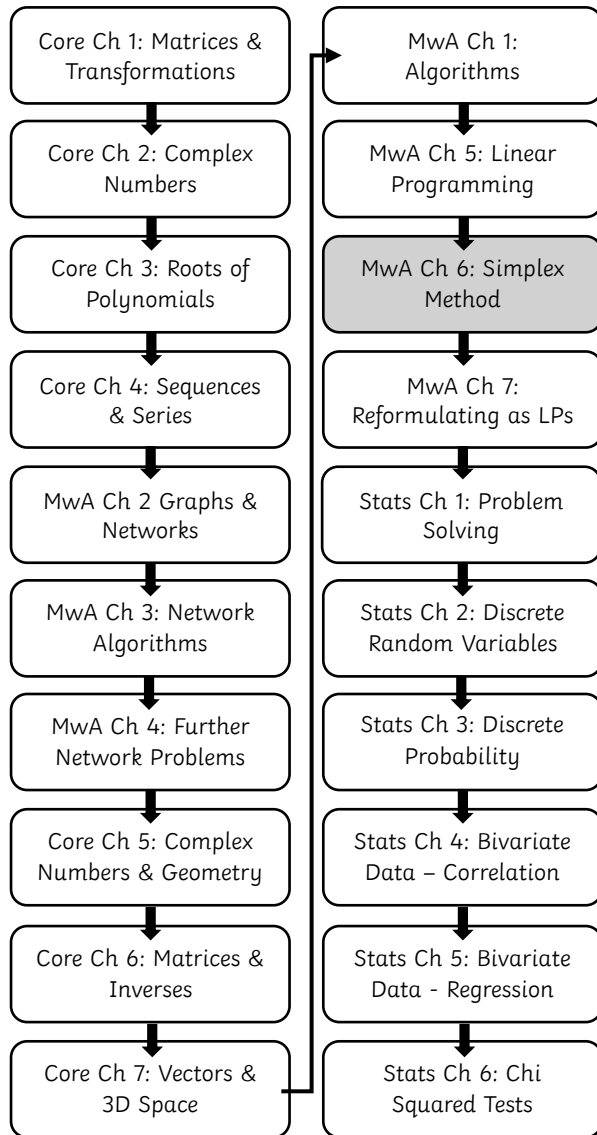


<b>Algorithms Chapter 5: Linear Programming</b>	
Formulate a linear programming problem in standard form: in terms of non-negative variables as a linear objective function to be maximised subject to linear constraints, each of which is less than or equal to a non-negative constant	
Use non-negative slack variables to convert an LP in standard form to augmented form, with constraints becoming equations	
Solve a 2D LP problem graphically	
Know that the feasible region is the set of points which satisfy all the constraints	
Know the vertices of the feasible region are called basis feasible solutions, and that two of the variables will be 0 here	
Know that variables which are 0 are called non-basic variables and non-zero variables are called basic variables	
Know that the optimal solution can be found using a profit line or by checking each vertex	
Recognise when an LP is an integer linear programming problem with integer variables	
Solve simple 2D integer LP problems graphically and find the solution at a grid point within the feasible region	
Consider the effect of modifying constraints or the objective function	
Reduce a 3D LP to a 2D LP when one constraint is an equality, and solve graphically	

Specification: OCR Further Mathematics B (MEI) H645

Specification	Ref.	Learning outcomes	Notes	Notation	Exclusions
<b>MODELLING WITH ALGORITHMS: LINEAR PROGRAMMING</b>					
Formulating a problem	AL1	Understand and use the language associated with linear programming.	Linear programming, objective, maximisation, minimisation, optimisation, constraints.	LP is an abbreviation for linear program.	
	L2	Be able to identify and define variables from a given problem. Be able to formulate a problem as a linear program.	Variables should be clearly identified as representing numerical values. E.g. 'Let $x$ be the number of ...'. Problem may be given in context.		
	L3	Be able to recognise when an LP is in standard form.	A linear function to be maximised, constraints with ' $\leq$ constant' and non-negative, continuous variables.		
	L4	Be able to use slack variables to convert an LP in standard form to augmented form.	Also called slack form. As standard form, but using non-negative slack variables to convert inequalities to equalities.	State variables. Slack variables. Basic and non-basic variables.	
	L5	Recognise when an LP requires an integer solution.	E.g. when a variable is discrete. E.g. a shortest path problem, because the variables take the values 1 or 0, depending on whether the corresponding arc is in the path or not. If an LP requires an integer solution this should be stated in the formulation.	ILP is an abbreviation for integer LP.	
	L6	Be able to formulate a range of network problems as LPs.	Shortest path problems; network flows; critical path (longest path) problems; matching, allocation and transportation problems. See after L18 for examples.		

Specification	Ref.	Learning outcomes	Notes	Notation	Exclusions
<b>MODELLING WITH ALGORITHMS: LINEAR PROGRAMMING</b>					
Graphical solution of an LP	L7	Be able to graph inequalities in 2-D and identify feasible regions. Be able to recognise infeasibility.	No particular shading convention is expected, but learners must make clear which is the feasible region.		Drawing diagrams in more than 2-D.
	L8	Be able to solve a 2-D LP graphically.	By finding at least one optimal feasible point and the value of the objective function at this point. Using the gradient of the objective function or by enumeration.		
	L9	Be able to consider the effect of modifying constraints or the objective function.		Post-optimal analysis.	
	L10	Be able to solve 2-D integer LP problems graphically.	The optimal lattice point may or may not be near the LP solution.		
	L11	Be able to use a visualisation of a 3-D LP to solve it. Be able to reduce a 3-D LP to a 2-D LP when one constraint is an equality.	Diagram will be given. Regions will be defined by an inequality based on the cartesian equation of a plane.		

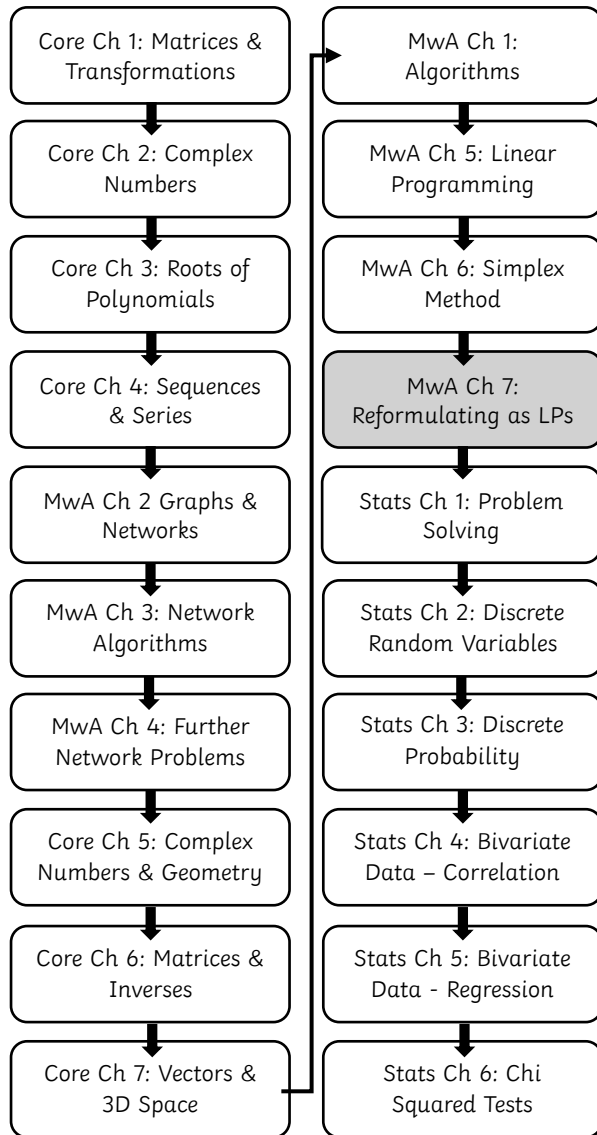


<b>Algorithms Chapter 6: Simplex Method</b>	
Formulate a linear programming problem as an initial simplex tableau	
Use a ratio test to choose a pivot and carry out an iteration of the simplex algorithm	
Recognise that an optimal solution has been achieved when there are no negative entries in the objective row	
Understand the geometrical basis for the simplex method	
Know that a degenerate problem is one where more than one vertex is optimal, and simplex will find one of the solutions	
Know that if there is a non-basic column with a 0 in the objective row, iterating on a pivot from this column will lead to a different solution	
Solve problem in non-standard form using a substitution	
Deal with negative variables and minimisation problems	
Understand how to use the two-stage simplex method to deal with greater than constraints by using surplus variables and artificial variables	
Understand how to use the two-stage simplex method to deal with equality constraints, by replacing them with two inequalities, one using $\leq$ and one using $\geq$	
Recognise that the artificial variable has been minimised when the right hand side is 0	
Interpret the output from computer software or spreadsheet	

Specification: OCR Further Mathematics B (MEI) H645

Specification	Ref.	Learning outcomes	Notes	Notation	Exclusions
<b>MODELLING WITH ALGORITHMS: LINEAR PROGRAMMING</b>					
Simplex method	L12	Be able to use the simplex algorithm on an LP in augmented form.	Setting up an initial tableau, choosing a pivot, transforming the tableau, interpreting a tableau, recognising when a tableau represents an optimal solution. Problems may be infeasible or have multiple solutions (degeneracy).	Initial, intermediate, final tableau. Slack variables. Pivot. Basic/non-basic variables.	Knowledge of complexity of the simplex algorithm.
	L13	Understand the geometric basis for the simplex method.	Interpret a tableau in terms of the vertex and value of the objective function.		

Specification	Ref.	Learning outcomes	Notes	Notation	Exclusions
<b>MODELLING WITH ALGORITHMS: LINEAR PROGRAMMING</b>					
Simplex and non-standard form	L14	Recognise that if an LP includes $\geq$ constraints then the two-stage simplex method may be used; understand how this method works and be able to set up the initial tableau in such cases.			Big-M method.
	L15	Be able to reformulate an equality constraint as a pair of inequality constraints.	E.g. replace $x = 4$ by $x \geq 4$ and $x \leq 4$ .		
	L16	Recognise that if an LP has variables which may take negative values or requires the objective function to be minimised then some initial reformulation is required before the simplex algorithm may be applied.	Learners need only know that such reformulation is possible.		Be able to apply simplex in these situations.
Use of software	L17	Understand that some LPs can be solved using graphical techniques or the simplex method, but for practical problems computing power needs to be applied. Know that a spreadsheet LP solver routine, or other software, can solve an LP given in standard form or, in some cases, in non-standard form.			
	L18	Be able to interpret the output from a spreadsheet optimisation routine, or other software, for the simplex method or ILPs.	Select the appropriate information to solve the original problem. This may lead to further analysis of the problem.		



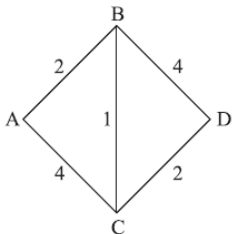
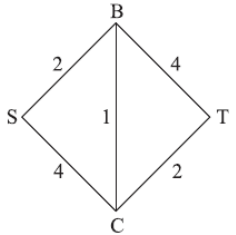
<b>Algorithms Chapter 7: Reformulating network problems as LPs</b>		
Know that an indicator variable records whether the variable is 'off' = 0 or 'on' = 1		
Formulate a shortest path problem as a linear programming problem: Minimise $\sum$ (arc weight x indicator) summed over all arcs Subject to $\sum$ (indicators from start vertex) = 1 $\sum$ (indicators to finish vertex) = 1 $\sum$ (indicators to vertex - indicators from vertex) = 0 for other vertices and each indicator takes the value 0 or 1		
Formulate a network flow problem as a linear programming problem: Maximise $\sum$ (flows from source) Subject to $\sum$ (flow to vertex - flow from vertex) = 0 for all vertices apart from source and sink $0 \leq \text{flow} \leq \text{capacity}$ for each arc		
Formulate a longest path problem as a linear programming problem Maximise $\sum$ (arc weight x indicator) summed over all arcs Subject to $\sum$ (indicators from start vertex) = 1 $\sum$ (indicators to finish vertex) = 1 $\sum$ (indicators to vertex - indicators from vertex) = 0 $\sum$ (indicators to vertex) $\leq$ 1 for other vertices and each indicator takes the value 0 or 1		
Formulate matching problems as linear programming problems Maximise $\sum$ (indicators) summed over all arcs Subject to $\sum$ (indicators from vertex) $\leq$ 1 for each vertex $\sum$ (indicators to vertex) $\leq$ 1 for each vertex and each indicator takes the value 0 or 1		
Formulate allocation problems as linear programming problems Minimise $\sum$ (arc weight x indicator) summed over all arcs Subject to $\sum$ (indicators from vertex) $\leq$ 1 for each 'out' vertex $\sum$ (indicators to vertex) $\leq$ 1 for each 'in' vertex and each indicator takes the value 0 or 1		
Formulate transportation problems as linear programming problems Minimise $\sum$ (arc weight x indicator) summed over all arcs Subject to $\sum$ (indicators from vertex) $\leq$ availability for each 'out' vertex $\sum$ (indicators to vertex) $\leq$ demand for each 'in' vertex and each indicator takes the value 0 or 1		

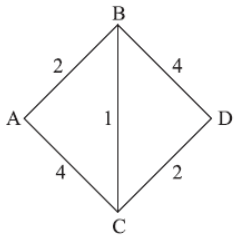
Specification: OCR Further Mathematics B (MEI) H645

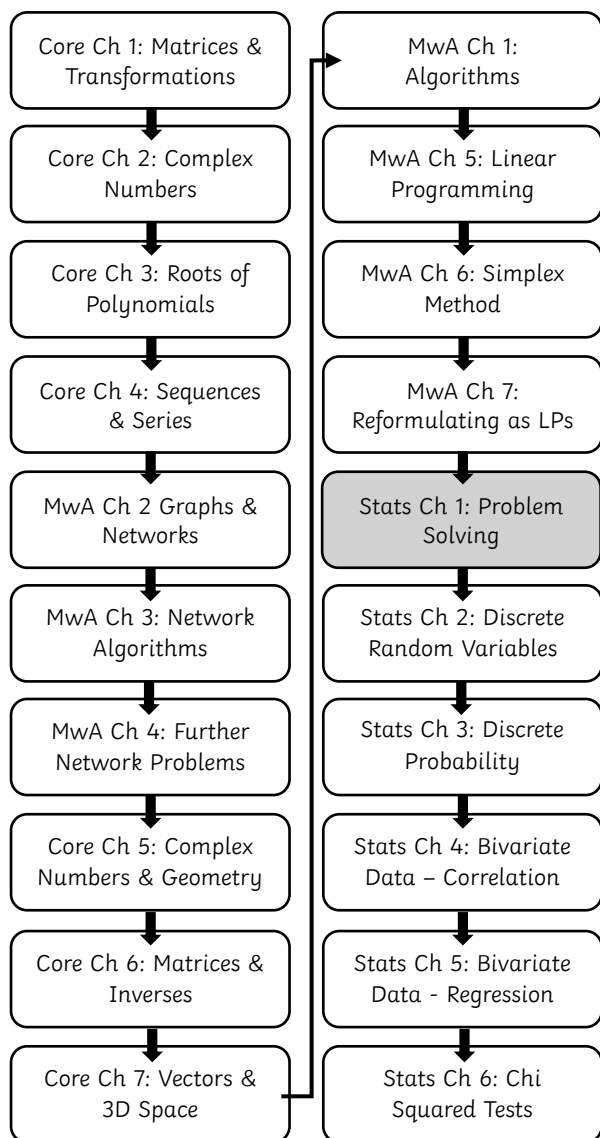
Specification	Ref.	Learning outcomes	Notes	Notation	Exclusions
<b>MODELLING WITH ALGORITHMS: LINEAR PROGRAMMING</b>					
Use of software	L17	Understand that some LPs can be solved using graphical techniques or the simplex method, but for practical problems computing power needs to be applied. Know that a spreadsheet LP solver routine, or other software, can solve an LP given in standard form or, in some cases, in non-standard form.			
	L18	Be able to interpret the output from a spreadsheet optimisation routine, or other software, for the simplex method or ILPs.	Select the appropriate information to solve the original problem. This may lead to further analysis of the problem.		

**Examples of reformulating network problems as LPs**

These examples show how six types of network problems can be reformulated as LPs. They illustrate the sort of notation that will be used in questions. They do not show the level of difficulty of problem that will be examined.

		<p>Shortest path</p> <p>Find a shortest path from A to D. Variables take the value 1 if the corresponding arc is used in a shortest path, and 0 otherwise.</p>		<p>Minimise</p> $2AB + 4BD + 4AC + 2CD + BC + CB$ <p>subject to</p> $AB + AC = 1$ $AB + CB - BC - BD = 0$ $AC + BC - CB - CD = 0$ $BD + CD = 1$
		<p>Network flow</p> <p>Find a maximum flow from S to T through the network.</p>		<p>Maximise</p> $SB + SC$ <p>subject to</p> $SB + CB - BC - BT = 0$ $SC + BC - CB - CT = 0$ $SB \leq 2$ $BT \leq 4$ $SC \leq 4$ $CT \leq 2$ $BC \leq 1$ $CB \leq 1$

	<p>Longest path Find a longest path from A to D. Variables take the value 1 if the corresponding arc is used in a shortest path, and 0 otherwise. This can be used to solve critical path problems on a directed network.</p>		<p>Maximise <math>2AB + 4BD + 4AC + 2CD + BC + CB</math> subject to <math>AB + AC = 1</math> <math>AB + CB - BC - BD = 0</math> <math>AC + BC - CB - CD = 0</math> <math>BD + CD = 1</math> <math>AB \leq 1</math> <math>BD \leq 1</math> <math>AC \leq 1</math> <math>CD \leq 1</math> <math>BC \leq 1</math> <math>CB \leq 1</math></p>																																				
	<p>Matching problem Possible associations between elements of {A, B, C, D} and {1, 2, 3, 4} are shown in the table. In a matching each element of one set is associated with at most one element of the other. The LP tries to find a maximal matching, i.e. a matching with as many associations as possible. Each variable (e.g. C3) takes the value 1 (if C and 3 are associated) or 0.</p>	<table border="1" data-bbox="674 507 1005 663"> <thead> <tr> <th></th> <th>1</th> <th>2</th> <th>3</th> <th>4</th> </tr> </thead> <tbody> <tr> <th>A</th> <td>x</td> <td></td> <td></td> <td>x</td> </tr> <tr> <th>B</th> <td>x</td> <td></td> <td>x</td> <td></td> </tr> <tr> <th>C</th> <td></td> <td>x</td> <td>x</td> <td></td> </tr> <tr> <th>D</th> <td></td> <td></td> <td>x</td> <td></td> </tr> </tbody> </table> <p>“x” indicates a possible matching</p>		1	2	3	4	A	x			x	B	x		x		C		x	x		D			x		<p>Maximise <math>A1 + A4 + B1 + B3 + C2 + C3 + D3</math> subject to <math>A1 + A4 \leq 1</math> <math>B1 + B3 \leq 1</math> <math>C2 + C3 \leq 1</math> <math>D3 \leq 1</math> <math>A1 + B1 \leq 1</math> <math>C2 \leq 1</math> <math>B3 + C3 + D3 \leq 1</math> <math>A4 \leq 1</math></p>											
	1	2	3	4																																			
A	x			x																																			
B	x		x																																				
C		x	x																																				
D			x																																				
	<p>Allocation problem This is like a matching problem, except that (usually) every association is possible, and each association has a cost. The LP minimises the total cost for a maximal matching. Each variable (e.g. A1) takes the value 1 or 0, depending on whether A is associated with 1 or not in the matching.</p>	<table border="1" data-bbox="674 876 1005 1032"> <thead> <tr> <th></th> <th>1</th> <th>2</th> <th>3</th> <th>4</th> </tr> </thead> <tbody> <tr> <th>A</th> <td>5</td> <td>2</td> <td>3</td> <td>6</td> </tr> <tr> <th>B</th> <td>1</td> <td>7</td> <td>2</td> <td>4</td> </tr> <tr> <th>C</th> <td>5</td> <td>8</td> <td>3</td> <td>1</td> </tr> <tr> <th>D</th> <td>4</td> <td>4</td> <td>2</td> <td>6</td> </tr> </tbody> </table>		1	2	3	4	A	5	2	3	6	B	1	7	2	4	C	5	8	3	1	D	4	4	2	6	<p>Minimise <math>5A1 + 2A2 + 3A3 + 6A4 + B1 + 7B2 + 2B3 + 4B4 + 5C1 + 8C2 + 3C3 + C4 + 4D1 + 4D2 + 2D3 + 6D4</math> subject to <math>A1 + A2 + A3 + A4 = 1</math> <math>B1 + B2 + B3 + B4 = 1</math> <math>C1 + C2 + C3 + C4 = 1</math> <math>D1 + D2 + D3 + D4 = 1</math> <math>A1 + B1 + C1 + D1 = 1</math> <math>A2 + B2 + C2 + D2 = 1</math> <math>A3 + B3 + C3 + D3 = 1</math> <math>A4 + B4 + C4 + D4 = 1</math></p>											
	1	2	3	4																																			
A	5	2	3	6																																			
B	1	7	2	4																																			
C	5	8	3	1																																			
D	4	4	2	6																																			
	<p>Transportation problem The body of the table shows the costs per item of transporting from one set of locations {A, B, C, D} to another {1, 2, 3, 4}. The margins show the availability of items at locations A, B, C and D and the demands at 1, 2, 3 and 4. The LP minimises the total cost of delivering all the required items.</p>	<table border="1" data-bbox="674 1198 1005 1385"> <thead> <tr> <th></th> <th></th> <th>5</th> <th>5</th> <th>5</th> <th>5</th> </tr> <tr> <th></th> <th></th> <th>1</th> <th>2</th> <th>3</th> <th>4</th> </tr> </thead> <tbody> <tr> <th>3</th> <th>A</th> <td>5</td> <td>2</td> <td>3</td> <td>6</td> </tr> <tr> <th>6</th> <th>B</th> <td>1</td> <td>7</td> <td>2</td> <td>4</td> </tr> <tr> <th>9</th> <th>C</th> <td>5</td> <td>8</td> <td>3</td> <td>1</td> </tr> <tr> <th>2</th> <th>D</th> <td>4</td> <td>4</td> <td>2</td> <td>6</td> </tr> </tbody> </table>			5	5	5	5			1	2	3	4	3	A	5	2	3	6	6	B	1	7	2	4	9	C	5	8	3	1	2	D	4	4	2	6	<p>Minimise <math>5A1 + 2A2 + 3A3 + 6A4 + B1 + 7B2 + 2B3 + 4B4 + 5C1 + 8C2 + 3C3 + C4 + 4D1 + 4D2 + 2D3 + 6D4</math> subject to <math>A1 + A2 + A3 + A4 = 3</math> <math>B1 + B2 + B3 + B4 = 6</math> <math>C1 + C2 + C3 + C4 = 9</math> <math>D1 + D2 + D3 + D4 = 2</math> <math>A1 + B1 + C1 + D1 = 5</math> <math>A2 + B2 + C2 + D2 = 5</math> <math>A3 + B3 + C3 + D3 = 5</math> <math>A4 + B4 + C4 + D4 = 5</math></p>
		5	5	5	5																																		
		1	2	3	4																																		
3	A	5	2	3	6																																		
6	B	1	7	2	4																																		
9	C	5	8	3	1																																		
2	D	4	4	2	6																																		



<b>Statistics Chapter 1: Problem Solving</b>	
Use the statistics within the problem solving cycle	
Explain why sampling may be necessary in order to obtain information about a population, and give desirable features of a sample, including the size of the sample	
Know a variety of sampling methods, the situations in which they might be used and any problems associated with them <ul style="list-style-type: none"> <li>• Simple random sampling</li> <li>• Cluster sampling</li> <li>• Opportunity sampling</li> <li>• Stratified sampling</li> <li>• Quota sampling</li> <li>• Self-selected sample</li> <li>• Systematic sampling</li> </ul>	
Explain the advantage of using a random sample when inferring properties of a population	
Display sample data appropriately	
Calculate and interpret summary measures for sample data	

**Formula Booklet Extract**

**Sample Variance**

$$s^2 = \frac{1}{n-1} S_{xx} \text{ where } S_{xx} = \sum(x_i - \bar{x})^2 = \sum x_i^2 - \frac{(\sum x_i)^2}{n} = \sum x_i^2 - n\bar{x}^2$$

Standard deviation,  $s = \sqrt{\text{variance}}$

**Formula to Learn**

The mean of a set of data:  $\bar{x} = \frac{\sum x}{n} = \frac{\sum fx}{\sum f}$

Specification	Ref.	Learning outcomes	Notes	Notation	Exclusions
<b>STATISTICS MINOR: SAMPLING (a)</b>					
Sampling	Sx1	Be able to explain the importance of sample size in experimental design.	E.g. an informal explanation of how the size of a sample affects the interpretation of an effect size.		
	x2	Be able to explain why sampling may be necessary in order to obtain information about a population, and give desirable features of a sample.	Population too large or it is too expensive to take a census. Sampling process may be destructive. Sample should be unbiased, representative of the population; data should be relevant, not changed by the act of sampling.	A sample may also be considered as $n$ observations from a random variable.	
	x3	Be able to explain the advantage of using a random sample when inferring properties of a population.	A random sample enables proper inference to be undertaken because the probability basis on which the sample has been selected is known.		

### Problem Solving Cycle

#### Problem specification and analysis

- What problem are you going to address?
- What data do you need to collect?

#### Information collection

- How will you collect the data?

#### Processing and representation

- How are you going to process and represent the data?

#### Interpretation

- How will you be able to interpret the data and explain your conclusions?

## Key Terms and Concepts

**Raw data** – figures in their original state

**Clean data** – deal with outliers, missing data and errors

**Outliers** are extreme values – decide if they should be included or excluded (is it genuine?)

**Random variable** – a quantity that takes any of a set of values

**Population** – set of all individuals you are interested in

**Sample** – a subset of the population

A good sample should produce sample statistics (eg mean), which are in line with the population statistics

A parent population is described in terms of a **parameter** (eg the mean)

Greek letters are used to describe the population parameters, such as  $\mu$ ,  $\sigma$  etc

A **sampling frame** is a representation of items to be sampled, such as a list or grid

The **sampling fraction** is the proportion of available items that are sampled

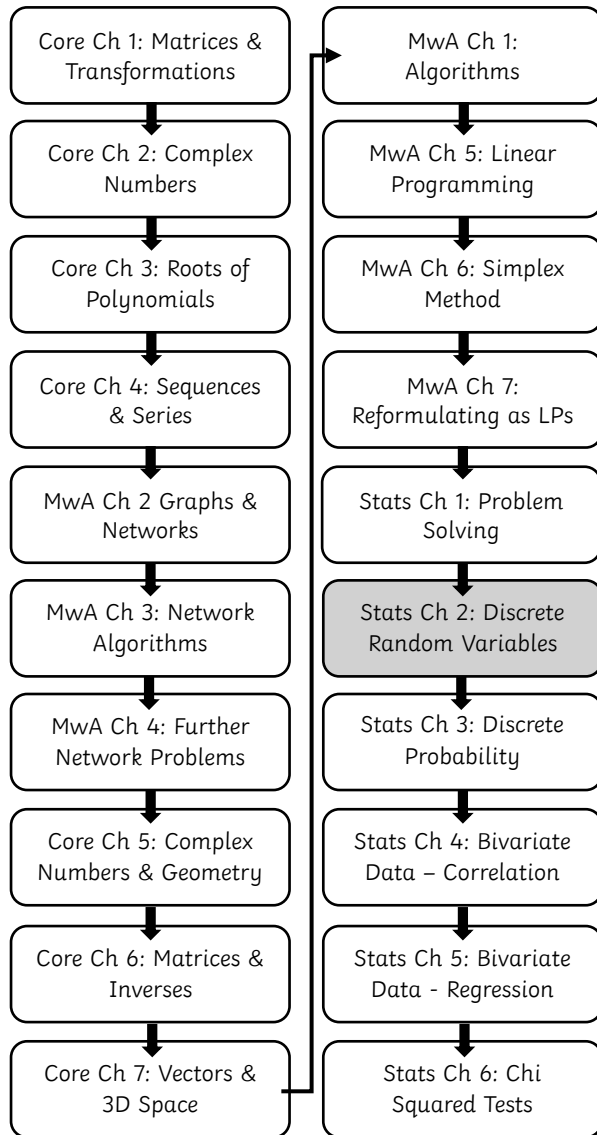
A **census** is a 100% sample

A sample is used to estimate values of a population parameter or to conduct a hypothesis test

A **sampling error** is the difference between estimated parameters from the sample and the true value

**Bias** is a systematic error, and skews results. A sample size should be big enough to produce reliable data

Sampling techniques sometimes run the risk of introducing bias into the sample.



<b>Statistics Chapter 2: Discrete Random Variables</b>	
Use probability functions, given algebraically or in tables	
Calculate the numerical probabilities for a simple distribution	
Draw and interpret graphs representing probability distributions, usually a vertical line chart	
Calculate the expectation (mean), $E(X)$ , and understand its meaning $E(X) = \mu = \sum rP(X = r)$ or $\sum x_i p_i$	
Calculate the variance $Var(X)$ , and understand its meaning $Var(X) = \sigma^2 = E(X - \mu)^2 = \sum (r - \mu)^2 P(X = r)$ or $E(X^2) - [E(X)]^2 = \sum r^2 P(X = r) - [\sum rP(X = r)]^2$	
Use and understand the meaning of the results (where $a, b, c$ are constants): $E(c) = c$ $E(aX) = aE(X)$ $E(aX + c) = aE(X) + c$ $E(X \pm Y) = E(X) \pm E(Y)$ $E(aX + bY) = aE(X) + bE(Y)$	
Use and understand the meaning of the results (where $a, b, c$ are constants): $Var(c) = 0$ $Var(aX) = a^2 Var(X)$ $Var(aX + c) = a^2 Var(X)$ And if $X$ and $Y$ are independent, $Var(X \pm Y) = Var(X) + Var(Y)$ $Var(aX + bY) = a^2 Var(X) + b^2 Var(Y)$	
Find the mean of any linear combination of random variables and the variance of any linear combination of independent random variables	

**Formula Booklet Extract**

**Discrete distributions**

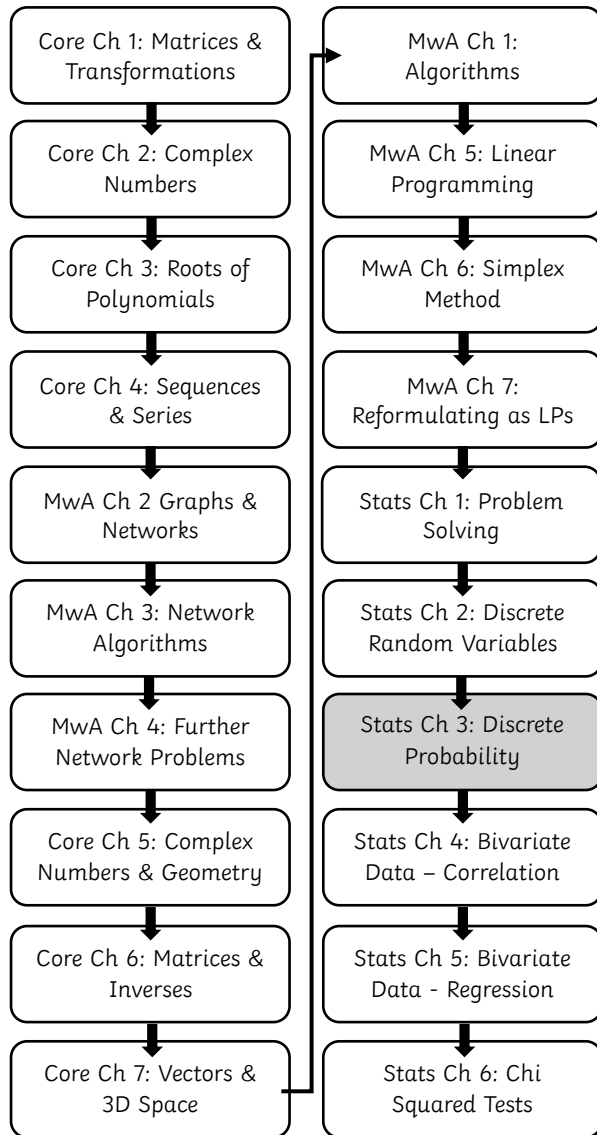
$X$  is a random variable taking values  $x_i$  in a discrete distribution with  $P(X = x_i) = p_i$

Expectation:  $\mu = E(X) = \sum x_i p_i$

Variance:  $\sigma^2 = Var(X) = \sum (x_i - \mu)^2 p_i = \sum x_i^2 p_i - \mu^2$

Specification	Ref.	Learning outcomes	Notes	Notation	Exclusions
<b>STATISTICS MINOR: DISCRETE RANDOM VARIABLES (a)</b>					
Probability distributions	SR1	Be able to use probability functions, given algebraically or in tables. Be able to calculate the numerical probabilities for a simple distribution. Be able to draw and interpret graphs representing probability distributions.	Other than the Poisson and geometric distributions, the underlying random variable will only take a finite number of values. An understanding that probabilities are non-negative and sum to 1 is expected.	$P(X = x)$	
Expectation and variance	R2	Be able to calculate the expectation (mean), $E(X)$ , and understand its meaning.		$E(X) = \mu$	
	R3	Be able to calculate the variance, $\text{Var}(X)$ , and understand its meaning.	Knowledge of $\text{Var}(X) = E(X^2) - \mu^2$ . Standard deviation = $\sqrt{\text{Var}(X)}$ .	$\text{Var}(X) = E[(X - \mu)^2]$	
	R4	Be able to use the result $E(a + bX) = a + bE(X)$ and understand its meaning.			
	R5	Be able to use the result $\text{Var}(a + bX) = b^2 \text{Var}(X)$ and understand its meaning.			

Specification	Ref.	Learning outcomes	Notes	Notation	Exclusions
<b>STATISTICS MINOR: DISCRETE RANDOM VARIABLES (a)</b>					
Expectation and variance (cont)	SR6	Be able to find the mean of any linear combination of random variables and the variance of any linear combination of independent random variables.	$E(X \pm Y) = E(X) \pm E(Y)$ $\text{Var}(X \pm Y) = \text{Var}(X) + \text{Var}(Y)$ $E(aX \pm bY) = aE(X) \pm bE(Y)$ $\text{Var}(aX \pm bY) = a^2 \text{Var}(X) + b^2 \text{Var}(Y)$		Proofs.



Statistics Chapter 3: Discrete Probability Distributions	
Recognise when the binomial distribution is likely to be an appropriate model	
Calculate probabilities using a binomial distribution	
Know and use the mean $np$ and variance $npq$ of a binomial distribution	
Recognise when the Poisson distribution is likely to be an appropriate model	
Calculate probabilities using the Poisson distribution	
Know and be able to use the mean $\lambda$ and variance $\lambda$ of a Poisson distribution	
Know the sum of independent Poisson distributions is also a Poisson distribution	
Recognise when both the Poisson distribution and the Binomial distribution might be appropriate models	
Recognise when the discrete uniform distribution is likely to be an appropriate model	
Calculate probabilities using a discrete uniform distribution	
Calculate the mean and variance of any given discrete uniform distribution	
Recognise when the geometric distribution is likely to be an appropriate model	
Calculate probabilities and cumulative probabilities using a geometric distribution	
Know and be able to use the mean and variance of a geometric distribution	

### Formula Booklet Extracts

#### The Binomial Distribution

If  $X \sim B(n, p)$  then  $P(X = r) = {}^n C_r p^r q^{n-r}$  where  $q = 1 - p$   
 Mean of  $X$  is  $np$

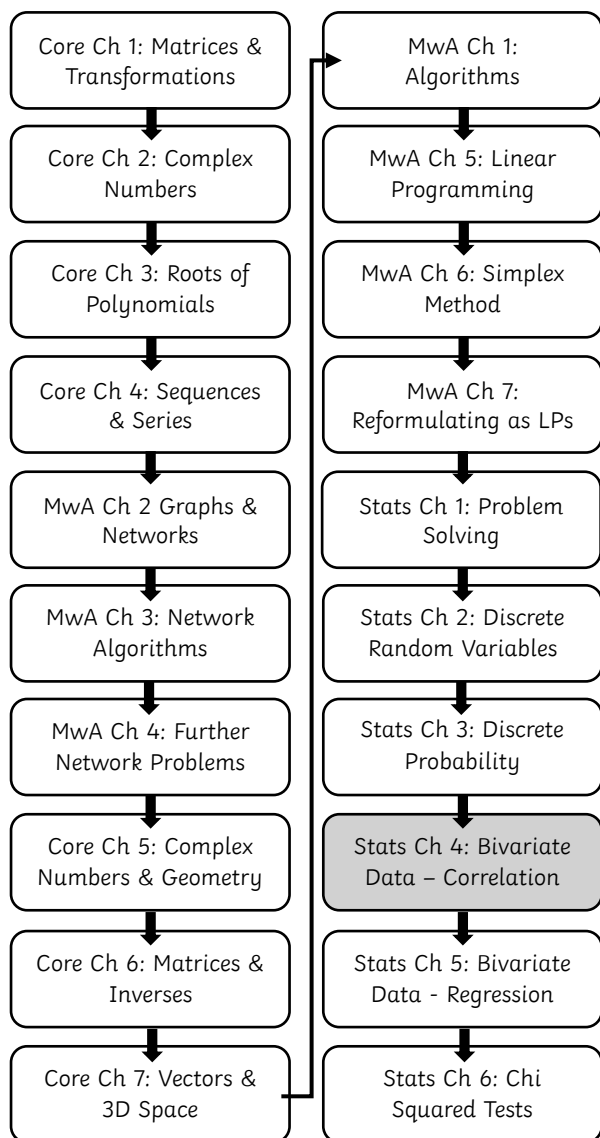
	Probability	E(X)	Var(X)
Uniform distribution over 1, 2, ..., n	$P(X = r) = \frac{1}{n}$	$\frac{n+1}{2}$	$\frac{1}{12}(n^2 - 1)$
Geometric distribution	$P(X = r) = q^{r-1}p$ $q = 1 - p$	$\frac{1}{p}$	$\frac{1-p}{p^2}$
Poisson distribution	$P(X = r) = e^{-\lambda} \frac{\lambda^r}{r!}$		

Specification: OCR Further Mathematics B (MEI) H645

Specification	Ref.	Learning outcomes	Notes	Notation	Exclusions
<b>STATISTICS MINOR: DISCRETE RANDOM VARIABLES (a)</b>					
The binomial distribution	R10	Recognise situations under which the binomial distribution is likely to be an appropriate model, and be able to calculate probabilities to use the model. Know and be able to use the mean and variance of a binomial distribution, $\mu = np$ and $\sigma^2 = np(1 - p)$ . Prove these results in particular cases.	E.g. prove results by considering a binomial random variable as the sum of $n$ independent Bernoulli random variables: $X = X_1 + X_2 + \dots + X_n$ where each $X_i$ takes the value 1 with probability $p$ and 0 with probability $1 - p$ . This proof assumes the relationship about variance in SR6.	$X \sim B(n, p)$	

Specification	Ref.	Learning outcomes	Notes	Notation	Exclusions
<b>STATISTICS MINOR: DISCRETE RANDOM VARIABLES (a)</b>					
The Poisson distribution	SR11	Recognise situations under which the Poisson distribution is likely to be an appropriate model.	Modelling the number of events occurring in a fixed interval (of time or space) when the events occur randomly at a constant average rate, and independently of each other. It is expected that these conditions can be applied to the particular context. If the mean and variance of the data do not have a similar value then the Poisson model is unlikely to be suitable.	$X \sim \text{Po}(\lambda)$ $X \sim \text{Poisson}(\lambda)$	
	R12	Recognise situations in which both the Poisson distribution and the binomial distribution might be appropriate models.	In a situation where the binomial model is appropriate, if $n$ is large and $p$ is small, then the conditions for a Poisson distribution to be appropriate are approximately satisfied. In the absence of guidance either model can be used.		Formal criteria. Using the Poisson distribution as a numerical approximation for calculating binomial probabilities.
	R13	Be able to calculate probabilities using a Poisson distribution.	Including use of a calculator to access Poisson probabilities and cumulative Poisson probabilities.		
	R14	Know and be able to use the mean and variance of a Poisson distribution.	$E(X) = \lambda, \text{Var}(X) = \lambda$		Proof.
	R15	Know that the sum of two or more independent Poisson distributions is also a Poisson distribution.	$X \sim \text{Po}(\lambda)$ and $Y \sim \text{Po}(\mu)$ $\Rightarrow X + Y \sim \text{Po}(\lambda + \mu)$ when $X$ and $Y$ are independent.		Proof.

Specification	Ref.	Learning outcomes	Notes	Notation	Exclusions
<b>STATISTICS MINOR: DISCRETE RANDOM VARIABLES (a)</b>					
The geometric distribution	SR16	Recognise situations under which the geometric distribution is likely to be an appropriate model.	Link with corresponding binomial distribution.	$X \sim \text{Geo}(p)$ , where $X$ = number of Bernoulli trials up to and including the first success.	The alternative definition which counts the number of failures.
	R17	Be able to calculate the probabilities within a geometric distribution, including cumulative probabilities.	$P(X = r) = (1 - p)^{r-1}p$ where $p$ = probability of success and $r \in \{1, 2, \dots\}$ . $P(X > r) = (1 - p)^r$ . An understanding of the calculation is expected.		
	R18	Know and be able to use the mean and variance of a geometric distribution.	$E(X) = \frac{1}{p}$ , $\text{Var}(X) = \frac{1-p}{p^2}$ .		Proof.



<b>Statistics Chapter 4: Bivariate Data – Correlation Coefficients</b>	
Understand what bivariate data are and know the conventions for choice of axis for variables in a scatter diagram	
Use and interpret a scatter diagram, including interpreting a scatter diagram produced by software	
Calculate Pearson's product moment correlation coefficient from raw data or summary statistics	
Know when it is appropriate to carry out a hypothesis test using Pearson's product moment correlation coefficient <ul style="list-style-type: none"> <li>Both variables must be random</li> <li>The data is drawn from a bivariate normal distribution, indicated by a roughly elliptical distribution on the scatter diagram</li> </ul>	
Carry out hypothesis tests using Pearson's product moment correlation coefficient and tables of critical values or the p-value from software	
Use the Pearson's product moment correlation coefficient as an effect size (a measure of the strength of relationship between the variables)	
Calculate Spearman's rank correlation coefficient from raw data or summary statistics	
Carry out hypothesis tests using Spearman's rank correlation coefficient and tables of critical values or the output from software	
Decide whether a test based on $r$ or $r_s$ may be more appropriate, or whether neither is appropriate	

**Formula Booklet Extract**

**Correlation and regression**

For a sample of  $n$  pairs of observations  $(x_i, y_i)$

$$S_{xx} = \sum x_i^2 - \frac{(\sum x_i)^2}{n}, S_{yy} = \sum y_i^2 - \frac{(\sum y_i)^2}{n}, S_{xy} = \sum x_i y_i - \frac{\sum x_i \sum y_i}{n}$$

product moment correlation coefficient:  $r = \frac{S_{xy}}{\sqrt{S_{xx}S_{yy}}} = \frac{\sum x_i y_i - \frac{\sum x_i \sum y_i}{n}}{\sqrt{\left(\sum x_i^2 - \frac{(\sum x_i)^2}{n}\right)\left(\sum y_i^2 - \frac{(\sum y_i)^2}{n}\right)}}$

Spearman's coefficient of rank correlation:

$$r_s = 1 - \frac{6 \sum d_i^2}{n(n^2 - 1)}$$

Specification	Ref.	Learning outcomes	Notes	Notation	Exclusions
<b>STATISTICS MINOR: BIVARIATE DATA (a)</b>					
<p>There are two kinds of bivariate data considered in A Level Mathematics and Further Mathematics and it is important to distinguish between them when considering correlation and regression. This note explains the reason for the distinction; learners will only be assessed on what appears under a specification reference below.</p> <p>Case A: Only <b>one of the variables</b> may be considered as a <b>random variable</b>. Often this occurs when one of the variables, the independent variable, is controlled by an experimenter and the other, the dependent variable, is measured. An example of this would be (weight, extension) in an investigation of Hooke's law for a spring. In this case certain fixed weights are used; this variable is <i>not</i> a random variable, any errors in measuring the weights are negligible. The extension <i>is</i> a random variable. There will be deviations from the 'true' value that a perfect experimenter would observe from a perfect spring as well as errors in the measurement. This case is referred to as '<b>random on non-random</b>'. The points on the scatter diagram are restricted to lie on certain vertical lines corresponding to the values of the controlled variable.</p> <p>Case B: The <b>two variables may both</b> be considered as <b>random variables</b>. An example of this would be (height, weight) for a sample from a population of individuals. For any given value of height there is a distribution of weights; for any given value of weight there is a distribution of heights. That is, there is no 'true' weight for a given height or 'true' height for a given weight. This case is referred to as '<b>random on random</b>'. The scatter diagram appears as a 'data cloud'.</p> <p>If a linear relationship between the variables is to be investigated and modelled using correlation and regression techniques then the two cases must be treated differently.</p> <p>If it is desired to test the significance of Pearson's product moment correlation coefficient then, as with all parametric hypothesis tests, probability calculations have to be performed to calculate the <math>p</math>-value or the critical region. These calculations rely on certain assumptions about the underlying distribution – <b>these assumptions can never be met in the 'random on non-random case'</b> – because one of the variables does not have a probability distribution – so <b>such a test is never valid in this case</b>. In fact the pmcc is not used in this case. In the '<b>random on random</b>' case the distributional assumptions <b>may</b> be met – see the specification below for details.</p> <p>If it is desired to calculate the equation of a line of best fit then the least-squares method is often used in both cases. However its interpretation is different in the two cases. In the example of the <b>random on non-random</b> case, (weight, extension), the line of regression is modelling the 'true' value of the extension for a given weight – the value that a perfect experimenter would observe from a perfect spring. In the example of the <b>random on random case</b>, (height, weight), the two <b>lines of regression are modelling the mean value of the distribution of weights for a given height and the mean value of the distribution of heights for a given weight</b>.</p>					

Specification	Ref.	Learning outcomes	Notes	Notation	Exclusions
<b>STATISTICS MINOR: BIVARIATE DATA (a)</b>					
Scatter diagrams	Sb1	Understand what bivariate data are and know the conventions for choice of axis for variables in a scatter diagram.	In the random on non-random case the independent variable is often one which the experimenter controls; the dependent variable is the one which is measured. The independent variable is usually plotted on the horizontal axis. In the random on random case (where both variables are measured), it may be that one is more naturally seen as a function of the other; this determines which variable is plotted on which axis.		
	b2	Be able to use and interpret a scatter diagram.	To look for outliers (by eye). To gain insight into the situation, for example to decide whether a test for correlation or association might be appropriate. Learners may be asked to add to a given scatter diagram in order to interpret a new situation.		
	b3	Interpret a scatter diagram produced by software.	Including where the software draws a trendline and gives a value for pmcc or $(pmcc)^2$ .		
Pearson's product moment correlation coefficient (pmcc)	b4	Be able to calculate the pmcc from raw data or summary statistics.	Use of a calculator expected for calculation from raw data. Summary statistics formulae given.	Sample value $r$ .	
	b5	Know when it is appropriate to carry out a hypothesis test using Pearson's product moment correlation coefficient.	The data must be random on random i.e. both variables must be random. There must be a modelling assumption that the data are drawn from a bivariate Normal distribution. This may be recognised on a scatter diagram by an approximately elliptical distribution of points. Learners will not be required to know the formal meaning of bivariate Normality but will be expected to know that where one or both of the distributions is skewed, bimodal, etc., the procedure is likely to be inappropriate. The test is for correlation, a linear relationship, so a scatter diagram is helpful to check that the data cloud does not indicate a non-linear relationship.		

Specification	Ref.	Learning outcomes	Notes	Notation	Exclusions
<b>STATISTICS MINOR: BIVARIATE DATA (a)</b>					
Pearson's product moment correlation coefficient (pmcc) (cont)	Sb6	Be able to carry out hypothesis tests using the pmcc and tables of critical values or the $p$ -value from software.	Only ' $H_0$ : No correlation in the population' will be tested. Both one-sided and two-sided alternative hypotheses will be tested. Learners should state whether there is sufficient evidence or not to reject $H_0$ and then give a non-assertive conclusion in context e.g. 'There is sufficient evidence to suggest that there is positive correlation between ... and ...'	Null hypothesis, alternative hypothesis $H_0, H_1$	
	b7	Use the pmcc as an effect size <sup>1</sup> .	Sensible informal comments about effect size are expected, either alongside or instead of a hypothesis test.		Any formal rules for judging effect size will be given.

**<sup>1</sup>Note on effect size for correlation**

For a large set of random on random bivariate data a small non-zero value of the pmcc is likely to lead to a rejection of the null hypothesis of no correlation in the population; the test is uninformative. In some contexts it is more important to consider the size of the correlation rather than test whether the population correlation is non-zero. The phrase 'effect size' is sometimes used in this context for the value of the pmcc. In social sciences Cohen's guideline is often used: small effect size 0.1; medium effect size 0.3, large effect size 0.5. Learners are not expected to know this rule; this or any other formal rule will be given if necessary.

Effect sizes for other situations, e.g. for the difference of two means, are beyond the scope of this specification.

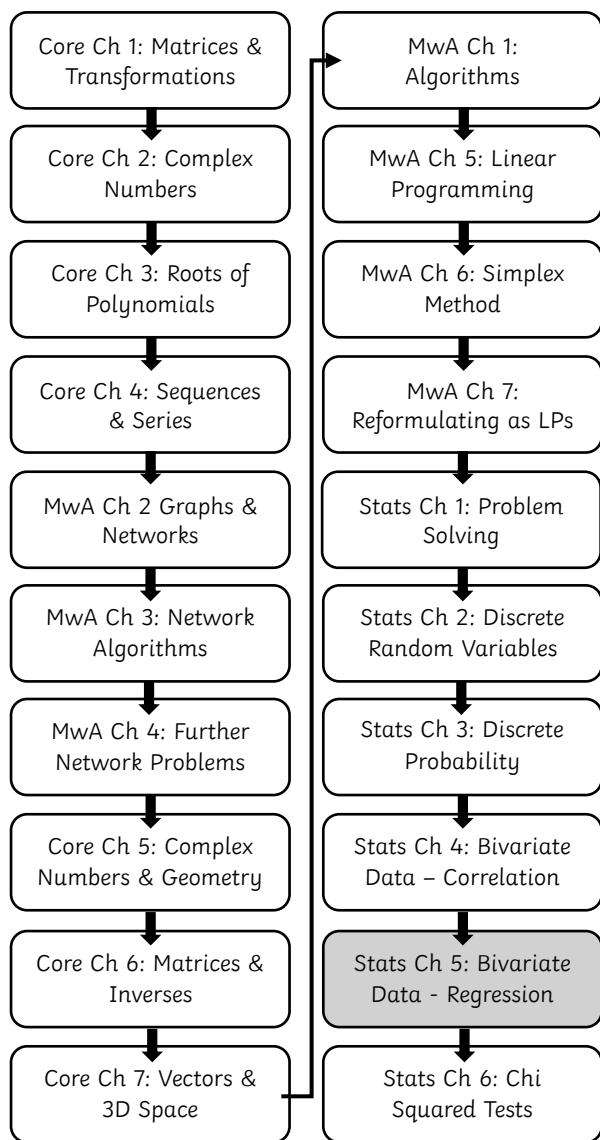
Specification	Ref.	Learning outcomes	Notes	Notation	Exclusions
<b>STATISTICS MINOR: BIVARIATE DATA (a)</b>					
Spearman's rank correlation coefficient	Sb8	Be able to calculate Spearman's rank correlation coefficient from raw data or summary statistics.	Use of a calculator on the ranked data is expected.	Sample value $r_s$ .	Tied ranks.
	b9	Be able to carry out hypothesis tests using Spearman's rank correlation coefficient and tables of critical values or the output from software.	Hypothesis tests using Spearman's rank correlation coefficient require no modelling assumptions about the underlying distribution. Only 'H <sub>0</sub> : No association in the population' will be tested. Both one-sided and two-sided alternative hypotheses will be tested. Learners should state whether there is sufficient evidence or not to reject H <sub>0</sub> and then give a non-assertive conclusion in context e.g. 'There is insufficient evidence to suggest that there is an association between ... and ...'.		
Comparison of tests	b10	Decide whether a test based on $r$ or $r_s$ may be more appropriate, or whether neither is appropriate.	Considerations include the appearance of the scatter diagram, the likely validity of underlying assumptions, whether association or correlation is to be tested for. Spearman's test is not appropriate if the scatter diagram shows no evidence of a monotonic relationship i.e. one variable tends to increase (or decrease) as the other increases. Understanding that ranking data loses information, which may affect the outcome of a test.		

Critical values for the product moment correlation coefficient,  $r$

$n$	1-Tail Test				$n$	1-Tail Test			
	5%	2½%	1%	½%		5%	2½%	1%	½%
	2-Tail Test					2-Tail Test			
	10%	5%	2%	1%		10%	5%	2%	1%
1	-	-	-	-	31	0.3009	0.3550	0.4158	0.4556
2	-	-	-	-	32	0.2960	0.3494	0.4093	0.4487
3	0.9877	0.9969	0.9995	0.9999	33	0.2913	0.3440	0.4032	0.4421
4	0.9000	0.9500	0.9800	0.9900	34	0.2869	0.3388	0.3972	0.4357
5	0.8054	0.8783	0.9343	0.9587	35	0.2826	0.3338	0.3916	0.4296
6	0.7293	0.8114	0.8822	0.9172	36	0.2785	0.3291	0.3862	0.4238
7	0.6694	0.7545	0.8329	0.8745	37	0.2746	0.3246	0.3810	0.4182
8	0.6215	0.7067	0.7887	0.8343	38	0.2709	0.3202	0.3760	0.4128
9	0.5822	0.6664	0.7498	0.7977	39	0.2673	0.3160	0.3712	0.4076
10	0.5494	0.6319	0.7155	0.7646	40	0.2638	0.3120	0.3665	0.4026
11	0.5214	0.6021	0.6851	0.7348	41	0.2605	0.3081	0.3621	0.3978
12	0.4973	0.5760	0.6581	0.7079	42	0.2573	0.3044	0.3578	0.3932
13	0.4762	0.5529	0.6339	0.6835	43	0.2542	0.3008	0.3536	0.3887
14	0.4575	0.5324	0.6120	0.6614	44	0.2512	0.2973	0.3496	0.3843
15	0.4409	0.5140	0.5923	0.6411	45	0.2483	0.2940	0.3457	0.3801
16	0.4259	0.4973	0.5742	0.6226	46	0.2455	0.2907	0.3420	0.3761
17	0.4124	0.4821	0.5577	0.6055	47	0.2429	0.2876	0.3384	0.3721
18	0.4000	0.4683	0.5425	0.5897	48	0.2403	0.2845	0.3348	0.3683
19	0.3887	0.4555	0.5285	0.5751	49	0.2377	0.2816	0.3314	0.3646
20	0.3783	0.4438	0.5155	0.5614	50	0.2353	0.2787	0.3281	0.3610
21	0.3687	0.4329	0.5034	0.5487	51	0.2329	0.2759	0.3249	0.3575
22	0.3598	0.4227	0.4921	0.5368	52	0.2306	0.2732	0.3218	0.3542
23	0.3515	0.4132	0.4815	0.5256	53	0.2284	0.2706	0.3188	0.3509
24	0.3438	0.4044	0.4716	0.5151	54	0.2262	0.2681	0.3158	0.3477
25	0.3365	0.3961	0.4622	0.5052	55	0.2241	0.2656	0.3129	0.3445
26	0.3297	0.3882	0.4534	0.4958	56	0.2221	0.2632	0.3102	0.3415
27	0.3233	0.3809	0.4451	0.4869	57	0.2201	0.2609	0.3074	0.3385
28	0.3172	0.3739	0.4372	0.4785	58	0.2181	0.2586	0.3048	0.3357
29	0.3115	0.3673	0.4297	0.4705	59	0.2162	0.2564	0.3022	0.3328
30	0.3061	0.3610	0.4226	0.4629	60	0.2144	0.2542	0.2997	0.3301

Critical values for Spearman's rank correlation coefficient,  $r_s$

$n$	1-Tail Test				$n$	1-Tail Test			
	5%	2½%	1%	½%		5%	2½%	1%	½%
	2-Tail Test					2-Tail Test			
	10%	5%	2%	1%		10%	5%	2%	1%
1	-	-	-	-	31	0.3012	0.3560	0.4185	0.4593
2	-	-	-	-	32	0.2962	0.3504	0.4117	0.4523
3	-	-	-	-	33	0.2914	0.3449	0.4054	0.4455
4	1.0000	-	-	-	34	0.2871	0.3396	0.3995	0.4390
5	0.9000	1.0000	1.0000	-	35	0.2829	0.3347	0.3936	0.4328
6	0.8286	0.8857	0.9429	1.0000	36	0.2788	0.3300	0.3882	0.4268
7	0.7143	0.7857	0.8929	0.9286	37	0.2748	0.3253	0.3829	0.4211
8	0.6429	0.7381	0.8333	0.8810	38	0.2710	0.3209	0.3778	0.4155
9	0.6000	0.7000	0.7833	0.8333	39	0.2674	0.3168	0.3729	0.4103
10	0.5636	0.6485	0.7455	0.7939	40	0.2640	0.3128	0.3681	0.4051
11	0.5364	0.6182	0.7091	0.7545	41	0.2606	0.3087	0.3636	0.4002
12	0.5035	0.5874	0.6783	0.7273	42	0.2574	0.3051	0.3594	0.3955
13	0.4835	0.5604	0.6484	0.7033	43	0.2543	0.3014	0.3550	0.3908
14	0.4637	0.5385	0.6264	0.6791	44	0.2513	0.2978	0.3511	0.3865
15	0.4464	0.5214	0.6036	0.6536	45	0.2484	0.2945	0.3470	0.3822
16	0.4294	0.5029	0.5824	0.6353	46	0.2456	0.2913	0.3433	0.3781
17	0.4142	0.4877	0.5662	0.6176	47	0.2429	0.2880	0.3396	0.3741
18	0.4014	0.4716	0.5501	0.5996	48	0.2403	0.2850	0.3361	0.3702
19	0.3912	0.4596	0.5351	0.5842	49	0.2378	0.2820	0.3326	0.3664
20	0.3805	0.4466	0.5218	0.5699	50	0.2353	0.2791	0.3293	0.3628
21	0.3701	0.4364	0.5091	0.5558	51	0.2329	0.2764	0.3260	0.3592
22	0.3608	0.4252	0.4975	0.5438	52	0.2307	0.2736	0.3228	0.3558
23	0.3528	0.4160	0.4862	0.5316	53	0.2284	0.2710	0.3198	0.3524
24	0.3443	0.4070	0.4757	0.5209	54	0.2262	0.2685	0.3168	0.3492
25	0.3369	0.3977	0.4662	0.5108	55	0.2242	0.2659	0.3139	0.3460
26	0.3306	0.3901	0.4571	0.5009	56	0.2221	0.2636	0.3111	0.3429
27	0.3242	0.3828	0.4487	0.4915	57	0.2201	0.2612	0.3083	0.3400
28	0.3180	0.3755	0.4401	0.4828	58	0.2181	0.2589	0.3057	0.3370
29	0.3118	0.3685	0.4325	0.4749	59	0.2162	0.2567	0.3030	0.3342
30	0.3063	0.3624	0.4251	0.4670	60	0.2144	0.2545	0.3005	0.3314



Statistics Chapter 5: Bivariate Data – Regression Lines	
Find the equation of the least squares regression line for a random variable on a non-random variable, using raw data or summary statistics	
Use the regression line as a model to estimate a value of the random variable and know when it is appropriate to do so	
Be able to calculate and interpret residuals For any data pair $(x, y)$ , the predicted value of $y$ is $\hat{y} = a + bx \Rightarrow residual \epsilon = y - \hat{y}$ The sum of residuals = 0	
Know that the least squares regression line minimises the sum of squared residuals	
Find the equations of the two least squares regression lines, $y$ on $x$ and $x$ on $y$ , where both variables are random, using raw data or summary statistics	
Use either regression line to estimate the expected value of one variable for a given value of the other and know when it is appropriate to do so	
Know that predicting using a value within the data values is called interpolation and is generally reliable, and predicting beyond the data values is extrapolation which is generally less reliable	
Check how well the model fits the data	
Know the relationship between two regression lines and when to use each one: Random on non-random – use $y$ on $x$ to predict $y$ values Random on random – use $y$ on $x$ to predict $y$ values or $x$ on $y$ to predict $x$ values	
Judge goodness of fit of a regression line using the coefficient of determination, $r^2$	

### Formula Booklet Extract

#### Correlation and regression

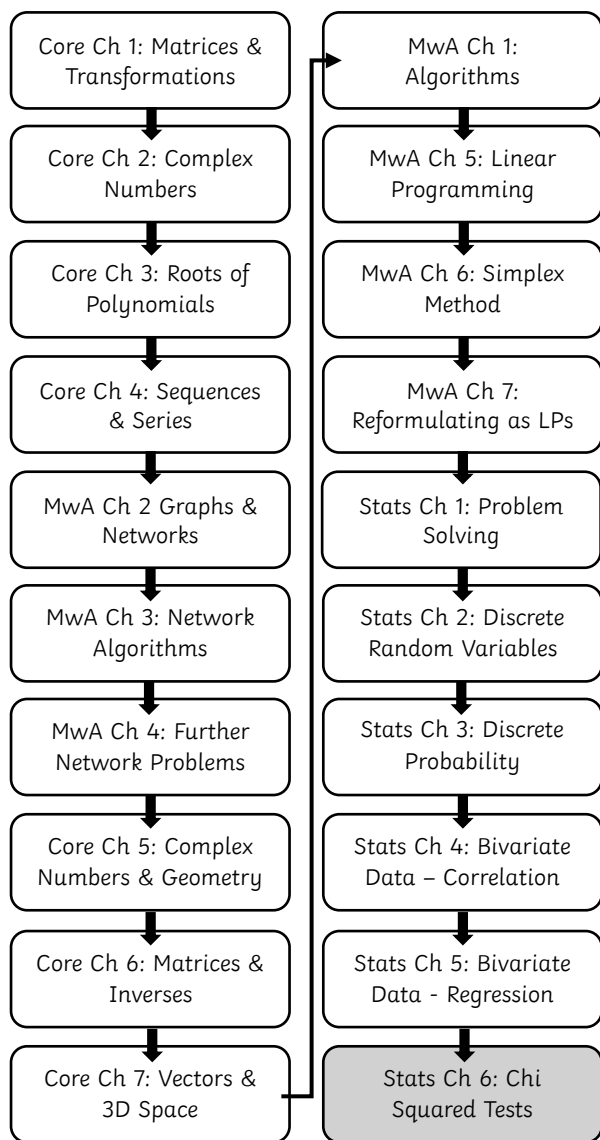
For a sample of  $n$  pairs of observations  $(x_i, y_i)$

$$S_{xx} = \sum x_i^2 - \frac{(\sum x_i)^2}{n}, S_{yy} = \sum y_i^2 - \frac{(\sum y_i)^2}{n}, S_{xy} = \sum x_i y_i - \frac{\sum x_i \sum y_i}{n}$$

$$\text{least squares regression line of } y \text{ on } x \text{ is } y - \bar{y} = b(x - \bar{x}) \text{ where } b = \frac{S_{xy}}{S_{xx}} = \frac{\sum x_i y_i - \frac{\sum x_i \sum y_i}{n}}{\sum x_i^2 - \frac{(\sum x_i)^2}{n}}$$

$$\text{least squares regression line of } x \text{ on } y \text{ is } x - \bar{x} = b'(y - \bar{y}) \text{ where } b' = \frac{S_{xy}}{S_{yy}} = \frac{\sum x_i y_i - \frac{\sum x_i \sum y_i}{n}}{\sum y_i^2 - \frac{(\sum y_i)^2}{n}}$$

Specification	Ref.	Learning outcomes	Notes	Notation	Exclusions
<b>STATISTICS MINOR: BIVARIATE DATA (a)</b>					
Regression line for a random variable on a non-random variable	Sb11	Be able to calculate the equation of the least squares regression line using raw data or summary statistics.	The goodness of fit of a regression line may be judged by looking at the scatter diagram. In this case examination questions will be confined to cases in which a random variable, $Y$ , and a non-random variable, $X$ , are modelled by a relationship in which the 'true' value of $Y$ is a linear function of $X$ . The use of a calculator is only expected for calculation from raw data. Summary statistics formulae will be given.		Derivation of the least squares regression line.
	b12	Be able to use the regression line as a model to estimate values and know when it is appropriate to do so. Know the meaning of the term residual and be able to calculate and interpret residuals.	residual = observed value – value from regression line Informal checking of a model by looking at residuals.	Interpolation extrapolation.	
Regression lines for a random variable on a random variable	b13	Be able to calculate the equation of the two least squares regression lines, $y$ on $x$ and $x$ on $y$ , using raw data or summary statistics. Be able to use either regression line to estimate the expected value of one variable for a given value of the other and know when it is appropriate to do so.	In the $y$ on $x$ case, the least squares regression line estimates $E(Y X = x)$ , that is the expected value of $Y$ for a given value of $X$ . Conversely for the $x$ on $y$ case. Only the use of a calculator is expected for calculation from raw data.		Derivation of the least squares regression lines.
	b14	Check how well the model fits the data.	Informal checking only of a model by visual inspection of a scatter diagram or consideration of $(\text{pmcc})^2$ .		Residuals in this case.
	b15	Know the relationship between the two regression lines and when to use one rather than the other. Be able to use the correct regression line to estimate the expected value of one variable for a given value of the other and know when it is appropriate to do so.	Both lines pass through $(\bar{x}, \bar{y})$ . Choice of line to use depends on which variable is to be estimated.	Interpolation extrapolation.	



Statistics Chapter 6: Chi-Squared tests	
Interpret bivariate categorical data in a contingency table	
Understand degrees of freedom, $\nu$	
Apply the chi-squared test to a $m \times n$ contingency table using $(m - 1)(n - 1)$ degrees of freedom	
Know expected frequencies must be at least 5, and classes must be combined where necessary to ensure this is the case	
Carry out a chi-squared test for goodness of fit of a uniform, binomial, geometric or Poisson model	
Calculate the degrees of freedom for a goodness of fit test, using $\nu = \text{number of classes} - \text{number of estimated parameters} - 1$	
Conclude a hypothesis test by rejecting $H_0$ if $\chi^2 > \text{critical value}$ or accepting $H_0$ if $\chi^2 \geq \text{critical value}$	
Interpret the results of a chi-squared test using tables of critical values or the output from software	

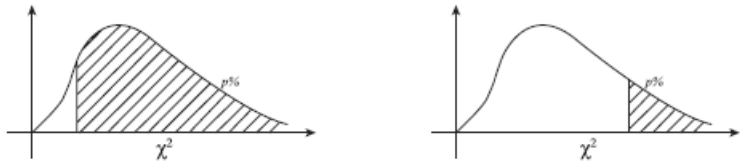
### Formula Booklet Extract

#### Hypothesis tests

Description	Test statistic	Distribution
$\chi^2$ test	$\sum \frac{(f_o - f_e)^2}{f_e}$	$\chi^2_\nu$

Specification	Ref.	Learning outcomes	Notes	Notation	Exclusions
<b>STATISTICS MINOR: CHI-SQUARED TESTS (a)</b>					
Contingency tables	Sb16	Be able to interpret bivariate categorical data in a contingency table.	Numerical data can be put into categories, but this loses information.		
$\chi^2$ test for a contingency table	SH1	Be able to apply the $\chi^2$ test (chi-squared) to a contingency table.	Only ' $H_0$ : No association between the factors' or ' $H_0$ : 'variables are independent' will be tested. Calculating degrees of freedom is expected. Knowing how to calculate observed values and contributions to the test statistic are expected, but repetitive calculations will not be required. Learners should state whether there is sufficient evidence or not to reject $H_0$ and then give a non-assertive conclusion in context e.g. 'There is not sufficient evidence to believe that there is association between ... and ...'.		Yates' continuity correction is not expected, though its appropriate use will not be penalised.
	H2	Be able to interpret the results of a $\chi^2$ test using tables of critical values or the output from software.	Output from software may be given as a $p$ -value. Interpretation may involve considering the individual cells in the table of contributions to the test statistic.		
$\chi^2$ test for goodness of fit	H3	Be able to carry out a $\chi^2$ test for goodness of fit of a uniform, binomial, or Poisson model.	Only ' $H_0$ : the given model fits the data' or ' $H_0$ : the given model is suitable' will be tested. Calculating degrees of freedom is expected. Knowing how to calculate observed values and contributions to the test statistic is expected, but repetitive calculations will not be required. Learners should be aware that cells are often combined when there are small expected frequencies, but will not have to make such decisions in examination questions. Learners should state whether there is sufficient evidence or not to reject $H_0$ and then give a non-assertive conclusion in context e.g. 'It is reasonable to believe that the ... model is suitable.'		
	H4	Be able to interpret the results of a $\chi^2$ test using tables of critical values or the output from software.	Output from software may be given as a $p$ -value.		

Percentage points of the  $\chi^2$ (chi-squared) distribution



p%	99	97.5	95	90	10	5	2.5	1	0.5
v = 1	.0001	.0010	.0039	.0158	2.706	3.841	5.024	6.635	7.879
2	.0201	.0506	0.103	0.211	4.605	5.991	7.378	9.210	10.60
3	0.115	0.216	0.352	0.584	6.251	7.815	9.348	11.34	12.84
4	0.297	0.484	0.711	1.064	7.779	9.488	11.14	13.28	14.86
5	0.554	0.831	1.145	1.610	9.236	11.07	12.83	15.09	16.75
6	0.872	1.237	1.635	2.204	10.64	12.59	14.45	16.81	18.55
7	1.239	1.690	2.167	2.833	12.02	14.07	16.01	18.48	20.28
8	1.646	2.180	2.733	3.490	13.36	15.51	17.53	20.09	21.95
9	2.088	2.700	3.325	4.168	14.68	16.92	19.02	21.67	23.59
10	2.558	3.247	3.940	4.865	15.99	18.31	20.48	23.21	25.19
11	3.053	3.816	4.575	5.578	17.28	19.68	21.92	24.72	26.76
12	3.571	4.404	5.226	6.304	18.55	21.03	23.34	26.22	28.30
13	4.107	5.009	5.892	7.042	19.81	22.36	24.74	27.69	29.82
14	4.660	5.629	6.571	7.790	21.06	23.68	26.12	29.14	31.32
15	5.229	6.262	7.261	8.547	22.31	25.00	27.49	30.58	32.80
16	5.812	6.908	7.962	9.312	23.54	26.30	28.85	32.00	34.27
17	6.408	7.564	8.672	10.09	24.77	27.59	30.19	33.41	35.72
18	7.015	8.231	9.390	10.86	25.99	28.87	31.53	34.81	37.16
19	7.633	8.907	10.12	11.65	27.20	30.14	32.85	36.19	38.58
20	8.260	9.591	10.85	12.44	28.41	31.41	34.17	37.57	40.00
21	8.897	10.28	11.59	13.24	29.62	32.67	35.48	38.93	41.40
22	9.542	10.98	12.34	14.04	30.81	33.92	36.78	40.29	42.80
23	10.20	11.69	13.09	14.85	32.01	35.17	38.08	41.64	44.18
24	10.86	12.40	13.85	15.66	33.20	36.42	39.36	42.98	45.56
25	11.52	13.12	14.61	16.47	34.38	37.65	40.65	44.31	46.93
26	12.20	13.84	15.38	17.29	35.56	38.89	41.92	45.64	48.29
27	12.88	14.57	16.15	18.11	36.74	40.11	43.19	46.96	49.64
28	13.56	15.31	16.93	18.94	37.92	41.34	44.46	48.28	50.99
29	14.26	16.05	17.71	19.77	39.09	42.56	45.72	49.59	52.34
30	14.95	16.79	18.49	20.60	40.26	43.77	46.98	50.89	53.67
35	18.51	20.57	22.47	24.80	46.06	49.80	53.20	57.34	60.27
40	22.16	24.43	26.51	29.05	51.81	55.76	59.34	63.69	66.77
50	29.71	32.36	34.76	37.69	63.17	67.50	71.42	76.15	79.49
100	70.06	74.22	77.93	82.36	118.5	124.3	129.6	135.8	140.2