NUMBER & LOGIC

TUTORS

These notes are designed to assist teachers of the course and are in a condensed format. Teachers should also consult the syllabus for this module and adapt these notes accordingly by using extra examples and by filling out this material with detail. Students will be expected to apply the material in a BUSINESS environment and with reference to the particular situation specified in the examination question. Repeating these or any other notes in a generalised form is unlikely to satisfy the required answers for any question.

NUMBER SYSTEMS

Candidate answers frequently show that they have little understanding of what constitutes a number system.

TEACHING NUMBER SYSTEMS

Determine the characteristics of the DECIMAL SYSTEM with contributions from the students. TEN is the BASE of the number system. Some book use the term RADIX. The following points should be established. The decimal system has:

1. TEN digits, numbered from ZERO to TEN-1.
2. Numbers are written down with place value. The right-hand digit represents units. The next digit is TENS.
3. From the right, the values of digits in the columns are 1, TEN^1, TEN^2, TEN^3….
4. Each digit in the decimal system has a value TEN times that of its right neighbour.
5. In adding numbers, when a column total reaches TEN or more, carry occurs to the column on the left.
6. In subtracting numbers, if the digit being subtracted in a column is more than the digit from which it is subtracted, then a borrow of TEN occurs from the column to the left. In turn, there will be a payback when the left hand column is processed.
7. Multiplying a number by TEN results in number with the same digits but with an additional ZERO attached to the right hand end. 345 x TEN = 3450
8. Dividing by TEN results in a number with the digits moved back and possibly a decimal point introduced. 345 /TEN = 34.5

All these should be obvious to the students of this course but if the basic rules are established, it is more likely that other number systems will be understood. The word TEN has deliberately been emphasised so that it can be replaced by another value in a different number system.

Candidates should also be aware of the basic patterns in the decimal system. e.g. 999+1 = 1000. We are well aware of this result and do not need to write it down by performing carries at each step. Consequently, candidates should be aware of similar situations in other number systems. e.g. in OCTAL, 777+1 = 1000 (see below) – here the "1" is not a thousand (10^3) but has the value 8^3.

OTHER NUMBER SYSTEMS

In drawing up the rules for other number system (with a base of N), all the above rules apply except TEN is replaced by N.

By convention, to avoid confusion when numbers are being written in different number systems within the same calculation, the precise number system is identified by writing the base as a subscript at the end of the number. This base is always written in decimal.
e.g. Decimal 579 is written 579_{10}.
Where no base appears, the number is assumed to be decimal.

The only number systems treated in any detail in this examination will be:

BINARY

BASE = TWO so only the digits 0 and 1 are valid. Carry therefore occurs more frequently when adding numbers.
The digits of an 8-bit byte have the place values:
256 128 64 32 16 8 4 2 1

Multiplying a binary number by 8 (= 2^3) is a similar process to multiplying a decimal number by one thousand(10^3). So 11011001_2 → $11011001\underline{000}_2$
Candidates in the past have been asked to WRITE DOWN the answer (for one mark) and have converted the original number to decimal, multiplied it by 8 and then converted back. Typical results were - two pages of calculations, much examination time wasted, errors and no mark.

Interestingly, 2^{10} = 1024 which is close enough to 1000 for approximate values of binary values to be calculated quickly.

OCTAL

BASE = EIGHT so only the digits 0 to 7 are valid.
The first five digits of an octal number have the place values:
512 128 64 8 1
As for binary, multiplying by 8 results in a shift of digits. 345_8 x EIGHT = 3450_8
There will clearly be a relationship between binary and octal because $2^3 = 8$. When converting binary to octal, separate the digits into groups of 3 **from the right** and then convert each group using the binary place values 4 2 1 (here obviously in decimal).
So binary 11110010111000001_2 = 011 110 010 111 000 001_2 = 362701_8

HEXADECIMAL

BASE = SIXTEEN. There is a problem here because unlike binary and octal where a subset of decimal digits could be used. Instead, extra digits are needed to make the TEN (0 to 9) up to sixteen. Traditionally, A to F have been used. The sixteen digits in order are:
0 1 2 3 4 5 6 7 8 9 A B C D E F (A=10, B=11, C=12, D=13, E=14, F=15)
Just as the highest single digit 7 in octal is followed by 10, so also in hexadecimal F + 1 = 10.
Of course, the "1"'s in these two instance have different values. (EIGHT and SIXTEEN respectively)
The first four place values in hexadecimal are:
4096 256 16 1
Similarly, there is a relationship between binary and hexadecimal because $2^4 = 16$. When converting binary to hexadecimal, separate the digits into groups of 4 <u>from the right</u> and then convert each group using the binary place values 8 4 2 1 (decimal).
So binary 11110010111000001_2 = 0001 1110 0101 1100 0001_2 = $1E5C1_{16}$

WORKING IN BINARY

A major cause of problems in performing calculations totally within binary is that the numbers are lengthy and it is easy for digits to be lost. Examination questions nearly always represent binary numbers in groups of 3 or 4 digits. Candidates are advised to lay out binary numbers in WIDELY spaced columns and to ensure that, for instance when several binary numbers are added, the columns are clear – use the whole width of the page for the addition. It takes no longer to do this but the likelihood of an error is reduced.

CONVERSIONS

Candidates will be expected to be able to:

- Convert integers decimal TO binary, octal or hexadecimal.
- Convert integers from binary, octal or hexadecimal TO decimal
- Convert decimal fractions TO binary, octal or hexadecimal
- Convert binary, octal or hexadecimal fractions TO decimal

DECIMAL TO BINARY

There are basically two way.

1. The first is a standard way which works for decimal to octal and to hexadecimal.
   eg 57 to binary.
   Repeatedly, divide by 2 and record the remainders at each stage (these are the binary digits FROM THE LEFT) until the number is reduced to ZERO
   57 / 2 = 28 rem 1
   28 / 2 = 14 rem 0

14 / 2 = 7 rem 0
7 / 2 = 3 rem 1
3 / 2 = 1 rem 1
1 / 2 = 0 rem 1 the **last** remainder is the **first** in the answer → $57_{10} = 111001_2$

2. The shortcut method (knowing common powers of 2)
   Students will need to know powers of 2 but the practice of manipulating numbers can improve the range.
   Break the 57 down into powers of 2 from the highest downwards.
   The largest power of 2 below 57 = 32
   57 = 32 + 25 = 32 + 16 + 9 = 32 + 16 + 8 + 1
   These are all the powers of 2 up to 32 except the 2 and 4 → 111001 (as above)
   This method is much quicker for larger numbers such as 1234
   1234 = 1024 + 210 = 1024 + 128 + 82 = 1024 + 128 + 64 + 18
   = 1024 + 128 + 64 + 16 + 2 → $10\ 011\ 010\ 010_2$ (No 512, 256, 32, 8, 4, 1)

OCTAL/HEXADECIMAL TO BINARY

1. Instead of dividing by 2 and recording the remainders as in 1 above, divide by 8 (octal) or 16 (hexadecimal).
2. Powers of 8 and 16 are less well known and unlike binary where there is either a 1 or 0 in each place, there can be a range of values (0 to 7 for octal). The second method is not recommended. But a simple example will illustrate the problem. Note the powers of 8 listed above.
   1234 = 2x512 + 3x64 + 2x8 + 2 → 2322_8

DECIMAL FRACTION TO BINARY

The method is:

1. Double the fraction.
2. Write down the value of the integer part - this is the next digit of the BINARY fraction.
3. Now repeat 1 above but ONLY doubling the fractional part.
4. Continue until the accuracy is sufficient (see Precision below)

e.g. 0 . 3 4 5
0 . 6 9 0 once a zero appears in the right hand column, this can be dropped
1 . 3 8
0 . 7 6 note only 0.38 is doubled
1 . 5 2
1 . 0 4 0.04 must be doubled many times before it will produce a 1.
0 . 0 8

So, $0.345_{10} = 0.010110_2$ etc - Are there enough binary places? (see below)
Check: 0.345 is between 0.5 and 0.25 The first binary place = ½ = 0.5, the second = ¼ = 0.25 so the answer will be 0.01etc... This agrees with the answer above.
For octal or hexadecimal, instead of doubling, multiplying by 8 or 16 respectively.

BINARY FRACTION TO DECIMAL

Probably the fastest way is to ignore the binary point and form a fraction.
e.g. 0.1101010_2 to decimal.

1. Drop any zeroes at the right hand end
2. Convert 110101 to decimal as above = 53
3. Count the number of binary places from the point to the last 1 = 6.
   The right hand "1" is effectively $1/2^6 = 1/64$.
4. The fraction is 53/64.
5. Use a calculator or long division to form a decimal value = 0.828125

PRECISION

How many decimal or binary places should be calculated during a conversion?
Note: $8 = 2^3$ and $16 = 2^4$ so $10 = 2^{3.???}$
This means that for every DECIMAL PLACE, calculate about $3\ ^1/3$ BINARY places.
It means the answer calculated above $0.345_{10} = 0.010110_2$ is incomplete. 3 decimal places requires about 10 binary places – work to 11 and round up.

BINARY IN MEMORY

When asked to show how a binary number is held in a memory location, ALL binary digits must be quoted.
In an examination question 8 or 16 bits will be used for convenience.
Decimal 25 held in 1 byte will be represented as 0001 1001

2's COMPLEMENT

2's complement is a method of representing NEGATIVE numbers in memory.
A good way to illustrate 2's complement is to consider a decimal equivalent. A 5-digit distance gauge on a car shows miles or kilometres. A new car initially showing 00000 when delivered will show 00001 if the car is driven 1 mile/km. Suppose the car had been driven backwards and suppose this meant the gauge moved backwards also. What would be the reading on the gauge? Clearly it would be 99999. It would show 99998 if driven this way for another mile/km. Similarly a very old car that had travelled 99999 miles would show 00000 after 1 further mile/km. Does this make it a new car?
Computer memory works similarly. Although it works in binary and not decimal, it is

- also a fixed length register and
- has no means of overflowing when the numbers become too big or become negative.

Few candidates seem to understand the 2's complement arithmetic. They perform arithmetic using it but in not understanding the full picture, make lengthy unnecessary calculations.

Consider a single byte for convenience. Normally, the byte could hold
0000 0000 = 0
or 1111 1111 = 255
The bit values for the 8 bits are 128 64 32 16 8 4 2 1 from the left. 255 can be calculated by adding these 8 numbers because all bits are set to 1. A much quicker way is to note that
1111 1111 + 1 = 1 0000 0000. This number **cannot** be held in the 8 bits but in pure binary, the relationship holds so the byte holds 256 –1 = 255. This method must be quicker and less prone to errors than adding the bits.

The above example is NOT using 2's complement and therefore 0 to 255 is the range possible with NO negative number representation.

However, with 2's complement arithmetic, the most significant bit (left hand end) is SIGN BIT which is

- 0 for positive numbers and
- 1 for negative numbers.

The sign bit DOES have a value – it is not just a signal for positive/negative. The values of the 8 bits in 2's complement are – 128 +64 +32 +16 +8 +4 +2 +1
In other words, the sign bit takes the value that it would normally take in non-signed numbers but it has a negative value instead of positive. Note that 64 + 32 + 16 + 8 + 4 + 2 + 1 = 127. So, if all bits are set to 1, the value of the 8 bits is –128 + 127 = – 1 Compare this with the analogy for the car distance gauge above. Adding 1 would produce all zeroes.

The **range** of the 2's complement number is now –128 to + 127. Half the numbers are positive (including zero) and the other half negative.

Example

Find the value of 1001 0100 held in 2's complement format. Three possible methods. Decide which is the better FOR THIS PARTICULAR NUMBER. Method 1 is definitely better when there are fewer 1's than 0's.

1. Using the above method, the answer is –128 + 16 + 4 = –108
2. Using the "textbook" method.
   a) The given number is 1001 0100
   b) Form the 1's complement 0110 1011 Reverse all bits

c) Add 1 0110 1100 This is now the numeric value of the original but with the reversed sign → 2's complement.

d) Calculate this value = 64 + 32 + 8 + 4 = 108.

e) The original number was therefore –108

Observations: Candidates often make mistakes at stage 3 adding the 1. They also frequently forget to reverse the sign in their final answer.

3. There is a third method that is a slight variation on 2 above. Stages b) and c) are combined into one step and without the possibility of the binary calculation error.

a) The given number is 1001 0100

bc) 2's complement 0110 1100

Locate the last "1", copy it and the digits to the RIGHT of it. Then reverse all the digits to the LEFT of this "1".

de) As before.

DATA IN MEMORY

1. **Binary-coded decimal** – If data is to be coded as decimal digits, 4 bits will be needed to hold the largest two digits 8 (=1000_2) and 9 (=1001_2). One byte can therefore hold 2 binary-coded digits. Codes from 1010 upwards are invalid. Codes 1010 and larger are therefore invalid in BCD.
2. **ASCII** (American Standard Code for Information Interchange) – This was adopted in the early years of computing to standardise how characters would be held in 8-bit bytes. 8 bits offers 256 different codes. Of these codes, 26 are for upper-letters (Capitals), 26 for lower-case letters and 10 for the digits 0 to 9. A number of codes are reserved for common punctuation and control characters (ASCII 13="NEW LINE"). By convention, codes 0 to 127 are used for the fixed requirements listed here. Codes 128 to 255 are called the **extended** ASCII set and are used internally for special purposes (e.g. graphics) and change from computer to computer. Of interest are:

   48 to 57 Digits (0 to 9)
   65 to 90 Upper-case letters (A to Z)
   97 to 122 Lower-case letters (a to z)

   Also note that the codes for lower-case are 32 bigger than for the equivalent capital letters. This means that an upper-case character can be converted to lower-case merely by setting the bit worth 32 to 1. This is obviously useful in word processing programs for changing case and for web site/email addresses where lower-case is the normal convention but where upper case can easily be recognised.
3. **Floating Point** – To accommodate very large numbers and very small fractional values using the same sized memory areas requires a different approach. The number is converted into 2 parts effectively in "standard form".
   - mantissa – fractional
   - exponent – the number of powers of 2 (many candidates think this is 10)

   Either or both could be negative and so 2's complement is used in either part.
   Consider a 16-bit memory location for convenience. The number could be split up as
   - 10-bit mantissa where the first bit is the sign bit.
   - 6-bit exponent

   Floating point numbers are **normalised** in storage. This means that for positive numbers, the fractional part is held with a 1 in the second bit (decimal equivalent = ½). Negative numbers use two's complement notation with the second bit set to 0.
   NOTE: A normalised floating point number will always have opposite values for the first two bits. The lowest value that can be held in the mantissa of a positive normalised number is ½

   Example the binary value 1101.0011 would be converted to 0.1101 0011 x 2^4 and held in 16 consecutive bits as:

   01101 00110 00100
   mantissa exponent

   The mantissa bit values range from ½ (bit-2) to 1/512 (bit-10). In normalised floating point:

- **Highest Positive** number possible = 01111 11111 011111 both sign bits 0 (= positive)
  HIGH mantissa but POSITIVE and HIGH POSITIVE exponent
  Decimal value = 1023/1024 x 2^{31}
  Note the quick method conversion → exponent = 100000_2 - 1
- **Lowest Positive** number possible = 01000 00000 100000 (most negative exponent)
  LOWEST mantissa but POSITIVE and HIGH NEGATIVE exponent
  Decimal value = 1/2 x 2^{-32}
- **Most Negative** number possible (furthest from zero) = 10000 00001 011111
  HIGH mantissa but NEGATIVE and HIGH POSITIVE exponent. The mantissa is the same as for the Highest Positive number but made negative by 2's complement.
  Decimal value = – 511/512 x 2^{31}
- **Least Negative** number possible (closest to zero) = 10111 11111 100000
  LOW mantissa but NEGATIVE and HIGH NEGATIVE exponent.
  Decimal value = –257/512 x 2^{-32}

Explanation of the exponents

Remember that HIGH POSITIVE exponents produce large powers of 2 and hence large numbers.

HIGH NEGATIVE exponents have large negative powers of 2 meaning that the number is divided by a high power of 2 and hence produces small fractions.

4. **Fixed-point Numbers** – They represent the integer and the fraction parts as consecutive bits but with an assumed point - often at the mid-point. Using mid-point notation here in 16 bits, 8 bits hold the integer (including the sign) and 8 bits hold the fraction. The two parts therefore need only one sign bit.
   In the examples below, the point has been included for clarification.

Highest positive number	= 0111 1111.1111 1111	(= 127 and 255/256)
Lowest positive number	= 0000 0000.0000 0001	(= 1/256)
Most negative number	= 1000 0000.0000 0000	(= – 127 and 255/256)
Least negative number	= 1111 1111.1111 1111	(= – 1/256)

5. **Arrays and matrices** – An array is a table of numbers. Powerful array instructions enable whole arrays to be manipulated in one instruction. Arrays can be 1-dimensional, 2-dimensional or of higher degree.
   A 1-dimensional array could be a list of numbers, one after the other and held in consecutive locations
   in memory.
   A 2-dimensional array is a human concept such as a rectangular table. Memory is, however, 1-dimensional. Suppose it is required to hold a 3 (rows) x 4 (columns) table in memory from location 100 onwards. Using the convention R_2C_3 = row 2 column 3, a possible layout could be:

100	101	102	104	105	106 …….	111
R_1C_1	R_1C_2	R_1C_3	R_1C_4	R_2C_1	R_2C_2	R_3C_4

   This stores the first row in the first 4 locations and then follows it with row 2. An alternative arrangement could be to store the 3 elements of the first column in 100 to 102 and then follow it by column 2 data. Only the first and last would be in exactly the same locations in the two arrangements.
   A 2-dimensional table is therefore being held as a 1-dimensional table in memory. A 3-dimensional table with rows, columns and layers could hold the top layer as above and then follow it in memory with the second layer.
   Candidates could be asked to derive a formula/algorithm to determine the address in memory of an element of a table in row R column C.
   Graphics bit-maps are usually 2-dimensional and would be held this way when they are stored in a disc file.

INSTRUCTIONS IN MEMORY

It is common to refer to machine code instructions as being held in **words** in memory. A word is merely a specified number of associated bits. In their simplest form, an instruction must include at least

- an **operation code**
  If 8 bits are allocated for the operation code, then 256 different coded operations are possible.
- an **address**
  If 24 bits are allocated for addresses, then this allows accesses to locations 0 up to $2^{24}-1$ = 16777215 or 16 Mb. Even this is too small for the latest huge software programs and large memory computers.

LDA 64 could be a machine code to LOAD the contents of location 64 into an accumulator register. This might be coded 00000000 00000000000000001000000 in 32 bitswhere the initial is the code for LDA 00000000 (because it is the most commonly used instruction) and 64 is held in pure binary in the other 24 bits.

2-instruction addresses – such as ADD A,B which would add the contents of location B to location A. The addresses of A and B would probably need an equal number of bits**.**

Registers – The instruction could include part of the word to identify the address of a particular register. There are few registers compared with the number of address in memory and so registers tend to use low numbered addresses and therefore require only a small number of bits to address them.

MATRIX ARITHMETIC

The syllabus requires that candidates can:

1. Add/Subtract matrices – they therefore need to be the same shape.
2. Multiply two matrices – ORDER MATTERS.

   e.g. $\begin{vmatrix} a & b \\ c & d \\ e & f \end{vmatrix} \times \begin{vmatrix} p & q \\ s & t \end{vmatrix} = \begin{vmatrix} ap+bs & aq+bt \\ cp+ds & cq+dt \\ ep+fs & eq+ft \end{vmatrix}$

   Use the convention A(r,c) to indicate the number of rows and columns in array A. This example is therefore A(3,2) x B(2,2)

   The test whether two matrices CAN be multiplied is that A x B requires <u>Columns of A = Rows of B (Remember the two INNER numbers when the matrices are written in row,col notation above)</u>

   The final shape will be <u>Rows = Rows of A</u> and <u>Columns = Columns of B.</u> (The two OUTER numbers).

   In this example, B x A is not possible.
3. Determine the inverse of a matrix (A^{-1}). A x A^{-1} = A^{-1} x A = I, the **identity** matrix where all elements are zero except the leading diagonal which contains 1's. The inverse is only defined for a square matrix.

   The inverse of a 2-dimensional matrix can be written down as:

   $\begin{vmatrix} a & b \\ c & d \end{vmatrix} = \frac{1}{D} \begin{vmatrix} d & -b \\ -c & a \end{vmatrix}$ where D = determinant of A = ad-bc
4. For a 3-dimensional array, candidates will NOT be expected to derive the inverse BUT could be asked to multiply two matrices, one of which is a multiple of the inverse of the other. i.e. n x A x A^{-1} The candidate should recognise that the inverse has been given and then use this result. The word "HENCE" in the question should give the hint.
5. Solve 2 or 3 variable simultaneous equations using the results of 3 and 4 above.

   e.g. A x X = N A is a square matrix of degree n, X is column vector (x,y..) with n rows and N is a numeric column vector with the same shape as X.

   Solution. A^{-1} x A x X = A^{-1} x N Since A^{-1} x A = I, X = A^{-1} x N and hence solutions for x,y…
6. Impose matrix notation on given tables of data and derive meaning to addition, subtraction or multiplication of two of the matrices.

<u>ALGORITHM</u>

Candidates may be asked to produce an algorithm for the addition, subtraction or multiplication of two matrices. Ideally this should be in **pseudo-code** format to avoid any ambiguity. Other means of algorithmic representation are permitted provided they are totally clear what each step involves. In pseudo-code, the following control structures are expected:

1. Double loop to access all elements of the rows and columns.

```
FOR col ← 1 to NoOfCols
    FOR row ← 1 to NoOfRows
        details of the array manipulation
    ENDFOR
ENDFOR
```

2. Array handling.
   Use A(row,col) to access a particular element of array A.
   In adding and subtracting two matrices, the same (row,col) subscripts can be used for all arrays because the two arrays AND the combined result array will all be the same shape. With multiplication of A x B, since the elements of a row in A are multiplied by the elements of a column in B, then the row subscript of A could be used in multiplying as the col subscript of B. Detailed consideration of the various steps involved in the multiplication of two particular matrices will make it clear what is needed.

ITERATIVE METHODS

An iterative method for solution of equations is one in which an estimate is fed into a formula and hopefully a better solution comes out. If this is repeated many times, the successive answers obtained "home in" on an accurate answer. Consider the function $F(x) = x^3 – 2x^2 + 3x – 4$. $F(x) = 0$ has 3 possible solutions because of the x^3. By substitution of different integer x values, a guide can be obtained approximately where these solutions occur. Here $F(0) < 0$, $F(1) < 0$ but $F(2) > 0$. Therefore there is a solution between x=1 and x=2 because the sign of F(x) changes. Rearrangement of $F(x) = 0$ to produce x = (new function) could give a suitable iterative solution.

$x^3 – 2x^2 + 3x – 4 = 0 \rightarrow x(x^2 – 2x + 3) = 4 \rightarrow x = 4 /(x^2 – 2x + 3)$ This could be a possible equation.

Trying x = 1 in the expression $4 /(x^2 – 2x + 3) = 2$.

Now trying x = 2 in the expression $4 /(x^2 – 2x + 3) = 4/3$.

This is a good sign since the third answer is between the first and second. This suggests convergence but it might be slow and it might be a false sign. The reader should continue and bear in mind the points made in the convergence paragraph below.

This process is repeated until:

1. Two successive answers are within the accuracy required. If 3 decimal place accuracy is needed then the difference between the two answers should be less than 0.001.
2. OR the successive solutions diverge. In which, a new iterative equation must be derived and the process repeated.

Other possible arrangements are:

a) $x = \sqrt{[(4-3x)/(x-2)]}$ requiring a calculator for square rooting.

@@@@ symbol after equals above should be SQUARE ROOT

b) $x = (4-3x)/(x^2 – 2x)$

The reader should check these can be rearranged to make the original $F(x) = 0$.

CONVERGENCE ?

Possible outcomes are that successive solutions:

1. Converge as above – the desired outcome.
2. Converge but very slowly and therefore a large number of iterations need to be performed.
3. Converge but oscillating about a final answer. i.e. if one solution is above the final answer, the next is below and vice versa. This is satisfactory when the higher alternate values are decreasing and lower alternate solutions are increasing. The accurate answer is between.
4. Diverge rapidly. While this will not produce a solution, little work is needed to come to the conclusion that divergence is taking place.
5. Diverge but then converge on one of the other solutions to F(x)=0. This will not solve the problem.
6. Diverge slowly. This is one of the worst situations because you cannot be sure if the divergence does not later converge on the correct value. It is probably better to try a new iterative equation that might save time by converging quickly.

ADVANTAGES OF ITERATIVE METHODS

- Some equations have no known mathematical solution – particularly those with mixed types of components such as power of x and sin(x).
- Known methods of solution may still involve lengthy calculations such as cube or higher roots.
- Manual methods may be very slow particularly if a high degree of accuracy is required.
- Computers can repeat a given method very quickly and so high accuracy can be achieved with little extra time.

LINEAR PROGRAMMING

Linear programming is the method used to determine the **best fit** for a practical situation where a number of conditions are imposed. It is recognised this could readily be implemented using computers but candidates are expected to understand the processes. They could be asked to solve:

1. a 2-variable problem by graphical means
2. a 2 or 3 variable problem using the Simplex Method.

GRAPHICAL SOLUTIONS

A practical situation will be given with a number of conditions. A full explanation is given here because this topic has not been answered well in the past. It is best illustrated by example.
"A company makes two types of products, X and Y. Production formation per week is given as for EACH UNIT as:

1. A minimum of 20 of X and 50 of Y must be produced.
2. Assembly time in hours per unit is 4 for X and 5 for Y with only 1000 employee man-hours available.
3. Testing needs equipment. It takes 3 hours to test X, 2 hours for Y and only 600 hours are available.
4. The profit of X is £100 and £90 for Y. This is to be maximised within the conditions and the profit this represents calculated."

Define x = number of X made per week, and y for Y.
Condition:
For minimum production levels $x \geq 20$ and so plot straight vertical line x = 20 using a ruler and pencil.
$y \geq 50$ and so plot horizontal line y = 50
For assembly $4x + 5y \leq 1000$ and so plot $y = 200 - 4x/5$
For testing $3x + 2y \leq 600$ and so plot $y = 300 - 3x/2$
In drawing the graph, as large a scale as possible should be used for both axes with the y axis vertical. Use intervals that can be accurately estimated for intermediate values as is normal when constructing graphs. The straight lines drawn represent the limiting situations in each case and so LIGHTLY shade the sides of each line that are on the **invalid** side of the equation. For x=20, shade left of it. The shading must not obscure any lines or written words. The conditions should give an area (often enclosed) which represents the valid conditions.
The maximising function is Profit P = 100x + 90y. Rearranging this produces and equation: $y = P/90 - 10x/9$. The only feature to note here is that this line has a downward gradient of 10/9. On the graph, draw ANY line with this slope – the easiest to draw is to join the 10 on the y-axis to the 9 of the x-axis (or multiples of these numbers). Now, maintaining this same slope, slide the line until it cuts across the valid area at its highest point. In this particular question, only whole numbers are permitted but in other situations, fractional values may be allowed.
For this problem, the optimum point on the profit line gives x=142 and y=85. Using the profit function, P = £21850 for these values.
In order to obtain the x and y values here, it is clearly necessary to have a graph where 142 and 85 can be read off exactly.
If the problem is about **minimising** costs, the line must be slid into a position showing the lowest value in the valid area.

SIMPLEX METHOD

A full example of the method is given, illustrated with the above problem, because it may be difficult to find an explanation in books. A detailed explanation going beyond the syllabus

requirements can be found in T.Lucey's "Quantitative Techniques" published by DP Publications (ISBN 1873981260).

1. Write the given information in table format (called a tableau) as follows:

1	SOLUTION VARIABLE	x	y	a	b	SOLUTION QUANTITY (S)	
2	a	4	5	1	0	1000	Assembly
3	b	3	2	0	1	600	Testing
4							
5	P	100	90	0	0	0	Profit

   $x \geq 20$ and $y \geq 50$ have been left at of the table for the moment.

2. a and b columns are **slack** variables. Their values are initially set to zero except the leading diagonal where 1's are inserted. In more involved problems, there could be many more conditions and hence the need for more slack variables.
3. In row 4, the assembly condition $4x + 5y \leq 1000$ is represented here as **equation** $4x + 5y + a = 1000$
   a represents the spare assembly time available.
4. Similarly, row 5 shows the equation $3x + 2y + b = 600$ and here b gives the spare testing time available.
5. The two delivery conditions in this particular problem each create difficulties because the conditions imposed are not "≤" but "≥". Simplex can handle a "greater than" inequality but requires a different approach not attempted here. Conditions cannot be mixed. To overcome the x problem, the 20 of X that MUST be made are taken out of the problem and at the end, added into the final answer. If 20 of X are made, they will need 20 x 4 (= 80) hours of assembly and 20 x 3 (= 60) hours of testing. These two figures are subtracted from the numbers in the solution column in rows 2 and 3 respectively. 1000 – 80 = 920 and 600 – 60 = 540. Thus, AFTER these 20 of X have been made, there are 920 hours for assembly available and 540 hours for testing.
6. This is repeated for the Y delivery condition. 5 x 50 = 250 hours for assembly and 2 x 50 = 100 hours for testing. This reduces the assembly still further to 670 and testing to 440. If these figures are built into the table, then at the end, 20 must be added to x and 50 to y.

1	SOLUTION VARIABLE	x	y	a	b	SOLUTION QUANTITY (S)		
2	a	4	5	1	0	670	Assembly	
3	b	**3**	2	0	1	440	Testing	**PR**
4								
5	P	100	90	0	0	0	Profit	
		PC						

7. Row 5 shows the profit parameters. Zeroes are place in each of the other table entries. Now, look at the P line and find the highest POSITIVE number. Here x=100 so x is the PIVOT COLUMN (PC). This is chosen because it is x that makes the higher contribution to profit - we should make as many of X as possible. Divide the x column values into their corresponding S quantities to find the **lowest** number. 670/4=167.5 and 440/3=146.67. There is no need to perform this stage accurately because rarely are the numbers very close together. The second is the lower one here - row 3 is the PIVOT ROW (PR). Mark the x column value (x=3) in some way. This is known as the PIVOT VALUE.
8. Now, divide **all** the elements of this **row** (row 3) by this pivot value. Replace the variable (b) in solution variable column with x. This will leave this pivotal element with the value 1. The S value on this row shows that 146.7 units can be made of x

1	SOLUTION VARIABLE	x	y	a	b	SOLUTION QUANTITY (S)		
2	a	4	5	1	0	670	Assembly	
3	**x**	**1**	**0.67**	**0**	**0.3333**	**146.67**	Testing	PR
4								
5	P	100	90	0	0	0	Profit	

		PC						

9. Subtract MULTIPLES of the pivot values from EACH OTHER row to reduce all elements in the PC to zero. So perform row 2 - 4 x row 3 for all columns. This reduces the 4 to zero and reduces other elements in other columns. In the P row, perform row 5 – 100 x row3. The new value in the S column, row P shows the contribution made by x in the profit. (£14667) – take its positive value.

1	SOLUTION VARIABLE	x	y	a	b	SOLUTION QUANTITY (S)		
2	a	**0**	**2.33**	**1**	**-1.33**	**83.32**	Assembly	**PR**
3	x	1	0.67	0	0.33	146.67	Testing	
4								
5	P	0	23	0	-33	**-14667**	Profit	
			PC					

10. IF there are any positive numbers in the P row, the steps 7,8 and 9 are repeated. In this case, clearly y will be the new PC (the x value column has been completed).
It will be found that row 2 is the new PR and so the pivotal value is 2.33. As before, this will be reduced to 1 (step 8). With the pivotal value now 1, multiples of this value are subtracted from the other two rows to reduce all numbers in the pivotal column (except the pivotal value itself which is 1) to zero. The result will be:

1	SOLUTION VARIABLE	x	y	a	b	SOLUTION QUANTITY (S)		
2	y	**0**	**1**	**0.43**	**-0.57**	**35.714**	Assembly	**PR**
3	x	1	0	-0.29	0.33	122.86	Testing	
4								
5	P	0	0	0.99	-20.03	**-15300**	Profit	
			PC					

From this table, 122.86 should be made of X. x can only be a whole number so round down to the nearest whole number (rounding up will probably mean one of the original conditions is violated). Similarly 35 of Y should be made.

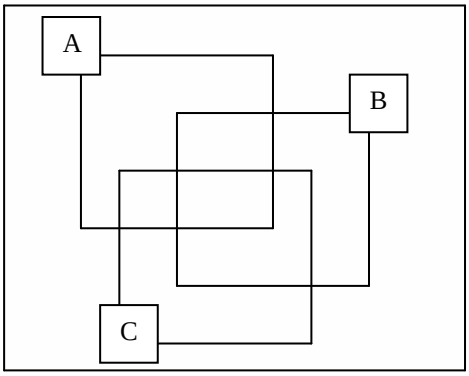
These, would bring in a profit of £15300. HOWEVER, it must be remembered that 20 should be added to X and 50 to Y. So the final result is make 142 of X and 85 of Y as predicted by the graphical method. The profit = 142 x 100 + 85 x 90 = £21850. Given that whole number solutions are required and the method does not take account of this, it is possible that the final result is slightly in error it could be x=143 and y=84. As a guideline, these figures would be good enough for any management decisions.

Important note: In practice in an examination, the candidate will only be asked to perform PART of this process but would only be able to do this if the whole process is understood.

VENN DIAGRAMS

These are used to count the number of people/items in particular categories. Questions will normally involve three categories. An example will illustrate.

"100 computer dealers in a survey provided the following information about printers. 36 sold model A, 32 sold B and 31 sold C. 9 sold A AND B. 12 sold B AND C. 11 sold A AND C. Only 4 sold all three models. Use a Venn diagram to determine how many of these dealers did NOT sell printers."



The candidates need to place the given numbers into the correct areas of the diagram. It is normally easier to draw **circles** for the three categories with the surrounding box outside representing the "population" of the survey. The area which is inside ALL 3 boxes then represents the number of dealers who sold all three printers (4 in this question). It is important to note that the phrase "9 sold A AND B" does not mention C and does not imply that these 9 dealers did not also sell C. Had the phrase been "9 sold ONLY A and B", then this would have excluded C. In fact we know that 4 sold all models so therefore 5 sold A AND B but not C. Given that there are 100 dealers in the survey, the number placed in the area inside the outer box but not in the A, B or C areas is the answer to the question. It can be found by determining the numbers in each section of the inner areas and subtracting the total from 100. The answer is 29.

Candidates frequently misinterpret the **exact** meaning of the information given. Questions set in class should be discussed to ensure all candidates understand this. Any information given in a statement tells the reader about the items mentioned and tells NOTHING about any items NOT mentioned. Candidates are expected to explain how they achieved their answers. It will often be necessary to place an x or y in certain areas of the diagram. For instance, a statement such as "1 more dealer sold A AND B compared with B AND C" might have been included. This means that an x can be placed in one area and an (x+1) in another to help to define the conditions. Finding x will determine both values.

NUMBER SERIES

Number series are important in computing to enable patterns in numbers to be determined.

Arithmetic Progression (AP) – These are sequences of numbers with the same **difference** between consecutive numbers. e.g. 2, 5, 8, 11….. Each number is usually called a term.

a is used to define the FIRST term - here a=2

d is used to define the difference - here d = 5-2 = 3

n is used to define the number of terms in a sequence.

$\mathbf{S_n}$ is the sum of those n terms. If a series has only 7 terms, this is written as S_7.

n-th term – clearly the 8^{th} term will be the first term with 7 differences added.

So a general formula for the nth term = a + (n-1)d

Sum of n terms - $S_n = n[2a – (n-1)d]/2$

It is possible to derive this formula. Write out all terms of the series:

S = a + a+d + a+2d……+ a+(n-2)d + a+(n-1)d

Underneath, write down the same series only **backwards**. Add the pair of corresponding terms. It will be found that each pair adds to 2a+ (n-1)d.

There are n of these terms so 2a+ (n-1)d needs to be multiplied by n but this gives the sum of TWO series - the result must be divided by 2.

In the special case for adding n successive integers starting always with 1, then a=1 and d=1 which reduces the formula to $S_n = n(n+1)/2$

Exercises should be set with practical meaning if possible and with a range of values for a and d including fractional and negative values. e.g. The second term is 22, the fourth term = 16. How many terms are needed for a total of 82? (clearly d < 0 here). Writing down formulae for the 2nd and 4th terms will produce two equations involving a and d and hence both can be found.

Geometric Progression (GP) – These are similar to APs but instead of a common difference, there is a common ratio **r**. e.g. 1, 3, 9, 27, 81….

n-th term – In this case, the 8th term will be the first term but **multiplied** 7 times by the ratio. So a general formula for the nth term = $a r^{(n-1)}$

Sum of n terms - $S_n = a(1-r^n)/(1-r)$

It is also possible to derive this formula. Write out all terms of the series:

$S = a + ar + ar^2 \ldots\ldots\ldots + ar^{n-2} + ar^{n-1}$

Underneath, write down the same series only multiplied by r. Subtract the two series to calculate

$S - rS = S(1 - r)$. All the terms except the first and last of the combined series cancel and hence the formula.

In the special case where r < 1 to determine the sum of an infinite number of terms (which are clearly getting smaller), then r^n tends to zero and so the sum to infinity is $a/(1-r)$. It must be stressed this only applies to decreasing series. A series with increasing terms would produce an infinite sum.

Exercises should be set with practical meaning if possible and with a range of values for a and r including fractional values. e.g. How far does a ball move if dropped from 10 metres when it bounces to half its previous height (a) after 10 bounces (b) if allowed to continue to bounce until it eventually comes to rest. Calculations involving depreciation and inflation over several periods of time require the use of GPs.

DEPRECIATION AND INFLATION

1. The value of a car will clearly depreciate every year as it wears out. Suppose that for taxation purposes, a person is allowed to offset the value of the car against tax assuming a 25% depreciation each year. The value in the second year will be 0.75 of its first year value. Similarly, the value reduces by 0.75 in each subsequent year. After 5 years, the value of a £10000 car would be 10000×0.75^5. This would also apply to a person earning the same wage each year over 5 years. The effective buying power decreases by that same factor.
2. Inflation works in the opposite sense in that the effective cost of an item increases each year. If a country has a constant inflation rate of 7% over 5 years, the cost of living would rise by a factor of 1.07^5

INTEREST

Simple interest is the gain that a sum of money achieves when invested over a period of time but where the interest paid each year DOES NOT ITSELF GAIN INTEREST. This occurs perhaps where the interest is paid to the investor directly so that the capital invested remains constant.

Interest = P x R x T / 100 where P is the Principal (Capital invested), R is the interest rate (%) and T is the number of time periods (usually years).

Value of the Investment after T years = P + Interest x T Interest is paid T times and the original Principal is still available. A general formula for this amount is $A = P(1 + RT/100)$

Compound Interest is the interest gained as a result of all interest paid each time being immediately reinvested with the original capital. It follows that the capital increases each year and so the interest earned each year increases assuming a constant rate. Compound Interest uses GPs.

The general formula for the value of an initial investment of P at R % for T years is $A = P (1 + R/100)^T$

ORDERING SYSTEMS

A vital part of business is to have an efficient **stock control**. A stock system could have

1. HIGH stock levels - ensures that there are sufficient items in stock to satisfy orders.

2. LOW stock levels - ensures lowest amount of money is invested in unsold stock.
   The secret is therefore to devise a system that is a happy medium of these extremes.
   Consider the situation for consumption in a company of particular items.
   Minimum usage = 50
   Normal daily usage = 100
   Maximum usage = 150
   Lead time (days to obtain new stock) = 20 to 30 days.
   Previous Economic Order Quantity (EOQ) = 4000
   From this, it is possible to calculate:
   i. Reorder Level = when stock level falls below this, reordering should take place
      = Maximum usage x Maximum lead time = 150 x 30 = 4500
   ii. Maximum level = above which, the stock level should not rise
      = Reorder Level – minimum usage + EOQ = 4500 – 50x20 + 4000 = 7500
   iii. Minimum level = warning of possible shortages if demand increases
      = Reorder Level – average usage in average time = 4500 – 100x25 = 2000
3. EOQ can be calculated from a formula but is only as reliable as the estimates of the various figures used to calculate it.
   EOQ = $\sqrt{(2 \times C \times A/H)}$ where C is the cost of each order, A is the annual demand and H is the holding cost per year for each item (to hold in stock).

STATISTICS

Candidates must be able to provide a clear, unambiguous definition of the various averages in common use. They must also have an awareness of the meaning and use than can be applied to each. If the data is a series of x values and there are n values in the population, then:

mean = Total of the values/count of the number of values = $\sum x/n$. The symbol "$\sum x$" represents "the sum of all x values". It should be remembered that if each x value was replaced by the mean value, the total of all values would not change.
Example: The mean of 6 values is 100. After one more number is included, the mean rises to 105. What is the value of this additional number? Note: It must be significantly higher than 100 to raise the mean by 5%.
Solution 1: (Arithmetic) The total of the 6 numbers is 6x100=600. The total of the 7 numbers is 7x105=735. Therefore the seventh number must be 735 – 600 = 135.
Solution 2: (Analytical) For the mean to be 105 after 7 numbers, the 7^{th} must be 105 + an extra 5 for EACH of the other 6 = 105 + 6x5 = 135.
Uses: To produce a single value that typically represents each of the items in the population.
e.g. The average number of cattle in all farms across the country would enable estimates to be made of national milk yields, beef production, dairy feed needed, number of veterinary experts needed.

median = The middle value. If the data is arranged in ascending order of value, the median is positioned the same distance from each end. If there are an even number of data items and hence there are two middle values, the median is the mean of these two.
Uses: A useful representative value where data is clustered with few extreme large and small values. e.g. If the number of cars passing a junction is measured every hour between 6am and midnight, there will be short periods when the junction is very busy (everybody going to or from work) and other times when few users are on the road (very early morning/late evening). The median could be a good representation of normal flow of traffic.

range = the difference between the highest and lowest x values.
Uses: Most items that people use fall between fairly distinct large and small values.
e.g. Cupboards are made to hang clothes between small and large. A low cupboard could hang children's clothes and a high one adult clothes. There would be little need outside this range because people would have little use for them and they would difficult to access.

mode = the most common value. This is only appropriate where distinct x values are possible such as in shoe sizes. People's heights could take any value within a given range and hence would not have a modal value (unless all heights were measured to the nearest cm.).

Uses: A clothes shop would be conscious of the modal value of people sizes because it would sell more clothes of this size.

FREQUENCY TABLES

Large quantities of data for analysis are usually collected in distinct bands rather than considering each as an individual value. The results are then summarised in a table where again ∑ represents "SUM OF". e.g.

Class	Frequency (f)	Midpoint (m)	f x m		
50-59	6	54.5	327.0		
60-69	8	64.5	516.0		
70-79	10	74.5	745.0		
80-89	9	84.5	760.5		
90-99	5	94.5	472.5		
TOTALS	38		2821.0		

Row 1 of these figures shows that there were 6 data items in the class 50 to 59. The exact values of each of the 6 items are unknown. However, their effect is that they can each be considered to take the mean of the class (= 54.5). Some would be higher and some lower. By multiplying the frequency by the class midpoint, a total (327.0) is produced which would be very close to the actual total of the original 6 numbers.

Mean. There are 38 items in the sample (∑ f). The total of all items is 2821 (∑fm). Therefore by definition, the mean of this data = (∑fm)/(∑ f) = 2821/38 = 74.2 This result looks reasonable because of the distribution of the data with lower values in the two extreme classes and the largest value in the middle class. A mean close to the midpoint of the middle class is reasonable.

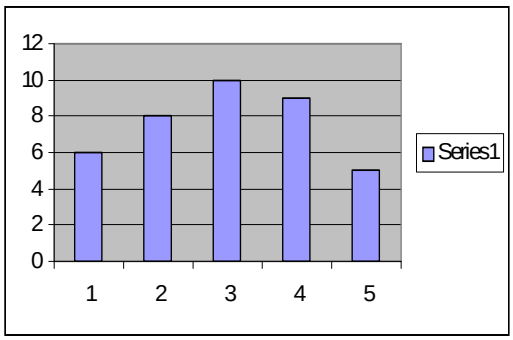
Mode. The modal class is 70 to 79 and the modal value can be taken to be its mean (74.5)

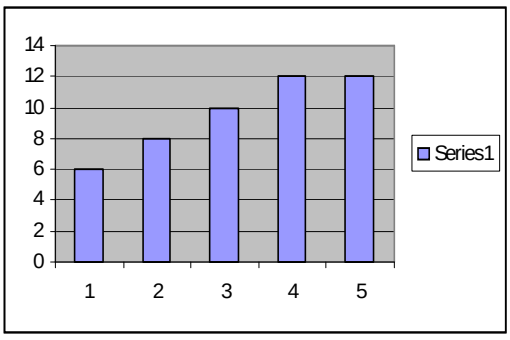
Median. There are 38 numbers and so the median will be between to 19th and the 20th. Counting down from the top of the table, 6 are below 60, 6+8 = 14 are below 70 and 14+10=24 are below 80. The 19th and 20th are therefore in the 70 to 79 class and its mean can be taken to be the median (74.5).

Range. The data varies from 50 to 99.

This distribution of data is fairly symmetrically arranged about the 70 to 79 class and so not surprisingly the three averages, mean, mode and median are close.

Had the 5 frequencies been 6, 8, 10, 12, 12, then clearly the data is **skewed** towards the higher values and higher values would have been obtained for the averages. However, the median would have been more complex. There are 48 in this new list and so the 24th and 25th would be the middle values. In this case, the 24th from the lower end is the last in the 70 to 79 class and the 25th is the first in the 80 to 89 class. The mean of 60 and 79.5 should be taken (lowest and highest of the combined two classes).





The two **barcharts** show, on the left, the distribution of data in the original table and on the right, second set with the skew towards the larger values.

STANDARD DEVIATION

Consider the following two tables.

TABLE 1			TABLE 2	
x	f		x	f
10	6		6	2
11	8		8	5
12	12		10	8
13	8		12	10
14	6		14	8
			16	5
			18	2

Because the frequency distributions are the symmetrical about x=12 in both cases, it is clear that 12 is the mean value of the x for both. Both samples have 40 items in them ($\sum f = 40$). In the first table, the highest value (14) is only 40% higher than the lowest. However, in table 2, the highest is 200% bigger than the lowest. Despite the fact that two key elements are identical, the second set of data has a wider **distribution** when each is plotted on a graph - with f (vertically) and x (horizontally). It would be useful to have a measure that would show this width of distribution.

Standard deviation (**σ**) does this. It is defined as:

$\sigma = \sqrt{[\sum (x-m)^2 / (n-1)]}$ where m is the mean and n the number of items in the sample.

Where a frequency table is provided, the standard deviation is defined as:

$\sigma = \sqrt{[\sum f(x-m)^2 / (\sum f - 1)]}$ if the mean m has already been calculated,

otherwise, it can be calculated as:

$\sigma = \sqrt{[[\{ \sum x^2 - (\sum x)^2 / \sum f) \} / (\sum f - 1)]}$ where the mean is effectively calculated within the formula.

All these versions of **σ** are best calculated using additional columns of the data table. Clearly, a spreadsheet is an ideal to perform this number manipulation because of the ease in which columns can be totalled.

CORRELATION

Candidates are not expected to be familiar with correlation formulae but should have an understanding of the meaning of correlation and that there is a **correlation coefficient** (**r**) that measures the extent of correlation.

Correlation is a measure of whether there is a relationship between one set of data and another set.

e.g.

1. Can it be stated with confidence that a person able in mathematics would also be able in physics?
2. Do tall people have larger feet?

Having two sets of data for as large a sample as possible could test both of these. These tables would hold:

1. Maths and physics scores in parallel examinations where candidates take both subjects.
2. Heights of people and their shoe size (or length of foot measured in cms).

The correlation coefficient can be calculated from the two sets of data using an involved formula not required here. This coefficient will have a value between –1 and +1.

Interpretation of the coefficient.

r = +1 Perfect correlation. This implies that if you knew ONE of the values, you could accurately predict the other. If there were perfect correlation between ability in mathematics and physics, then if the candidates were stood in order of their marks for mathematics, they would be in the same order for physics. Perfect correlation means more than this. A straight-line graph could be drawn for these marks with an

upward slope. If the maths mark for a candidate is known, the physics mark could be read from the graph. The candidate would only need to sit one examination to obtain two marks! Researchers always hope to reach this situation in their work.

r = -1 Perfect INVERSE correlation. This implies the reverse effect. If one value is high, the other will be low. Examples of this are more difficult to find in the real world. Perhaps if the air temperature was recorded at different heights up a mountain, there might be perfect correlation? Perfect inverse correlation would still produce a straight line if the two sets of data were plotted on a graph. However, the line would slope downwards instead of up. This is also helpful to researchers because at least results are predictable.

r = 0 No correlation. This implies that there is no relationship between the two sets of data. Knowing one value would not enable the other to be predicted. Examples are easy to find. The length of a person's hand would have no bearing on how many books he read last year. Obtaining a coefficient of 0 is the researcher's nightmare because it means nothing is predictable. It could only be useful if there had been a suggestion that two events were related – zero correlation coefficient would strongly suggest there in NO relationship.

In practice, NONE of these situations will occur. There will always be variations. The coefficient for physics- mathematics example may be 0.95. This is a high value and clearly implies the two are closely linked. A person opting to study for higher physics qualifications would be advised not to if his/her mathematics ability were poor. Note: there will always be some a small number of people who are good at one and not the other.

Classwork. Try to set examples that are related to life rather than provide a series of meaningless numbers. With real data, the final results can be discussed as to their meanings and whether the calculated values are actually reasonable. Students could select their own proposition to investigate provided it is chosen sensibly so that the data can be measured or obtained.

PROBABILITY

Candidates must be aware that it is not possible to calculate exact figures for most aspects of life. Is it likely that children of tall parents will be tall. How do you define "tall"?

Introductory class exercises.

1. Discuss with the class how the theory that tall parents produce tall children could be tested.
2. Discuss each of the following statements with emphasis on how likely each is. Are there any special factors which would influence the result in each case?
   a) The next baby born in this town will be a boy.
   b) I will receive a letter tomorrow.
   c) My country will win the next World Cup in football.
   d) Using a standard 6-sided die with the faces holding the numbers 1 to 6,
      i. I will roll a 6 in the next throw
      ii. I will roll a 6 within the next 6 throws.
   e) I have a rectangular piece of card, which if placed on an 8 x 8 chessboard, will exactly cover two squares. The chessboard squares are alternatively black and white. I place the card onto the board to cover two squares. The squares covered are:
      i. Both black or both white.
      ii. One is black and one is white.

Experimental tools

1. 6-sided die as above
2. Coin with a "head" or some national symbol on one side and another design on the back. In many countries these are referred to as "heads and tails".
3. Pack of 52 playing cards divided into 4 suits – Spades and Clubs (black), Diamonds and Hearts (red). Each suit has 13 numbered cards 1 to 10 and three higher valued cards Jack/Knave (J), Queen (Q) and King(K) – these three are often called COURT cards. In many games, The 1 is usually referred to as the Ace. In some games, Aces are considered to be of highest value card and rank above the Kings. The Court cards AND Aces together are sometimes called HONOURS.

Definition:

Probability of an event = $\frac{\text{Number of favourable outcomes}}{\text{Number of ALL possible outcomes}}$

Probabilities can be expressed as fractions to give an exact value or as a decimal fraction which may be rounded to a limited number of decimal places.

Examples:

1. Probability of drawing a card at random from the pack and that card is
   a) "7" = 4/52 = 1/13 - There are four 7's in the 52 card pack, one in each suit.
   b) Heart = 13/52 = ¼ - There are 13 cards in each suit.
   c) an "Honour" (defined as A,K,Q,J) = 16/52 = 4/13. There are 4 of each, one in each suit.
2. Vegetable seeds are planted in two areas of ground. The two areas are A (6m by 4m) and B (2m by 3m). What is the probability that the first growth will show in A? This assumes the areas are equally likely to produce the growth – similar light, watering etc and the seeds are sown in the same manner. The probability for A is related here to area. The bigger the area, the more seeds can be sown. In terms of areas, A = 6x4 =24m^2 and B = 2x3 = 6m^2.

   Probability that A shows the first growth = Area of A / Total Area = 24/30 = 4/5 = 0.8

Discussion

At this stage, the meaning of probability can be discussed with the following points being brought out.

1. Probability is a notional idea that can give a reasonable prediction but usually over many trials. If a coin is tossed and produces three heads in succession, it does not mean that tails is more likely to appear on the next throw. Each event is totally independent. If 10 heads appear in succession, there is some suspicion that the coin is not normal – does it have two heads?
2. The probability that an event WILL occur PLUS the probability that it WILL NOT occur is 1. In tossing a coin, the probabilities of a head and a tail are each ½ because there is no alternative – or is there? See 3.a) below.
3. Probability is a simple measure and cannot take into account all possible factors that could affect the outcome.
   a) It is possible (although unlikely) that a coin could land vertically on its edge and not on one face - it would not then be considered to be a head or a tail.
   b) In the "seed" example above, all the seeds could be faulty and no plants grow. A small patch of one of the two areas could be more fertile because of previous use of the land.
   c) A card in a much-used pack could be sticky. This might prevent it from being selected because it sticks to the card above or make it more likely to be picked because it sticks to the fingers of the experimenter.

Complex Situations

Sometimes, probabilities involve several events.

Independent events are events where one has no effect on the other. In this case the probability of both the events occurring is the PRODUCT of their individual probabilities.
e.g. Throw a die and take a card from a pack. What is the probability of throwing a 6 on the die and also selecting a 6 on the card?
$P_{die=6}$ = 1/6 $P_{card=6}$ = 4/52 = 1/13 $P_{die=6 \& card=6}$ = 1/6 x 1/13 = = 1/78
In general terms P(AB) = P(A) x P(B) where P(AB) is the probability of both events and P(A) and P(B) are the two probabilities that the two events will occur **independently**.

Conditional Probability relates to one event occurring as a result of a previous event already having occurred.
e.g. Probability of selecting two cards from a pack which are both Hearts.
First card is a Heart P_H = 13/52 = ¼ as above. There are now only 51 cards left
In considering the second card, there is an assumption the first WAS a Heart for both to be Hearts. This means there are only 12 Hearts left. This probability is 12/51 = 4/17
So the probability of 2 Hearts is ¼ x 4/17 = 1/17
This situation is usually written P(AB) = P(A) x P(B/A)

P(AB) is the probability of two Hearts, P(A) is here the probability that the first card is a Heart, P(B/A) is the probability the second is a Heart given that event A (the first was a Heart) has already occurred.

Had a similar experiment performed by throwing two dice, this would be independent events. A selected card is not replaced in the pack but with dice, a number cannot be taken off the die leaving it as a 5-sided die with only 5 numbers.
The question implies that once the first has been selected, it is NOT replaced. If it is, a different result is obtained because the SAME card could be drawn the second time.

Another situation occurs where a situation is satisfied provided at least ONE OF TWO events occurs. This is sometimes written as P(A+B) = P(A) + P(B) – P(AB).

Consider the two situations:
A and B are independent: An example best illustrates this.
e.g. Find the probability of drawing a Heart from a pack OR throwing a 6 on a die.

P_H = 1/13 as before. P_6= 1/6. Clearly the probability rises in this example because there are two opportunities to satisfy the conditions. The probability of one or both occurring is 1/13 + 1/6. There is no dependence of one on the other and P(AB), the probability of selecting a Heart from a pack will not cause a die to show a 6. P(AB) = 0 and is not really defined for this situation.
This can be easily understood by considering throwing a die once. Probability of throwing each of the 6 numbers is 1/6. The probability of throwing 1, 2 or 3 is clearly half the possibilities and 1/6 + 1/6 + 1/6 = ½. The probability of ANY number is 1/6+1/6+1/6+1/6+1/6+1/6 = 6/6 = 1 which is fairly obvious.

A and B dependent on each other is a different situation. Again an example illustrates.
e.g. Find the probability of selecting a SINGLE card which is either a Court card (JQK) or a Heart. 3 of the 12 Court cards are Hearts so there is overlap. If P(A)= probability of selecting a Heart and P(B)= probability of a court card, then P(A) = 13/52 =1/13 P(B) =12/52 = 3/13
P(B/A)= 3/13 (three of the 13 hearts)
and so from above P(AB) = P(A) x P(B/A) = 13/52 x 3/13 = 3/52.
Also from above, P(A+B) = P(A) + P(B) – P(AB) = 1/13 + 3/13 – 3/52 = 22/52 = 11/26
In this instance, it would have been better NOT to cancel down the probabilities but to work in fractions of 1/52 until the final line.

Mutually Exclusive events are events that CANNOT occur at the same time. In throwing a die, a 4 AND a 5 cannot occur at the same time in a SINGLE throw. It therefore follows that P(A) x P(B)= 0 where event A is a 4 being thrown and B is a 5 being thrown.

It sometimes pays to consider reverse situations, particularly where two separate events are involved. To find the probability that neither A nor B occur = 1 – probability of one AND/OR the other occurring.
It is a useful exercise to try this and calculate the various parts for A = throw of 6 on a die and B = select a red card from a pack.

EXPECTED VALUE

Throw a standard die 120 times and theoretically, you would expect to count 20 occurrences of each of the 6 numbers. In practice, this will probably not occur. The results could be:

Number thrown	Observed frequency (O)	Expected frequency (E)	O - E	$(O - E)^2$	$(O - E)^2/E$
1	12	20	-8	64	3.2
2	22	20	2	4	0.2
3	18	20	-2	4	0.2
4	28	20	8	64	3.2
5	26	20	6	36	1.8
6	14	20	-6	36	1.8
Σ	120	120			10.4

Column 2 shows the actual frequency of each number occurring in 120 throws. Column 3 shows how many are expected. In throwing a die, each side is equally likely. A measure of discrepancy COULD be shown by column 4, the difference between actual and expected values. However, to add this column would be pointless because it clearly totals zero in this situation – over-throwing on one number will compensate any under-throwing on another number. Column 5 removes the compensation by squaring the differences. This column has no real meaning BUT it is a measure of the discrepancy from the expected.

However, this is unsatisfactory because it measures actual discrepancy without comparing it with the data. We would obtain the same answer if all the numbers in columns 2 and 3 were 100 larger. We expected 20 1's but we only threw 12. The difference 8 is quite large compared with the expected 20. Throwing 112 1's when we expected 120 is not such a relative difference. Hence, column 6 is calculated dividing the squared differences by expected frequency. The total is a measure of discrepancy.

NOTE: In other situations, the figures could be different. If we had a 6-sided die but instead of a 6 on one face, another 5 were painted in, the expected number of 6's would be 0 (so would the observed because it would not be possible to throw a 6). However, the expected number of 5's would now be 40.

Exercise

Repeat this exercise with each value in columns 2 and 3 being increased by 100 and determine the measure of discrepancy with these figures. The actually differences are identical to the above figures but relatively, the discrepancy in the second case is less of a problem.

Performing an exercise like this assists the statistician to decide whether there might be something wrong in the way that the testing performed or that original assumptions were wrong. For the die situation, surely the die should show an equal number occurrences of each face? Surely the die is totally fair with each face the same? In fact, if the numbers on each face are represented by a series of dots produced by drilling small indentations, then the faces are not identical. The centre of gravity of the die would not be EXACTLY at the centre of the cube however well that was made. This very small difference is bound to have some long term effect, if small, on results.

PERMUTATIONS AND COMBINATIONS

Permutations relate to the number of different ways in which items from a population can be arranged where ORDER is taken into account.

1. The letters A, B and C can be arranged in the following ways ABC, ACB, BAC, BCA, CAB, CBA.

   This can be worked out as follows:
   First letter: There are 3 ways to select – A , B or C.
   Second letter: Once one has gone, there are two left.
   Third letter: There is no choice. There is only one letter left.
   So, Permutation of 3 from 3 is 3 x 2 x 1 = 6 as above. The numbers are multiplied because the choice of 2nd and 3rd places are not affected by earlier ones (given that we have already taken into account the reduction in numbers after each selection)

2. There are 10 runners in a race. How many can ways are there to award the three medals (gold, silver and bronze)?
   As above, 10 could gain the gold leaving 9 who could each earn the silver and hence 8 to win the bronze.
   A general formula for this situation uses the symbol $^{10}P_3$ meaning permutate any 3 from 10.

   $^{n}P_r = \frac{n!}{(n-r)!}$ where n! is called FACTORIAL n and represents n x (n-1) x (n-2) x…x 3 x 2 x 1

   Many of the factors cancel in most problems.
   In the race, $^{10}P_3 = 10! / 7! = \frac{10 \times 9 \times 8 \times 7 \times 6 \times 5 \times 4 \times 3 \times 2 \times 1}{7 \times 6 \times 5 \times 4 \times 3 \times 2 \times 1} = 10 \times 9 \times 8 = 720$

3. Duplicates can sometimes occur.
   e.g. How many ways are there to arrange the letters of the word BANANA? Suppose the two N's are relabelled as N_1 and N_2. Similarly, the A's could be labelled A_1 to A_3. IF the letters had been all different, the answer would be $^{6}P_6 = 6! / 0! = 6! = 720$. Factorial 0 should be taken to have the value 1. However, the N_1 and N_2 are not different – we could not distinguish between them. We can arrange these two N's AMONGST THEMSELVES in 2! ways. Similarly, the 3 A's can be arranged amongst themselves in 3! ways.
   So there are 6! / (2! x 3!) ways to arrange the letters of BANANA = 60

4. Combining the last two situations.
   e.g. How many ways are there of arranging 4 letters taken from EMBARKATION?
   We need a general method to calculate the answer because it is easy to miss one or two if an attempt is made to write out all the possibilities.
   Note, one letter (A) is repeated. The results in a:
   Code with 2 A's will have $^{11}P_4 / 2!$ (because there is no distinction between the 2 A's)
   Code with 1 A will have $^{10}P_4$ - there are 9 different letters + an A.
   Code with no A will have $^{9}P_4$ - only 9 letters when no A's are considered.
   The total number of code = sum of these three quantities.
   Question: How many different 4-letter codes can be made from BANANA? Not easy!

Exercises

Practical problems to investigate

1. Codes (e.g. PIN numbers)
2. Travel routes from A to C via B when there are several ways A to B and several ways B to C.
3. Arranging people round a round table. What if they have to alternate man, woman?
4. Selecting a football team of 11 players from a squad of 20 players. Some positions in the team (e.g. goalkeeper) would have to be filled with a specialist player.

Combinations **relate to the number of ways in which items can be selected from a population where order does NOT matter.**

1. Consider question 2 above. How many different groups of 3 people could take the 3 medals? Now we are making no distinction between gold, silver and bronze. If the runners are labelled A to J, then whereas ABC is a different result from ACB when order matters, it is the same if order does not matter – the same three runners win the medals. The combination result is therefore the same as the permutation but divided by the number of ways in which the 3 runs can be arranged.
   ${}^{10}C_3 = 10! / 7! / 3!$
   A general formula for this situation uses the symbol ${}^{n}C_r$ meaning select any r from n.
   $${}^{n}C_r = \frac{n!}{(n-r)!\, r!}$$
2. Some competition entries wrongly use the phrase “permutate any 8 for 12” when they mean “select any 8 from 12” and the order does not matter. It is important to use the correct terminology.