MATH6038 — Mathematics for Science 2.2 with Maple

J.P. McCarthy

February 29, 2012

Contents

0.1 Introduction

Lecturer

J.P. McCarthy

Office

Meetings before class by appointment via email only.

Email & Web:

jippo@campus.ie and *http://irishjip.wordpress.com*
This page will comprise the webpage for this module and as such shall be the venue for course announcements including a definitive date for the test. This page shall also house such resources as links (such as to exam papers), as well supplementary material. Please note that not all items here are relevant to MATH6038; only those in the category 'MATH6038'. Feel free to use the comment function therein as a point of contact.

Module Objective

This modules involves the study of matrices, statistics and probability distributions.

Module Content

Matrix Algebra

Matrix operations, properties of matrix operations, determinants, properties of determinants, row operations, Gaussian elimination, inverse matrices, solving linear system of equations, investigation of the solution space of linear system of equations.

Probability and Statistics

Presentation and analysis of data. Measures of central tendency; mean, mode and median. Measures of dispersion; range variance and standard deviation. Sample space, compound events, conditional probability, independent events, reliability block diagrams,Bayes Rule. Random variables, binomial, Poisson and Normal distributions. Introduction to sampling, confidence intervals for large and small samples. Construct and interpret quality control charts.

Assessment

Total Marks 100: End of Year Written Examination 70 marks; Continuous Assessment 30 marks.

Continuous Assessment

The Continuous Assessment will be divided between a one hour written exam in Week 7 worth 20% and your weekly participation in the Maple Lab (worth 10%).

Absence from a test will not be considered accept in truly extraordinary cases. Plenty of notice will be given of the test date. For example, routine medical and dental appointments will not be considered an adequate excuse for missing the test.

Lectures

It will be vital to attend all lectures as many of the examples, proofs, etc. will be completed by us in class.

Maple Labs

Maple Labs will commence next week and are designed to explore and reinforce mathematical concepts.

Exercises

There are many ways to learn maths. Two methods which aren't going to work are

1. reading your notes and hoping it will all sink in
2. learning off a few key examples, solutions, etc.

By far and away the best way to learn maths is by doing exercises, and there are two main reasons for this. The best way to learn a mathematical fact/ theorem/ etc. is by using it in an exercise. Also the doing of maths is a skill as much as anything and requires practise.
I will present ye with a set of exercises every week. In this module the "Lecture-Supervised Learning" is comprised of you doing these exercises, giving them to me on a weekly basis, me marking them, and returning them. In addition I will provide a set of solutions online. Everyone shall have access to the solution sets however.
The webpage may contain a link to a set of additional exercises. Past exam papers are fair game. Also during lectures there will be some things that will be *left as an exercise*. How much time you can or should devote to doing exercises is a matter of personal taste but be certain that effort is rewarded in maths.

Reading

Your primary study material shall be the material presented in the lectures; i.e. the lecture notes. Exercises done in tutorials may comprise further worked examples. While the lectures will present everything you need to know about MATH6038, they will not detail all there is to know. Further references are to be found in the library. Good references include:

- J. Bird, 2006, *Higher Engineering Mathematics*, Fifth Ed., Newnes.
- A. Croft & R. Davison, 2004, *Mathematics for Engineers — A Modern Interactive Approach*, Pearson & Prentice Hall,

The webpage may contain supplementary material, and contains links and pieces about topics that are at or beyond the scope of the course. Finally the internet provides yet another resource. Even Wikipedia isn't too bad for this area of mathematics! You are encouraged to exploit these resources; they will also be useful for further maths modules.

Exam

The exam format will roughly follow last year's. Acceding to the maxim that learning off a few key examples, solutions, etc. is bad and doing exercises is good, solutions to past papers shall not be made available (by me at least). Only by trying to do the exam papers yourself can you guarantee proficiency. If you are still stuck at this stage feel free to ask the question come tutorial time.

0.2 Motivation: Network Flows & How to Make a Decision.

Coughs seem very common here, especially among the children, though people look strong and healthy, but in the absence of proper statistics one cannot undertake to say whether the district is a healthy one or not.

Edward Burnett Tylor

Network Flows

Suppose we have a network of one-way streets as shown in the diagram:

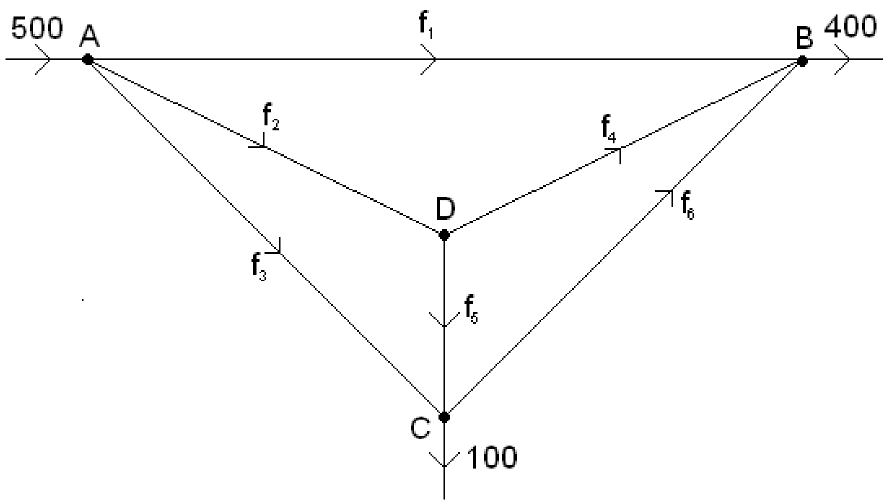


Figure 1: The flow of cars into junction A is 500 cars per hour, and 400 and 100 cars per hour emerge from B and C. Suppose the flows along the streets are $f_1, \ldots, f_6$ cars per hour. From the simple rule that the flow into a junction must equal to flow out, how much can we tell about the internal workings of the network?

Equating the flow in with the flow out at each junction we get:

$$\begin{array}{lccc} \text{Junction } A & 500 & = & f_1 + f_2 + f_3 \\ \text{Junction } B & f_1 + f_4 + f_6 & = & 400 \\ \text{Junction } C & f_3 + f_5 & = & f_6 + 100 \\ \text{Junction } D & f_2 & = & f_4 + f_5 \end{array} .$$

This gives four equations in six variables $f_1, \ldots, f_6$:

$$\begin{aligned} f_1 + f_2 + f_3 &= 500 \\ f_1 + f_4 + f_6 &= 400 \\ f_3 + f_5 - f_6 &= 100 \\ f_2 - f_4 - f_5 &= 0 \end{aligned}$$

In the first section of the module we will learn how to solve these types of equations. We will find out if there are many solutions, a unique solution, or indeed no solution at all.

How to Make a Decision.

Suppose that you are the 'supervisor' in a large factory with a production line. When attendance is low, the production line system doesn't work as well, there are issues with health and safety and efficiency is reduced. There is going to be more than just a production line in the factory; there will be a cleaning section, an admin section, a loading section, etc.; and if attendance is particularly low it might be prudent to call off production and instead reallocate these workers to assist in these other departments.

Now there is a balance to achieve — we can't afford to have the production line closed very often and neither do we want the production line to operating at a 'half-assed' capacity. Suppose the order come in from the main office that on 10% of the days, the production line is to be closed down, and the workers reallocated to other tasks. The decision we need to make is; what days should we close the production line? We want to close on days when attendance is low. We make assumptions such that we are not in Ireland so everyone isn't calling in sick on a Monday...

The answer is that we use statistics to do it. Essentially statistics is the science of data so the very first thing we need to do is get data on absenteeism from the HR department. From this data we will calculate the *average* attendance and the *standard deviation*. The standard deviation is a measure of, on average, how spread out the data is, are there large deviations from the average, or small deviations? Now a central plank of statistics theory: for many types of numerical data, there is an average which is also the outcome that occurred most often, and we are just as likely to be larger than the average as smaller than the average. When the number of employees is large, the attendance is like this. On most days there is a middling number of people in, some days a lot of people are absent, and on some days nearly everyone is in.

Many examples of this unimodal (peaked, bell-shaped), symmetric (about the peak) kind of data can be shown to be of a certain form. Such data is called *normal* and we say it has a *normal distribution*. This kind of idea can be illustrated in a picture:

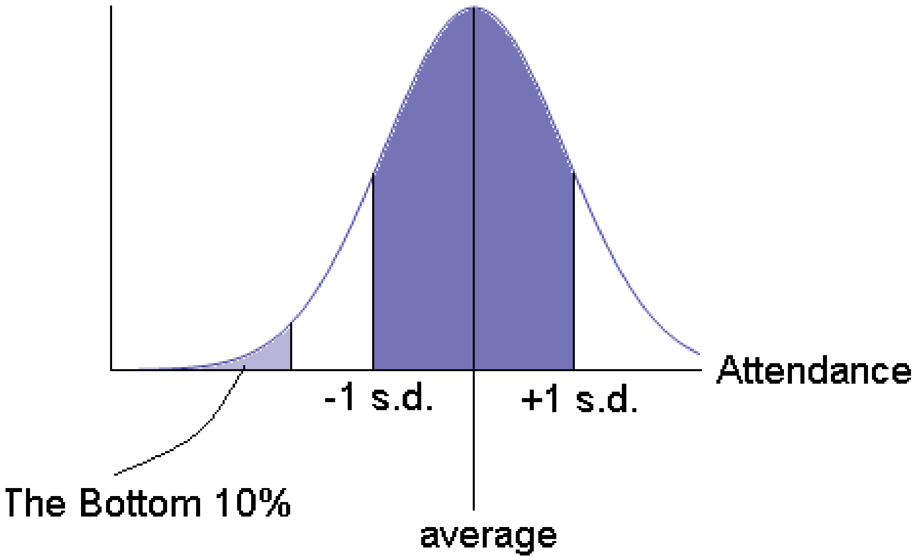


Figure 2: Note that for a normal distribution approximately 68% of the data is found within a distance of one standard deviation from the average.

We will discover in this module how to find the attendance level corresponding to the 'bottom' 10%. For example, if the average attendance is 100 with a standard deviation of 15 then we can show that if the attendance falls below 80.8 (i.e. 80 or less) workers then the production line should lay idle. So this gives a policy — head-count at 8.10 AM — if there are less than 80 workers, the production line is not operated. This will ensure that, *average*, the production line is closed one in every ten days.

Chapter 1

Matrix Algebra

It is my experience that proofs involving matrices can be shortened by 50% if one throws the matrices out.

Emil Artin.

In this chapter we learn how to solve and analyse equations such as the those generated by the network flow question. Such a set of equations is known as a *system of linear equations*. For now a matrix is just a rectangular array of numbers in a bracket but later we will see their true nature.

1.1 Systems of Linear Equations

If two lines intersect they will do so at a single point; if two planes intersect their intersection will be a line, a line can intersect a plane at one point, lie in the plane, or not intersect it at all. Three planes can have one point in common or no points in common. Some of these possibilities are illustrated as follows:

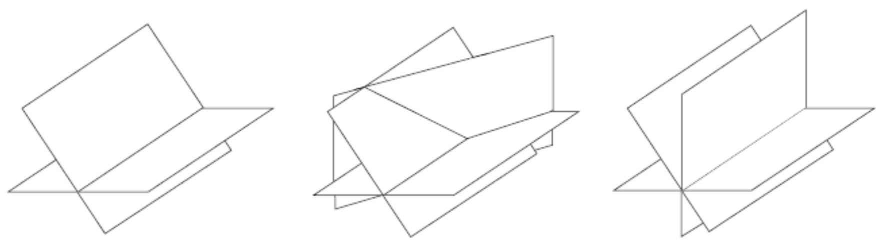


Figure 1.1: Three planes can intersect at a point, a line, or nowhere.

We can show that the *equation of a plane* is given by:

What the hell is the *equation of a plane*? In essence it is a membership card:

Hence to find the intersection of three planes we find points that are on all three curves — that is the satisfy their equations, at the same time, *simultaneously*. For example, we might have to find the points (x, y, z) that satisfy all of

$$\begin{aligned} 3x + 4y - z &= 7 \\ 2x - 6y + z &= -2 \\ x - y + z &= 3 \end{aligned}$$

1.1.1 Definitions

A *linear equation in* n *variables* is an equation of the form:

$$a_1x_1 + a_2x_2 + a_3x_3 + \cdots + a_nx_n = b, \tag{1.1}$$

where the *variables* are $x_1, x_2, \ldots, x_n$, the numbers $a_1, \ldots, a_n$ are called the *coefficients*, and b is the *constant*. A *system of* m *equations in* n *variables* has the form

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= b_2 \\ \vdots\ \vdots \quad &\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n &= b_m. \end{aligned}$$

Examples

Solve the following simultaneous equations.

1.

$$\begin{aligned} 2x - y &= 1 \\ 3x - 4y &= 9 \end{aligned}$$

Solution 1: This is the method taught in secondary schools. We will develop our method along similar lines. Multiply the top equation by -4 and add the equations together:

Now back-substitute to get into either equation to get $y = 1$.
Solution 2: This method is better for more general simultaneous equations (e.g. with x^2 and the like). Solve the first equation for $y = h(x)$:

Now *substitute* this into the second equation[1]:

Once again back-substitute to get $y = 1$.

2.

$$\begin{aligned} x + y + z &= 2 \\ 2x + y + z &= 3 \\ x - 2y + 2z &= 15 \end{aligned}$$

Solution: There is a method analogous to method 1 above but there is an easier method. Find the intersection (a line) between planes 1 & 2 by solving for z:

Now find the intersection between the planes 2 & 3 similarly:

[1]as an alternative we would have solved the second for $y = g(x)$ and set $y = y$ morryah $h(x) = g(x)$ — an equation in one variable which we can solve for x

No just back-substitute to find $z = -1$. Solution: $x = 1, y = 2, z = -1$.

1.2 Row Operations and Gaussian Elimination

While it is possible to solve systems with small numbers of equations in a few variables by *ad hoc* methods such as these, we would like a more systematic approach to solve more complex systems, and would also like to be able to program computers to do the task. We will develop an algorithmic method perfectly adapted to the task. First note that the variable names are irrelevant; the systems

$$\begin{aligned} 4x - 8y &= 1 \\ -3x + y &= -3 \end{aligned} \qquad \text{and} \qquad \begin{aligned} 4m - 9n &= 1 \\ -3m + n &= -3 \end{aligned}$$

have the same solutions[2]. Consequently all we actually need to look at are the coefficients and constants, which can be recorded in a rectangular array called a *matrix*:

$$\begin{array}{rcrcrcrcr} x_1 & + & 2x_2 & - & 6x_3 & - & x_4 & = & 0 \\ 2x_1 & + & 4x_2 & & & + & 7x_4 & = & 3 \\ 6x_1 & - & 2x_2 & + & x_3 & + & 2x_4 & = & -4 \\ 3x_1 & & & - & 8x_3 & + & 2x_4 & = & 9 \end{array} \quad \text{converts to} \quad \left[\begin{array}{cccc|c} 1 & 2 & -6 & -1 & 0 \\ 2 & 4 & 0 & 7 & 3 \\ 6 & -2 & 1 & 2 & -4 \\ 3 & 0 & -8 & 2 & 9 \end{array}\right]$$

Conversely, given such a matrix we can recover the corresponding system:

$$\left[\begin{array}{cccc|c} 4 & 3 & -7 & 1 & 0 \\ 2 & 9 & 1 & -1 & 10 \\ 8 & -2 & 0 & 5 & 0 \end{array}\right] \quad \text{converts to} \quad \begin{array}{rcrcrcrcr} 4x_1 & + & 3x_2 & - & 7x_3 & + & x_4 & = & 9 \\ 2x_1 & + & 9x_2 & + & x_3 & - & x_4 & = & 10 \\ 8x_1 & - & 2x_2 & & & + & 5x_4 & = & 0 \end{array}$$

1.2.1 Elementary Row Operations

We want to transform a given system into one which is easier to solve. There are three things which we can do to a linear system which will not change the solution, but possibly make it easier to see the solution.

- Swap equations — clearly

$$\begin{aligned} 4x - y &= 7 \\ 2x + 5y &= -2 \end{aligned}$$

  has the same solution as

$$\begin{aligned} 2x + 5y &= -2 \\ 4x - y &= 7 \end{aligned}$$

- Multiply an equation by a constant — neither will this change the solution; say multiplying the second equation by five:

$$\begin{aligned} 2x + 5y &= -2 \\ 20x - 5y &= 35 \end{aligned}$$

[2]namely $x = m = 26/23$ and $y = n = 9/23$.

- Add the equations together — why would this not change the solution?

$$2x + 5y = 1$$
$$(20x - 5y) + (2x + 5y) = 35 - 2$$

Now we would have[3] $22x = 33 \Rightarrow x = 3/2$. Note also that we could have put these last two transformations into the single *add a multiple of an equation to another*.

If we go back into the augmented matrix picture we see:

$$\left[\begin{array}{cc|c} 4 & -1 & 7 \\ 2 & 5 & -2 \end{array}\right] \overset{r_1 \leftrightarrow r_2}{\rightarrow} \left[\begin{array}{cc|c} 2 & 5 & -2 \\ 4 & -1 & 7 \end{array}\right] \overset{r_2 \rightarrow r_2 \times 5}{\rightarrow} \left[\begin{array}{cc|c} 2 & 5 & -2 \\ 20 & -5 & 35 \end{array}\right] \overset{r_2 \rightarrow r_2 + r_1}{\rightarrow} \left[\begin{array}{cc|c} 2 & 5 & -2 \\ 22 & 0 & 33 \end{array}\right],$$

and we can convert this into

$$2x + 5y = -2$$
$$22x = 33$$

Note the *row operations* we enacted. Suppose we have a system of linear equations in augmented matrix form $[A|B]$. From the discussions above we can show that the following row operations will leave the solution unchanged:

- swapping any two rows:

- multiplying any row by a constant:

- adding any row to any other:

- combining the last two: adding a multiple of a row to another row:

We call these the *elementary row operations* or EROs.

Example

Use the techniques above to simplify and hence solve the following simultaneous equation:

$$5x + 7y = 0$$
$$-3x + 4y = 2$$

Solution: First we write everything in augmented matrix form:

[3][Ex]: from this find y

What I am going to do is try to use the EROs to get the augmented matrix in the form

Hence we now have

Using the three EROs we want to take the augmented matrix form of the linear system and apply the EROs until the coefficient matrix is in *row-echelon form*. This looks like

In words,

1. all rows containing zeros are on the bottom.
2. all the *leading coefficients* (of the non-zero rows) are 1 and above zeros.

   The coefficient matrix is in *reduced row-echelon form* if, in addition

3. the leading coefficient or pivot is the only non-zero entry in its column.

The following matrices are in row-echelon form (where $\star$ denotes *any* number):

To be in reduced row-echelon form they must look like:

The following matrices are not in row-echelon form:

but can easily be brought into row-echelon form by applying EROs.

1.2.2 The Solution Space

As soon as the coefficient matrix is brought into row-echelon form we can tell if solutions exist for the system, and if so whether there is just one solution, or infinitely many. There will be no solution if there is a row looking like

This follows because this particular row corresponds to the equation

which has no solution since the left-hand side is zero but the right-hand side is $k \neq 0$.

If no such row appears then there is at least one solution. It is unique if every column contains a pivot; if this is not so then the variables corresponding to the columns without pivots are not determined and become parameters/free variables in the solution leading to infinitely many solutions.

Examples

The augmented matric for the three linear systems has been brought into reduced row-echelon form. Find the solutions:

$$(i)\left[\begin{array}{ccc|c}1&0&0&-1\\0&1&0&0\\0&0&1&2\end{array}\right],\quad(ii)\left[\begin{array}{cccc|c}1&3&0&1&3\\0&0&1&7&1\\0&0&0&0&2\end{array}\right],\quad(iii)\left[\begin{array}{ccccc|c}1&5&0&0&-2&3\\0&0&1&0&4&-5\\0&0&0&1&2&6\\0&0&0&0&0&0\end{array}\right].$$

Solution:

(i) We simply have

(ii) Note the third row...

(iii) Rewrite this set of equations:

$$\begin{aligned}x_1+5x_2-2x_5&=3\\x_3+4x_4&=5\\x_4+2x_5&=6\\0x_5&=0\end{aligned}$$

No solve from the bottom. Firstly x_5 could be anything. We write this by saying $x_5=t$ for $t\in\mathbb{R}$. This means that x_5 can take on any real number ($\mathbb{R}$) value. In this case, x_5 or t is called a *parameter* or *free variable*. For each value of the parameter (t), we get a different solution. As t can take on any value from minus to plus infinity, there are thus an infinite number of solutions. Now we look at the second last row:

Now at the third last:

Now look at the first equation:

Now for any fixed value of t, $x_1=-5x_1+(3+2t)$ actually represents a line and thus there are an infinite number of pairs (x_1,x_2) that satisfy this equation. We need another parameter/free variable. In this we could choose x_1 or x_2 but usually we will be better off if we pick the x_2 (e.g. the x_5 over the x_3 etc.) Hence now call $x_2 = s$ — where again $s \in \mathbb{R}$:

Hence we have the solution(s):

$$\begin{aligned} x_1 &= 3 + 2t - 5s \\ x_2 &= s \\ x_3 &= -19 + 8t \\ x_4 &= 6 - 2t \\ x_5 &= t \end{aligned}$$

where t, $s \in \mathbb{R}$. You might (not?) be interested to show that we have shown that three five-dimensional (hyper) planes can intersect along a plane...

How did we know that x_5 and x_2 were parameters/free variables? Why did we need two parameters? Why did we need any? The following theorem is useful in this case. We will not provide a proof.

1.2.3 Theorem

Consider a linear system of m equations in n variables. Suppose that the coefficient matrix has r non-zero rows when put in row-echelon form. Then if there are solutions, *the set of solutions has $n - r$ parameters.* **In particular, if $r < n$,** *then there will be infinitely many solutions.*

Remark

It follows that we have three possibilities:

(i) there is no solution (the system is *inconsistent*), or

(ii) there is exactly one solution ($n = r$), or

(iii) there are infinitely many solutions ($n > r$)

The number r represents the number of independent equations. Consider the three equations:

$$\begin{aligned} 2x - y &= 4 \\ x - 6y &= 1 \\ 2x - 12y &= 2 \end{aligned}$$

Although there are three equations here, equations 2 and 3 are actually equivalent — in row echelon form these would form a row of zeros.

Hence once we have the augmented matrix in row-echelon form we must see how many parameters/free variables there are (in this module it will usually be zero, one (t) or two (t and s)). Usually we look at the augmented matrix and correspond rows to variables. If there is no row for the last equation we can usually take that variable to be a parameter/free variable.

Note that this will not always be possible. For example,

Here there is one parameter/free variable but we can't say that x_3 is a parameter/free variable — $x_3 = 1$.

The coefficient matrix can always be brought into row-echelon form by using the following *Gaussian Elimination* algorithm.

1. If possible, swap rows such that the first entry of the first row is $a \neq 0$.
2. Multiply the first row by $1/a$ in order to get a leading 1.
3. Subtract multiples of this row from those below to make each entry below this into a zero.
4. Repeat steps 1-3 for the second entry in the second row, third entry in the third row etc.

When $n = r$ (so essentially each row has a leading 1), we will also be able to put the matrix in *reduced row-echelon form* by the *Gauss-Jordan Elimination* algorithm.

1. Apply Gaussian Elimination.
2. Assuming $n = r$, delete all the zero rows. Add minus the last row to the second last row.
3. Add minus the last row and minus the second last row to the third last row.
4. Repeat this procedure for the rest of the rows so that the coefficient matrix is all zeros apart from ones along the diagonal.

Gauss-Jordan Elimination looks like this:

Now the solutions are easy to see.

Examples

Solve the following systems of linear equations using row reduction.

1.

$$\begin{aligned} x + 2y &= 2 \\ 2x - y &= 1 \end{aligned}$$

Solution:

2.

$$\begin{aligned} x + 2y - z &= 2 \\ 2x + 5y + 2z &= -1 \\ 7x + 17y + 5z &= -1 \end{aligned}$$

Solution:

3.

$$
\begin{aligned}
x + 10z &= 5 \\
3x + y - 4z &= -1 \\
4x + y + 6z &= 1
\end{aligned}
$$

Solution:

4.

$$
\begin{aligned}
x + 2y - 4z &= 10 \\
2x - y + 2z &= 5 \\
x + y - 2z &= 7
\end{aligned}
$$

Solution:

Now the number of non-zero rows, $r = 2$; while the number of variables, $n = 3$. Hence there is $n - r = 1$ parameter. Let $z = t \in \mathbb{R}$:

Hence we have the solution $x = 4 - 8t$, $y = 3 + 2t$ and $z = t$ for $t \in \mathbb{R}$.

Summer 2011 Question 2(a)

Use only the Gauss-Jordan method to determine the solution set S for each of the following systems of linear equations. Clearly describe the solution set S in each of the three cases.

$$(A)\ \begin{array}{ccccc} x & + & 3y & = & 4 \\ 4x & + & 12y & = & 17 \end{array}\ (B):\ \begin{array}{ccccc} x & + & 2y & = & 3 \\ 2x & + & 4y & = & 6 \end{array}\ (C):\ \begin{array}{ccccc} x & + & 3y & = & 2 \\ 4x & + & 18y & = & 16 \end{array}.$$

Solution: (A)

Hence the solution set is empty.
(B)

Now $n - r = 2 - 1 = 1$ so we have one parameter. Let $y = t \in \mathbb{R}$. Hence $x = 3 - 2t$. Ans: $S = (x = 3 - 2t, y = t)$.
(C)

Summer 2011 Question 2(c)

Given the following row reduced augmented matrix, write down the associated linear system of equations in terms of the variables x_1, x_2, x_3 and x_4. Identifying the free variable and express the solutions set in terms of the parameter t.

$$\left[\begin{array}{cccc|c} 1 & 0 & 4 & 0 & 5 \\ 0 & 1 & 9 & 0 & 3 \\ 0 & 0 & 0 & 1 & 8 \end{array}\right].$$

Solution:

As $n - r = 4 - 3 = 1$ there is a parameter/free variable. Clearly this can't be x_4 as $x_4 = 8$. Let $x_3 = t$. Now

Exercises

1. Write a system of linear equations corresponding to each of the following augmented matrices.

$$(i)\ \left[\begin{array}{ccc|c} 1 & -1 & 6 & 0 \\ 0 & 1 & 0 & 3 \\ 2 & -1 & 0 & 1 \end{array}\right] \quad (ii)\ \left[\begin{array}{ccc|c} 2 & -1 & 0 & -1 \\ -3 & 2 & 1 & 0 \\ 0 & -1 & 1 & 3 \end{array}\right].$$

2. Find all the solutions (if any) of each of the following systems of linear equations using augmented matrices and Gaussian elimination:

$$(i)\ \begin{array}{r} x + 2y = 1 \\ 3x + 4y = -1 \end{array} \quad (ii)\ \begin{array}{r} 3x + 4y = 1 \\ 4x + 5y = -3 \end{array} \quad (iii)\ \begin{array}{r} 3x - 2y = 5 \\ -12x + 8y = 16 \end{array}$$

$$(iv)\ \begin{array}{r} 2x + y + z = -1 \\ x + 2y + z = 0 \\ 3x - 2z = 5 \end{array} \quad (v)\ \begin{array}{r} -2x + 3y + 3z = -9 \\ 3x - 4y + z = 5 \\ -5x + 7y + 2z = -14 \end{array} \quad (vi)\ \begin{array}{r} 3x - 2y + z = -2 \\ x - y + 3z = 5 \\ -x + y + z = -1 \end{array}$$

3. Consider the following statements about a system of linear equations with augmented matrix A. In each case decide if the statement is true, or give an example for which it is false:

   (a) If the constants are all zero then the only solution is the zero solution (all variables equal to zero).

(b) If the system has a non-zero solution, then the constants are not all zero.

(c) If the constants are all zero and there exists a solution, then there are infinitely many solutions.

(d) If the constants are all zero and if the row-echelon form of A has a row of zeros, then there exists a non-zero solution.

1.3 Matrices

For now, a *matrix* is a rectangular area of numbers in a bracket.

Examples

$$A = \begin{pmatrix} 1 & 0 \\ 2.6 & -8 \end{pmatrix}$$

$$B = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

$$C = \begin{pmatrix} 1 & 0 & 3 \\ -16 & 0 & \sqrt{26} \end{pmatrix}$$

Remarks

1. A matrix with n rows and m columns is said to have *dimension* $n \times m$ or be an $n \times m$ matrix. For example, A is a 2×2 matrix; B is a 2×1 matrix, and C is a 2×3 matrix.
2. A *square* matrix is an $n \times n$ matrix.
3. The (i, j)-entry of a matrix is the number in the ith row and jth column.

1.3.1 Addition of Matrices

Two matrices of equal dimension may be added together to produce another matrix of the same dimension. This sum is a matrix whose elements are obtained by adding corresponding elements.

The *zero matrix* is denoted $\mathbf{0}$, and has only 0 as its entries. It satisfies

Just like zero for the real numbers.

1.3.2 Scalar Multiplication of a Matrix

Any matrix may be multiplied by a scalar (some $k \in \mathbb{R}$) by multiplying each element by the number.

By definition $-A = (-1)A$, so that $A - B$ means $A + (-B)$. Properties of matrix addition and scalar multiplication include:

$$A + B = B + A; \quad (A + B) + C = A + (B + C); \quad k(A + B) = kA + kB;$$

$$(k + l)A = kA + lA; \quad (kl)A = k(lA); \quad A - A = 0; \quad 0A = \mathbf{0}.$$

Note that these mirror the properties of ordinary addition and multiplication.

If A is an $m \times n$ matrix then the *transpose* of A, denoted A^T, is the $n \times m$ matrix whose got by exchanging the rows and columns of A. Properties of the transpose operation include:

$$(A^T)^T = A; \quad (kA)^T = kA^T; \quad (A + B)^T = A^T + B^T.$$

1.3.3 Equality of Matrices

Two matrices are *equal as matrices* if they have same dimension and each corresponding element is equal.

Example

Suppose

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}, \text{ and}$$
$$B = \begin{pmatrix} 1 & 0 \\ 2.6 & -8 \end{pmatrix}$$

and one is told $A = B$. Thence $a = 1, b = 0, c = 2.6$ and $d = -8$.

Examples

Solve the following equations in each case to find the matrix A.

1. $A^T + \begin{bmatrix} 0 & 9 & 5 \\ 3 & -7 & 19 \end{bmatrix}$.

2. $\left(3A + 2\begin{bmatrix} 2 & 2 \\ -1 & 6 \\ 4 & 0 \end{bmatrix}\right)^T$.

1.3.4 Definition

A matrix A is *conformable* with a matrix and B if the dimension of A is $n \times k$ and the dimension of B is $k \times m$ for some $k \in \mathbb{N}$.

Remarks

1. In LC, only a notion of multiplication between conformable matrices is considered. In this case the product of an $n \times k$ matrix and a $k \times m$ matrix is a $n \times m$ matrix.
2. This means that a matrix A may be multiplied by a matrix B to form the product AB if and only if the number of columns in A is equal to the number of rows in B.
3. Note also that if A is conformable with B it does not follow that B is conformable with A. For example, a 2×3 matrix be multiplied by a 3×4 matrix to produce a 2×4 matrix but a 3×4 matrix may not be multiplied by a 2×3 matrix
4. Two square matrices of equal dimension may be multiplied together to produce another square matrix of the same dimension. However note that the order of multiplication is important. It will be seen in general that for square matrices A and B;

$$AB \neq BA \tag{1.2}$$

   That is the axiom of commutivity for real numbers $xy = yx$, $\forall\, x, y \in \mathbb{R}$; fails in general for an algebra of matrices.

1.3.5 Definition

Let $A := [A]_{ij} = a_{ij}$ of dimension $n \times r$; and $B := [B]_{ij} = b_{ij}$ of dimension $r \times m$. Then the matrix product $AB = C = [C]_{ij}$ has matrix entries

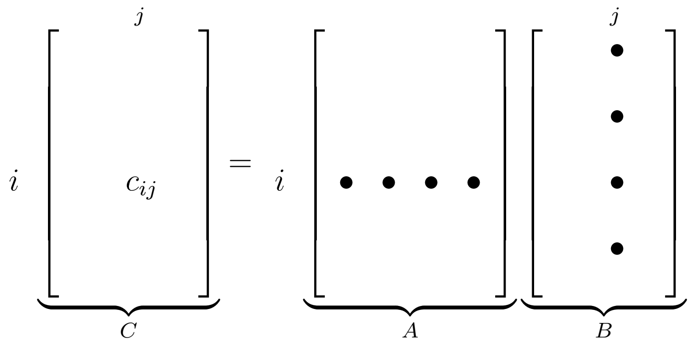
Remarks

This is the technical definition for any two conformable matrices A and B. The meaning of (**??**) will be discussed for the general case of two conformable matrices; and for the cases of $n \times m$ matrices with $n, m \leq 2$.

(i) The General Case;
Let A be a $n \times r$ matrix and B be a $r \times m$ matrix. What are the entries of $C = AB$? Well take the general entry that is in the i-th row and j-th column of C. This is the number c_{ij}. This is by (**??**):

$$c_{ij} = a_{i1}b_{1j} + a_{i2}b_{2j} + a_{i3}b_{3j} + \cdots + a_{ir}b_{rj}$$

So to find the (ij)-th element sum the numbers along the i-th row of A multiplied by the numbers along the j-th column of B:



(ii) A 1×2 matrix by a 2×1 matrix.

Note a 1×2 matrix by a 2×1 matrix is a 1×1 matrix. This is equivalent to a real number; in this case $ac + bd$.

(iii) A 1×2 matrix by a 2×2 matrix.

Note a 1×2 matrix by a 2×2 matrix is a 1×2 matrix.

(iv) A 2×2 matrix by a 2×1 matrix.

Note a 2×2 matrix by a 2×1 matrix is a 2×1 matrix.

(v) A 2×2 matrix by a 2×2 matrix.

I find the best way to remember is as follows:

$$C = AB = \begin{pmatrix} \text{1st row by 1st column} & \text{1st row by 2nd column} & \cdots & \text{1st row by last column} \\ \text{2nd row by 1st column} & \text{2nd row by 2nd column} & \cdots & \text{2nd row by last column} \\ & & & \\ \text{last row by 1st column} & \text{last row by 2nd column} & \cdots & \text{last row by last column} \end{pmatrix} \tag{1.3}$$

Example

If

$$A = \begin{bmatrix} 1 & 8 \\ 3 & -2 \\ 0 & 4 \end{bmatrix} \text{ and } B = \begin{bmatrix} 5 & 9 \\ -2 & 7 \end{bmatrix},$$

then

Other properties of matrix multiplication include:

$$A(BC) = (AB)C; \quad A(B + C) = AB + AC; \quad (A + B)C = AC + BC$$

$$k(AB) = (kA)B; \quad (AB)^T = B^T A^T.$$

Summer 2011 Question 1(b)

Given the matrices

$$A = \begin{bmatrix} 5 & -3 \\ -2 & -4 \\ 2 & 0 \end{bmatrix}, \; B = \begin{bmatrix} 1 & -2 & 3 \\ -3 & 0 & 2 \end{bmatrix},$$

determine the following sums/products if defined

1. $2A + B$
2. $2A + B^T$
3. BA

Solution:

Exercises

1. Let $A = \begin{bmatrix} 2 & 1 \\ 0 & -1 \end{bmatrix}$, $B = \begin{bmatrix} 3 & -1 & 2 \\ 0 & 1 & 4 \end{bmatrix}$, $C = \begin{bmatrix} 3 & -1 \\ 2 & 0 \end{bmatrix}$ and $D = \begin{bmatrix} 1 & 3 \\ -1 & 0 \\ 1 & 4 \end{bmatrix}$. Compute the following (where possible):

$$(i)\, 3A - 2B \;\; (ii)\, 5C \;\; (iii)\, 4A^T - 3C \;\; (iv)\, B + D \;\; (v)\, (A + C)^T \;\; (vi)\, A - D.$$

2. Find A if

   (a) $5A - \begin{bmatrix} 1 & 0 \\ 2 & 3 \end{bmatrix} = 3A - \begin{bmatrix} 5 & 2 \\ 6 & 1 \end{bmatrix}$.

   (b) $3A + \begin{bmatrix} 2 \\ 1 \end{bmatrix} = 5A - 2\begin{bmatrix} 3 \\ 0 \end{bmatrix}$.

   (c) $\left(3A^T + 2\begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix}\right) = \begin{bmatrix} 8 & 0 \\ 3 & 1 \end{bmatrix}$.

   (d) $\left(2A^T - 5\begin{bmatrix} 1 & 0 \\ -1 & 2 \end{bmatrix}\right)^T = 4A - 9\begin{bmatrix} 1 & 1 \\ -1 & 0 \end{bmatrix}$.

3. Compute the following matrix products (if possible):

   (a) $\begin{bmatrix} 1 & -1 & 2 \\ 2 & 0 & 4 \end{bmatrix}\begin{bmatrix} 2 & 3 & 1 \\ 1 & 9 & 7 \\ -1 & 0 & 2 \end{bmatrix}$.

   (b) $\begin{bmatrix} 1 & 3 & -3 \end{bmatrix}\begin{bmatrix} 3 & 0 \\ -2 & 1 \\ 0 & 6 \end{bmatrix}$.

   (c) $\begin{bmatrix} 3 & 1 \\ 5 & 2 \end{bmatrix}\begin{bmatrix} 2 & -1 \\ -5 & 3 \end{bmatrix}$.

   (d) $\begin{bmatrix} a & 0 & 0 \\ 0 & b & 0 \\ 0 & 0 & c \end{bmatrix}\begin{bmatrix} a' & 0 & 0 \\ 0 & b' & 0 \\ 0 & 0 & c' \end{bmatrix}$.

   (e) $\begin{bmatrix} 1 & 2 & 4 \\ 0 & 1 & -1 \end{bmatrix}\begin{bmatrix} -1 & 6 \\ 1 & 0 \end{bmatrix}$.

   (f) $\begin{bmatrix} 1 & 2 & 4 \\ 0 & 1 & -1 \end{bmatrix}\begin{bmatrix} 2 & 0 \\ -1 & 1 \\ 1 & 2 \end{bmatrix}$.

4. Let A, B and C be matrices.

   (a) If A^2 can be formed, what can be said about the size of A.

   (b) If AB and BA can both be formed, describe the sizes of A and B.

   (c) If ABC can be formed, A is 3×3 and C is 5×5, what size is B.

1.4 Matrix Inverses

In arithmetic in $\mathbb{R}$, every non-zero number x has a *multiplicative inverse* x^{-1} given by the number $1/x$ with the property:

where '1' is the *multiplicative identity* with the special property that for all $x \in \mathbb{R}$:

There is a special matrix I that is a multiplicative identity for matrix multiplication:

That is I is a matrix such that for any matrix A:

A natural question to ask is given a matrix A; does there exist a matrix A^{-1} such that:

Why might this be relevant for us (i.e. why are we studying matrices at all?).

Summer 2011 Question 1(c)

Use Gauss-Jordan elimination to find A^{-1} where

$$A = \begin{bmatrix} 1 & 0 & 8 \\ 2 & 5 & 3 \\ 1 & 2 & 3 \end{bmatrix}.$$

Solution:

Autumn 2011 Question 1(a)

Determine A^{-1} where $A = \begin{bmatrix} 1 & 1 & -1 \\ -3 & 2 & -1 \\ 3 & -3 & 2 \end{bmatrix}$.

Solution:

1.5 Determinants

1.5.1 Proposition: Properties of Determinants

1.5.2 Cramer's Rule

Autumn 2011 Question 2(a)

Apply ***only*** *the* Gauss-Jordan Method *to solve the system of linear equations*

$$\begin{aligned} -x + y + z - 3, \\ -2x - 3y - z = 2, \\ 2x - 3y - z = 1. \end{aligned}$$

Verify y using Cramer's Rule.
Solution :

1.5.3 Matrix Inverses

Analysis of Solution Space of Linear Equations

Summer 2011 Question 3(b)

Use only deteminants *to determine if the following homogenous system of linear equations has either the* trivial solution *or* non-trivial solutions.

$$\begin{aligned} 2x - 4y - 5z = 0 \\ 3x + y - 4z = 0 \\ x - 6y - z = 0. \end{aligned}$$

Solution:

Chapter Checklist

1. ...

Chapter 2

Statistics

Be able to analyse statistics, which can be used to support or undercut almost any argument.

Marilyn vos Savant

Statistics is the science of collecting, studying, analysing and making judgements based on numerical data. The subject divides broadly into two branches: *descriptive* and *inferential statistics*.

Descriptive statistics involves describing the main features of a collection of data. Descriptive statistics are distinguished from inferential statistics in that descriptive statistics aim to summarize a data set, rather than use the data to learn about the population that the data are thought to represent. Activities include graphing the data (putting a spin on things) and calculating key summary statistics such as the *average* or the *standard deviation*. The aim of descriptive statistics is to summarise the data. We could also include methods of collection of data.

Inferential statistics is the process of drawing conclusions from recorded data. Typically the data is not complete in that there are measurement errors or the data is just a sample from a much larger population. The outcome of statistical inference may be an answer to the question "what should be done next?", where this might be a decision about making further experiments or surveys, or about drawing a conclusion before implementing some organizational or governmental policy. Note that descriptive statistics precedes inferential statistics in the sense that data is necessary for inferential statistics and it is descriptive statistics that provides this.

2.1 Data Analysis

2.1.1 Types of Data

There are many different types of data and it is useful to be aware of this. As a quick example, the heights of the MATH6038 class is numerical and hence ordered data. However, the hometowns of the MATH6038 class is not numerical and not ordered so that these are fundamentally different types of data that require different presentations and summary statistics in order to summarise them.

Nominal Data — is data that cannot be ordered, for example the eye colours of ten children or the martial status of groups of individuals as single, married, widowed or divorced.

Ordinal Data — is data that can be ordered, for example satisfaction levels in a consumer survey: *very happy, happy, indifferent, unhappy, very unhappy.*

Numerical Data — is exactly what you think it is. Numerical data has the advantage of having a natural ordering. More 'mathematical' methods of data presentation can be used to demonstrate numerical data. An example: the heights of the MATH6038 class.

Whole Number Data — again, exactly what you think it is. Numerical data where the only possibilities are whole numbers; for example, the number of employees in a business.

Continuous Data — data that can take on a infinite number of values that are arbitrarily close to each other; for example, the time taken to serve customers could be anything from zero to an infinite number of seconds.

Obviously there is a lot of overlap here.

2.1.2 Average

When you hear the word average we often think of the mean. For example, if Ann is 28, Betty is 31 and Carolone is 31 we say that their average age is 30. However this is not the only 'average', To be more careful, an average is any 'measure of central tendency' — or if you will the 'middle value'. We might want to talk about the average of nominal data: for example who does the average Irish male support in soccer?

Mean

The mean is the 'usual' average that we usually use. It is used for numerical data only. It is calculated by adding up the all the data points and dividing by the number of data points:
Definition
Let $x = \{x_1, \dots, x_n\}$ be a collection of data. The *mean-average* of x, $\bar{x}$, is given by

$$\begin{aligned}\bar{x} &= \frac{x_1 + \cdots + x_n}{n} \\ &= \frac{\sum_{i=1}^{n} x_i}{n}.\end{aligned}$$

Mode

The mode(s) of a data set is (are) those data points which occur most often and is often the best average to use for nominal data. As an example which English soccer team do you support:

Median

The median is an average that is used for ordinal data. It is primarily used for a finite set of discrete data and when the number of data points is odd, it is that data point which divides the data into a 'bottom half' and 'lower half'.

Example
Find the median of the following data:

$$15, 4, 46, 23, 57, 3, 5, 34, 57, 243, 5.$$

Solution: First we order the data:

Hence 23 is the median.

When the number of data points is even, the median is given as the midpoint of the two 'middle' elements.

Example
Find the median of

$$2, 344, 23, 555, 643, 2, 542, 57$$

Solution

Which Average to Use?

By and large you have a choice but hard and fast rules are:

1. Use the mean unless there are large *outliers* in the data set
2. If there are large outliers in the data set use the median.

When the data is symmetric about it's mode all three will agree.

Example: Average Industrial Wage
The distribution of income looks like

A few millionaires will skew the mean to the right — so that more than half the population don't then make "the average industrial wage". So depending on what your vested interest is, you may use the mean or the median (or the mode). This is the origin of famous quip "*There are only three kinds of lies. Lies, damned lies and statistics*". If you are disingenuous with statistics you can support a lot of marginal arguments.

2.1.3 Deviation

Note that data can be spread out or quite concentrated (around the average). For example, consider soccer players vs rugby players:

The use the *standard deviation* to measure the spread of the data. Standard Deviation is a kind of average deviation/distance from the mean and is defined as follows.

Definition

Let $S = \{x_1, \dots, x_n\}$ be numerical data with a mean-average of $\bar{x}$. Then the *standard deviation*, σ, is given by:

$$\sigma = \sqrt{\frac{\sum_{i=1}^{n}(x-\bar{x})^2}{n}} \tag{2.1}$$

2.2 Data Presentation

There are many, many different data presentation styles: some of which come in and go out of fashion fairly quickly. Here we 'present' some of the classics.

2.2.1 Frequency Tables

Although usually used for numerical data, frequency tables, as generalisations of bar charts, lend themselves well to nominal data with not too many possible outcomes. Frequency tables arise from data that can be put into a number of frequency classes or *bins*. Usually the data will take the form of a list $x = \{x_1, \dots, x_n\}$ and we simply count the number of occurrences in the various bins. Next we calculate their relative frequency with respect to the number of data points.

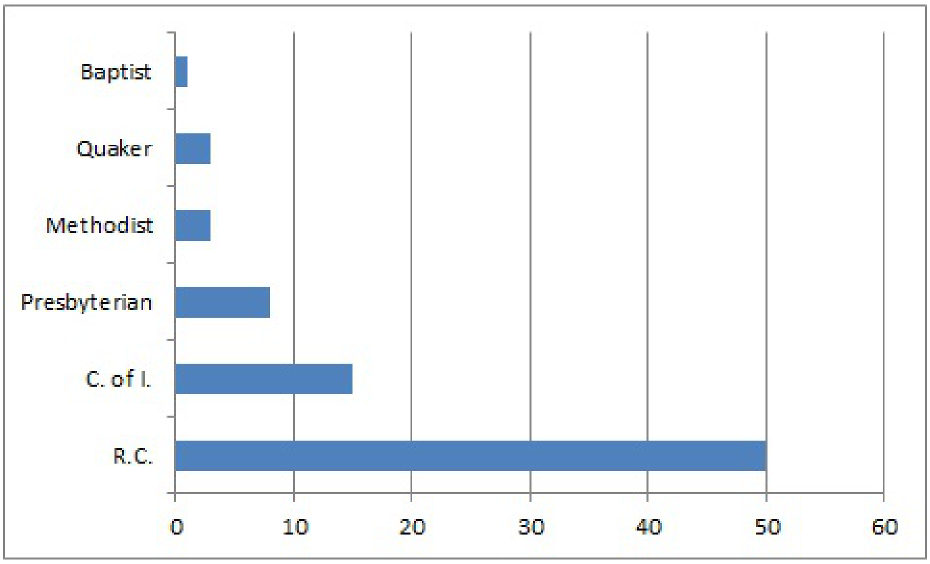
Example

Suppose that in a survey of a small village in Co. Down that of the 80 inhabitants we had 3 Quakers, 1 Baptist, 15 Church of Ireland, 9 Presbyterian, 50 Roman Catholic and 3

Methodist. We could summarise this data in a simple table (it is sometimes a good idea to order nominal data when we have the frequency distribution):

Religion	Frequency
Roman Catholic	50
Church of Ireland	15
Presbyterian	8
Methodist	3
Quaker	3
Baptist	1
Total	80

We could now summarise this data using a bar chart:



We don't have to graph the actual values if we want — instead we can graph their proportions of the total (their frequency).

Example

Suppose that forty individuals presented themselves to give blood in CUH yesterday and their blood types were given as follows:

Blood Type	Frequency	Relative Frequency
O	16	
A	18	
B	4	
AB	2	

The Mean for Grouped Data

Suppose that we want to find the mean number of aphids found on a particular plants in a particular garden. To estimate this we could take a random leaf from 100 plants and count the number of aphids on it.

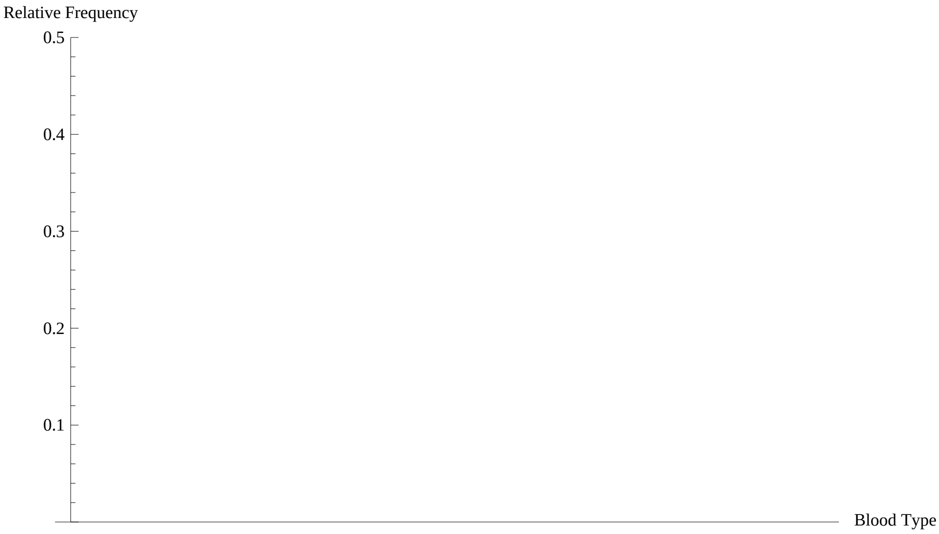


Figure 2.1: A frequency graph for the Blood Type data.

# Aphids	0	1	2	3	4	5
Frequency	10	36	27	16	7	4

If we want to calculate the mean of this we might note that there are 10 instances of 0, 36 instance of 1, etc. and 100 data points in total:

$$\bar{x} = \frac{(0+\cdots+0)+(1+\cdots+1)+(2+\cdots+2)+(3+\cdots+3)+(4+\cdots+4)+(5+\cdots+5)}{100}$$
$$= \frac{10(0)+36(1)+27(2)+16(3)+7(4)+4(5)}{10+36+27+16+7+4} = 1.86.$$

Note that this is noting but the following:

Formula

Suppose we have a discrete data set S in the form of a frequency distribution with n frequency classes with values $x_1, x_2, \ldots, x_n$. Suppose further that there are f_1 points in the frequency class x_1, f_2 points in the frequency class x_2, etc. Then the mean average of the data is given by

$$\bar{x} = \frac{\sum_{i=1}^{n} f_i x_i}{\sum_{i=1}^{n} f_i}. \tag{2.2}$$

Example

On average, which has more children per family: region A or region B?

# Children	0	1	2	3	4	5	6	7	8	9	10
Frequency in A	18	19	35	15	9	2	1	1	0	0	0
Frequency in B	8	11	13	15	26	19	3	2	0	1	2

We can calculate this as follows

When we actually have numerical or ordinal data we don't order the data from most common to least common because the data comes with a natural order. Usually in this case we would plot the frequencies rather than the actual values.

Example

Consider the following data on the amount of black pigmentation in sunfish:

Amount of Black Pigmentation	Number of Fish	Relative Frequency %
No Black Pigmentations	71	18
Faintly Speckled	169	42
Moderately Speckled	84	21
Heavily Speckled	20	
Solid Black Pigmentation	56	14
Total	400	100

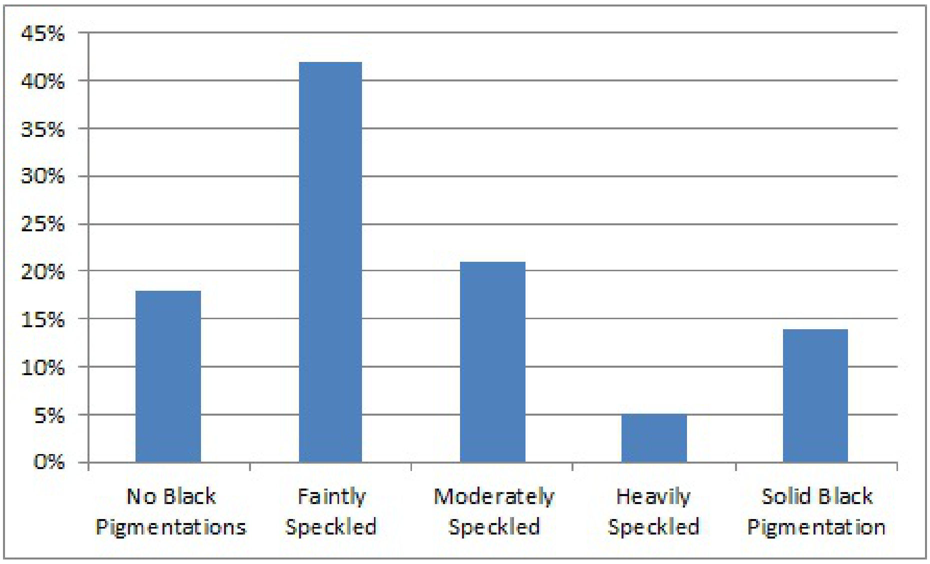


Figure 2.2: A frequency graph for the Sunfish data.

What if we get numerical data with many distinct values? An example would be the (annual) incomes of say 100 people. Usually there are very few repetitions of each data-value in this type of data set (especially continuous data). We illustrate the process with quite a small data-set. The data is of the time taken by each of 50 students to solve a puzzle: {60,33,85,52,65,77,84,65,57,74,71,81,35,50,35,64,74,47,68,54,80,41,61,91,55,73,59,53,45,77, 41,78,55,48,69,85,67,39,76,60,94,66,98,66,73,42,65,94,89,88}.

The steps for setting up a frequency table for such data is as follows:

1. Decide on the **number of frequency classes** to be used (typically from 5 to 20). The fewer the number of classes, the less detail there will be in the frequency graph. Let us hope to use 7.

2. Find the range the data: 33 to 98 — and divide this range by the required number of frequency classes to get the width of each frequency class/ bin.

   Here it makes sense to use 10 as the width of each frequency class. Then choose **convenient** classes of that width. In this case we could use 30 - 40, 40 - 50, etc.

3. Now that the frequency classes have been chosen, we proceed to **count** the number of observations falling into a class and summarising in a frequency table:

Frequency Class	Frequency	Relative Frequency %
30 - 40	4	
40 - 50	6	
50 - 60	8	
60 - 70	12	
70 - 80	9	
80 - 90	7	
90 - 100	4	
Total		100

The convention is that the frequency class 90 - 100 contains the 90 while 80 - 90 does not. Again this data may be summarised in a graph of the (grouped) frequency distribution:

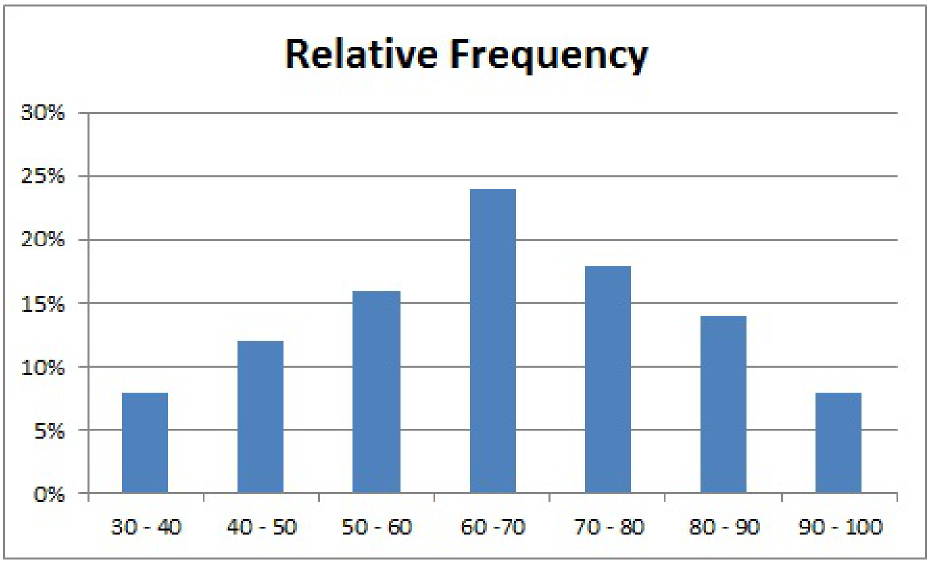


Figure 2.3: A graph of the frequency distribution for the puzzle data.

In this module we are going to make the *Assumption of Uniformity*. This assumption states that the 'sub'-distribution in a frequency class is the uniform distribution — the data points in the frequency class are spread out equally. This is not a very good assumption, particularly when there is a large distance between adjacent frequencies:

However, it is not *too* bad and eases our calculations greatly.

Example

Using the *Assumption of Uniformity*, describe the distribution of the 90 - 100 frequency class from the last example.
Solution:

The Mode for Grouped Data

The mode of grouped data is simply the class with the greatest frequency. If we like we can use the midpoint of this class

2.2.2 Histograms

A more widely used (and mathematically useful eventually) method of graphical representation is that of a *histogram*. A histogram is exactly like the graph of a frequency distribution except the data must be numerical and the *area* of the bars is made proportional to the relative frequency of each class:

Usually all of frequency classes will be of the same width and the histogram will be the same as the graph of a frequency distribution — although we put the bars/columns side-by-side.

Example

Draw a histogram for the puzzle data.
Solution:

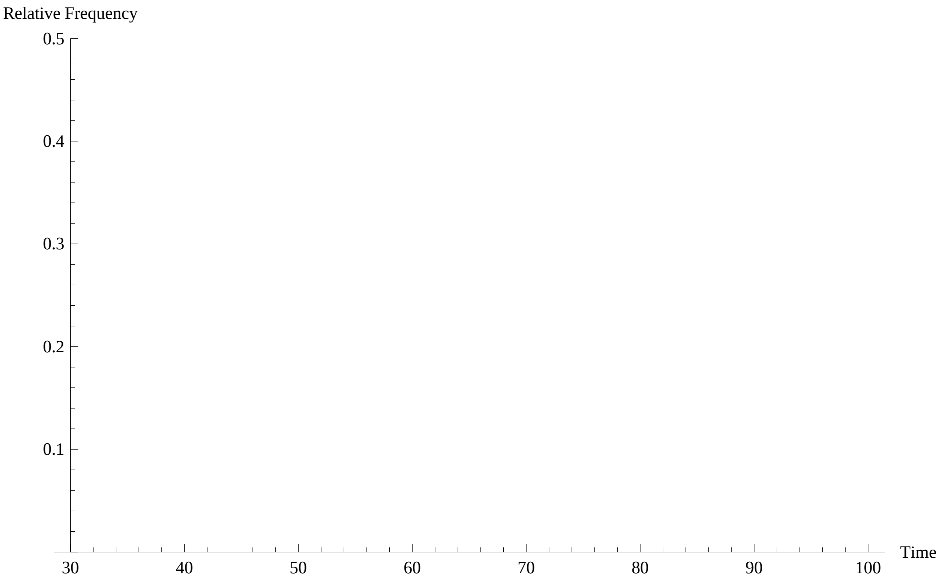


Figure 2.4: A histogram for the puzzle data.

2.2.3 The Assumed Mean Method

Autumn 2011 Question 3 (a)

A sample of 80 ball bearings was taken from the production of machine A and their diameters (in cm) were measured to give the following distribution.

x	0.80-0.82	0.83-0.85	0.86-0.88	0.89-0.91	0.92-0.94	0.95-0.97	0.98-1.00
f	6	8	15	23	18	6	4

1. *Find the* mode *of the above distribution.*
2. *Use the* assumed mean method *to determine the mean and standard deviation.*
3. *A second sample of 100 ball bearings was taken from machine B, giving a mean diameter of 0.73 cm with standard deviation of 0.04 cm. Compute the* coefficient of variation *for each machine. Which machine has the greater variation?*

Solution:

2.2.4 Other Methods of Data Presentation

Include

1. Stem and Leaf Diagrams
2. Pie Charts
3. Dot Plots
4. Line Plots

Chapter 3

Probability

The probable is what usually happens.

Aristotle.

3.1 Introduction

The study of probability provides us with concepts and terminology that may be applied...

Summer 2011 Question 1(a)

A fair coin is tossed five times. Find the probability of obtaining;

1. five heads,
2. at least one head.

Solution:

3.2 Random Variables

The central concept of probability theory is that of a *random variable*: a variable X that takes on different values 'at random' with respective probabilities. Examples of random variables include:

1. the second name of the first baby to be born in 2013
2. the amount of rainfall next Tuesday
3. the age I will be when I die

3.3 Independence

Example

A deck of 52 cards contains 13 cards in each of the spade ($\spadesuit$), heart ($\heartsuit$), diamond ($\diamondsuit$) and club ($\clubsuit$) suits. What is the probability that two cards dealt randomly from the deck will

both be spades?
Solution:

Exercises

1. Suppose that A, B and C are three independent events such that $P(A) = 1/4$, $P(B) = 1/3$ and $P(C) = 1/2$. Evaluate the probabilities of the following events:
   (a) none of these events will occur.
   (b) exactly one of these events will occur.
   (c) A or B; $A \cup B$.

3.4 Conditional Probability

3.5 Tree & Reliability Block Diagrams

Summer 2011 Question 1(d)

A manufacturer has three machines that produce fan heaters. Machine I produces 40%, machine II produces 35% while machine III produces the remaining amount of heaters. Machine I outputs 3% of its run as defective, machine II outputs 2% of its run as defective and machine III has 1% of its output defective. Represent this information in a tree diagram. A heater is found to be defective. Find the probability that this defective heater was produced by machine III, i.e. determine $P(III|D)$. Solution:

Autumn 2011 Question 1(c)

Mobile phones from the CIT shop come in three varieties, Samsung, Nokia *or* phone. *Of all such mobiles, 25% are Samsung, 35%* Nokia *and 40%* phone. *Further it is known that 3% of all* Samsung *mobiles are defective, 1% of* Nokia *are defective, and 2% of* iPhone *are defective.*

1. *Represent this information in a tree diagram.*
2. *Find the probability that a mobile chosen at random is a* Nokia *or an* phone.
3. *Find the probability that a mobile chosen at random is defective.*
4. *A certain mobile is found to be defective in the CIT shop. Find the probability that this defective mobile is a* Nokia.

Solution:

Summer 2011 Question 1(f)

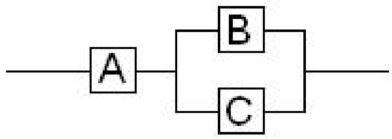
A system S consists of five identical components connected in parallel, each with reliability a.

1. *Express the overall reliability of the system in terms of a.*
2. *Determine a if the overall reliability of the system is* 0.96*.*

Solution:

Summer 2011 Question 3(a)

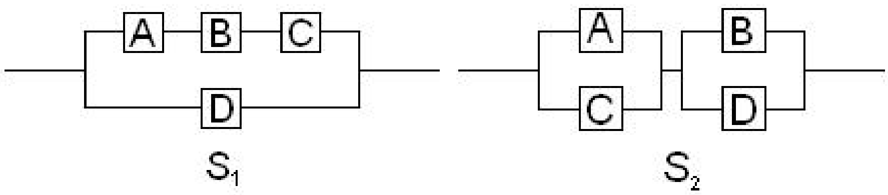
Consider the RBD that is described by the diagram below, where $P(A) = 0.9$, $P(B) = 0.8$ and $P(C) = 0.7$. Compute the system reliability showing all steps and intermediate calculations.



Solution:

Autumn 2011 Question 1(e)

A Reliability Block Diagram *is given for two systems S_1 and S_2. Determine $P(S_1)$ and $P(S_2)$ and hence identify the most reliable system where the reliabilities $P(A) = 0.9$, $P(B) = 0.7$, $P(C) = 0.8$ and $P(D) = 0.9$. Carefully show all steps and intermediate calculations.*



Solution:

3.6 Binomial Distribution

Summer 2011 Question 3(c)

A manufacturer estimates that 5% of his output of a small item is defective. Assuming a binomial distribution find the probability that in a sample of 25 items, more than two items will be defective.
Solution:

Autumn 2011 Question 3(b)

Suppose a large lot contains 8% defective fuses. Assuming a binomial *distribution find the probability that in a sample of 12 fuses, either five or six fuses are defective.*
Solution: Exercises

1. The following probabilities have been computed for the binomial distribution of a random variable X for which $n = 5$, but two of the probabilities have been omitted. Furthermore, the value of p has not been provided. Determine the missing probabilities, and explain your procedure.

X	0	1	2	3	4	5
Probability		0.4096	0.2048	0.0512	0.0064	

3.7 Poisson Distribution

Summer 2011 Question 2(b)

A production department has 35 similar milling machines. The number of breakdowns on each machine averages 0.06 per week. Determine the probabilities of having

1. *one machine breaking down in a week,*
2. *less than three machines breaking down in a week.*

Solution:

Autumn 2011 Question 1(d)

The number of vehicle failures satisfies a Poisson *distribution. Over the year the number of failures in a fleet of vehicles for each month is given by the following table.*

J	F	M	A	M	J	J	A	S	O	N	D
3	1	0	1	0	1	2	3	1	0	1	2

1. *What is the probability of two failures in a given month?*
2. *What is the probability of at most 2 failures in a given month?*

Solution:

3.8 Normal Distribution

Many examples of continuous data is unimodal (peaked, bell-shaped) and symmetric (about the peak) and can be shown to be of a certain form. Such data is called *normal* and we say it has a *normal distribution.* By and large, data with a dominant average with deviations from the mean just as likely to be positive or negative tend to have this shape:
Examples of random variables likely to conform to the normal distribution are:

1. A particular experimental measurement subject to several random errors.
2. The time taken to travel to work along a given route.
3. The heights of woman belonging to a certain race.

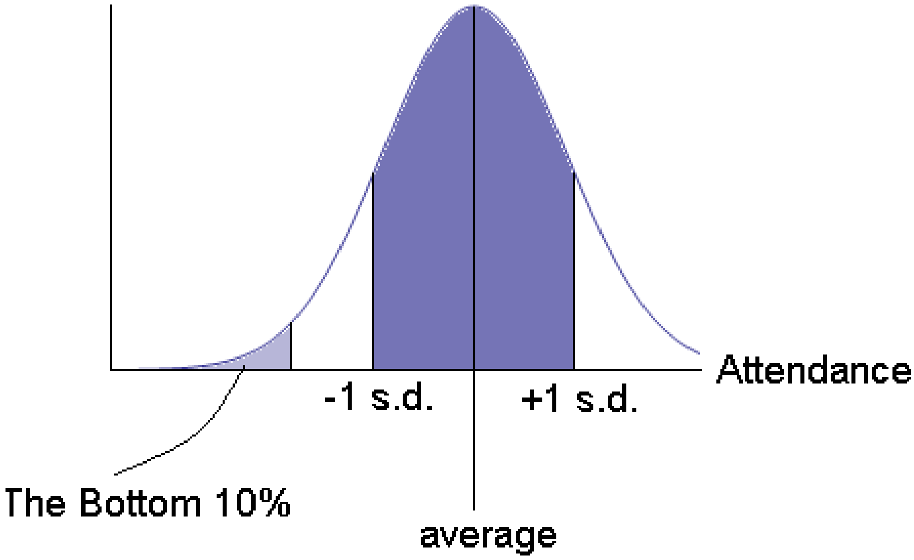


Figure 3.1: Note that for a normal distribution approximately 68% of the data is found within a distance of one standard deviation from the average.

3.8.1 The z-Distribution

In practise, the normal distribution is very useful in that real-life calculations are very easy to handle because all normal distributions are related to each other in that all of the are rescaling of a particular normal distribution — the Daddy Distribution if you will. This is the normal distribution with mean $\mu = 0$ and standard deviation $\sigma = 1$:
This ease comes from the fact that we can transform from $X = N[\mu, \sigma]$ to $z = N[0, 1]$ by the following *z-transform*

$$z = \frac{X - \mu}{\sigma} \tag{3.1}$$

This is a transform that converts the X-distribution to a z-distribution:

Note that $\mu \to 0$ by (3.1) as you would hope. It is not clear why we are doing this but the following fact makes it all clear:

Fundamental Calculation of Normal Distributions

Suppose that $X = N[\mu, \sigma]$ and we want to calculate the probability

$$\mathbb{P}[x_1 \leq X \leq x_2].$$

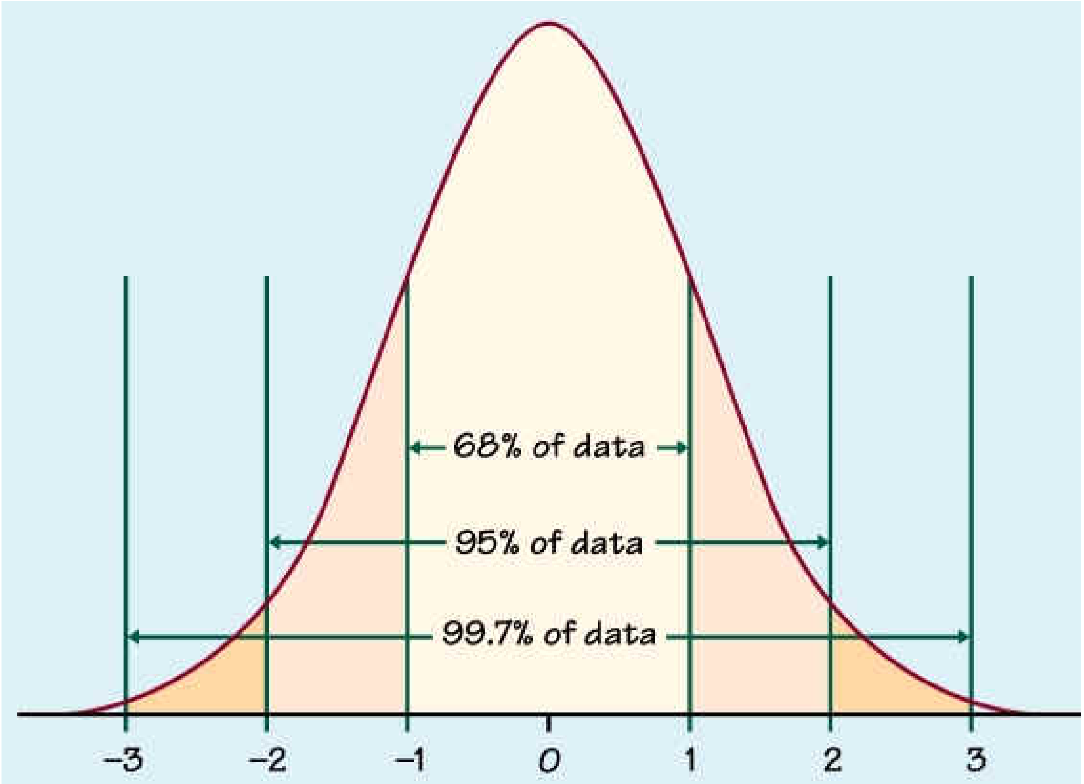


Figure 3.2: All normal distributions are just transforms of $z = N[0, 1]$.

Then we can transform $x_1 \to z_1$ and $x_2 \to z_2$ using (3.1); and in this case

$$\mathbb{P}[x_1 \leq X \leq x_2] = \mathbb{P}[z_1 \leq z \leq z_2] \tag{3.2}$$

Proof. A well-known, albeit difficult, integration •

Now how do we calculate $\mathbb{P}[z_1 \leq z \leq z_2]$? Well we can't right away but what we can do is integrate the frequency distribution of $N[0, 1]$ from $-\infty$ to z_1 to calculate $\mathbb{P}[z \leq z_1]$... too much work, too difficult? Yes absolutely: that is why we use a table of values:

Example

What z value has a 20% probability of being exceeded?
Solution: We are looking for the z_1 such that $\mathbb{P}[z \geq z_1] = 0.2$. Now we don't have the $\mathbb{P}[z \geq z_1]$ probabilities in this table (you have the other one in your department of education tables). Hence we calculate:

Summer 2011 Question 3(d)

Samples of 10 A fuses have a mean fusing current of 9.9 A and a standard deviation of 1.2 A. Assuming the fusing currents are normally distributed, determine the probability of a fuse blowing with a current between 8 A and 12 A.
Solution:

Autumn 2011 Question 1(b)

Wires manufactured for use in a certain electronic device are specified to have resistances between 0.16 Ω *and 0.18* Ω*. The actual measured resistances of the wires have a* normal *distribution with a mean of 0.17* Ω *and a standard deviation of 0.005* Ω*. What is the probability that a randomly selected wire will meet the specifications?*
Solution:

Autumn 2011 Question 2(b)

It is assumed that the weights of goods packed by a certain machine are normally *distributed with a mean weight of 8 kg and a standard deviation of 0.03 kg. Calculate the probability that a package taken at random will weigh*

1. *less than 8.07 kg*
2. *greater than 8.08 kg*
3. *between 7.98 kg and 8.05 kg?*

If 99.8% of readings are less than some critical weight, W*, find the value of* W*.*
Solution:

3.9 Sampling

Consider the problem of finding the mean-average height, μ, of the population of Irish males. Plainly this is impossible. However we could *approximate* this population mean-average by taking a random sample of say 1,000 males from the population. Suppose we measure these males to have heights

$$h_1, h_2, \dots, h_{1,000}.$$

We could then find the mean-average of the sample:

$$\bar{h} = \frac{h_1 + \cdots + h_{1,000}}{h} = \frac{1}{h}\sum_{i=1}^{1,000} h_i,$$

Now we could take $\bar{h}$ as an estimate of μ; $\bar{x} \approx \mu$. How accurate is this? Now consider all the possible samples we could have taken from the population:
So in this sense, the mean-average of the sample means is equal to the population mean. We can also show that the standard deviation of the sample means from the population mean is given by $\sqrt{\sigma}/\sqrt{1,000}$, where σ is the standard deviation of the population.

This general process is called *sampling* and is used to make inferences of a whole *population* just by looking at a *random sample* of the data.

3.9.1 The Central Limit Theorem

The *Central Limit Theorem* (CLT) is of crucial importance in sampling theory (as well as probability theory in general). It essentially says that in many cases, a large collection of independent identically distributed random variables has a normal distribution approximately. In particular this is true for the Binomial distribution:

$$\lim_{n\to\infty} B[n,p] \to N[np, \sqrt{np(1-p)}]. \tag{3.3}$$

Example

A random variable has binonomial distribution with $n = 200$ and $p = 0.2$. Use the Normal approximation $X = B[200, 0.2] \approx N[40, \sqrt{32}]$ to estimate the probability $\mathbb{P}[30 < X < 45]$.
Solution: First we do a z-transforms of $X = 30$ and $X = 45$:

Hence we must calculate $\mathbb{P}[-1.77 < z < 0.88]$:

As before, the best way to calculate this is to calculate

$$\mathbb{P}[z \leq 0.88] - 0.5 =$$
$$\mathbb{P}[z \leq 1.77] - 0.5 =$$

Much more importantly however is the rôle of CLT in Sampling Theory: it underpins the whole justification.

Hence we can be reasonably confident of using sample summary statistics to estimate population summary statistics. Hence we can make informed and statistically significant decisions (recall the example in the motivation of how often to close down the production line).

Example

The times taken to travel from your home to class on $n = 7$ different days are:

$$30, 33, 26, 23, 30, 35, 27 \text{ mins}$$

If you are willing to take a 1% chance of being late for class, how long before the class should you set out?
Solution:

3.10 Quality Control Charts

Summer 2011 Question 4

In order to monitor the quality of a production run of aluminium bolts, 8 samples, each containing 4 components, are taken at random and their diameter lengths are measured correct to the nearest 0.1 mm and tabulated as follows:

Sample	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**
	89.4	92.2	89.7	89.2	91.1	91.7	91.8	93.2
	89.9	90.1	90.1	89.4	91.0	89.9	91.8	90.1
	91.9	91.3	92.3	90.8	92.1	89.3	90.3	87.3
	90.8	91.4	90.9	89.8	91.3	80.2	91.9	89.3
Means, $\bar{x}_i$	90.50	91.25	$\bar{x}_3$	89.8	$\bar{x}_5$	$\bar{x}_6$	$\bar{x}_7$	$\bar{x}_8$
Ranges, w_i	2.5	2.1	w_3	1.6	w_5	w_6	w_7	w_8

1. *Use sample 3 to set up 95% and 99% confidence intervals for the population mean. Comment briefly on your answers.*

2. *Calculate the remaining sample means $\bar{x}_i$ and ranges w_i. Find the grand mean $\overline{\overline{x}}$ and the associated* outer *and* inner control limits. *Hence set up a control chart for the sample means. State, giving reasons, whether or not the process is under control.*

Solution:

Autumn 2011 Question 4

In order to monitor the quality of a production run of aluminium bolts, 8 samples, each containing 4 components, are taken at random and their diameter lengths are measured correct to the nearest 0.1 mm and tabulated as follows:

Sample	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**
	88.3	91.1	87.6	89.2	81.1	91.7	81.8	90.2
	91.0	91.2	85.2	89.4	82.0	89.9	81.8	90.2
	91.9	89.3	82.3	90.8	82.1	88.3	90.3	87.3
	90.8	93.4	92.9	86.8	81.3	80.2	81.9	89.3
Means, $\bar{x}_i$	$\bar{x}_1$	$\bar{x}_2$	$\bar{x}_3$	$\bar{x}_4$	81.65	87.5	83.95	89.25
Ranges, w_i	w_1	w_2	w_3	w_4	0.9	11.5	8.5	2.9

1. *Use sample 3 to set up 95% and 99% confidence intervals for the population mean. Comment briefly on your answers.*

2. *Calculate the remaining sample means* $\bar{x}_i$ *and ranges* w_i*. Find the grand mean* $\bar{\bar{x}}$ *and the associated* outer *and* inner control limits. *Hence set up a control chart for the sample means. State, giving reasons, whether or not the process is under control.*

Solution:

3.11 Bayesian Statistics

Summer 2011 Question 1(e)(ii)

A technical supervisor picks a sample of 20 light bulbs at random from a shipment of light bulbs known to contain 10% defective light bulbs. What is the probability that no more than two of the light bulbs are defective?

Remark

Solution:

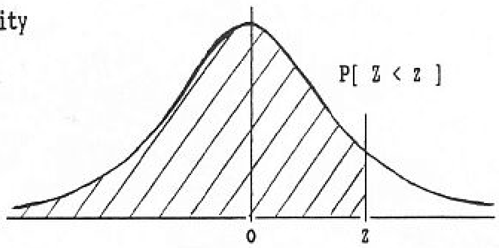
STANDARD STATISTICAL TABLES

1. Areas under the Normal Distribution

The table gives the cumulative probability up to the standardised normal value z

i.e.

$$P[\ Z < z\ ] = \int_{-\infty}^{z} \frac{1}{\sqrt{2\pi}} \exp(-\tfrac{1}{2}Z^2)\ dZ$$



z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.5000	0.5040	0.5080	0.5120	0.5159	0.5199	0.5239	0.5279	0.5319	0.5359
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480	0.6517
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224
0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517	0.7549
0.7	0.7580	0.7611	0.7642	0.7673	0.7704	0.7734	0.7764	0.7794	0.7823	0.7854
0.8	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.8078	0.8106	0.8133
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.8340	0.8365	0.8389
1.0	0.8413	0.8438	0.8461	0.8485	0.8508	0.8531	0.8554	0.8577	0.8599	0.8621
1.1	0.8643	0.8665	0.8686	0.8708	0.8729	0.8749	0.8770	0.8790	0.8804	0.8830
1.2	0.8849	0.8869	0.8888	0.8907	0.8925	0.8944	0.8962	0.8980	0.8997	0.9015
1.3	0.9032	0.9049	0.9066	0.9082	0.9099	0.9115	0.9131	0.9147	0.9162	0.9177
1.4	0.9192	0.9207	0.9222	0.9236	0.9251	0.9265	0.9279	0.9292	0.9306	0.9319
1.5	0.9332	0.9345	0.9357	0.9370	0.9382	0.9394	0.9406	0.9418	0.9429	0.9441
1.6	0.9452	0.9463	0.9474	0.9484	0.9495	0.9505	0.9515	0.9525	0.9535	0.9545
1.7	0.9554	0.9564	0.9573	0.9582	0.9591	0.9599	0.9608	0.9616	0.9625	0.9633
1.8	0.9641	0.9649	0.9656	0.9664	0.9671	0.9678	0.9686	0.9693	0.9699	0.9706
1.9	0.9713	0.9719	0.9726	0.9732	0.9738	0.9744	0.9750	0.9756	0.9761	0.9767
2.0	0.9773	0.9778	0.9783	0.9788	0.9793	0.9798	0.9803	0.9808	0.9812	0.9817
2.1	0.9821	0.9826	0.9830	0.9834	0.9838	0.9842	0.9846	0.9850	0.9854	0.9857
2.2	0.9861	0.9865	0.9868	0.9871	0.9874	0.9878	0.9881	0.9884	0.9887	0.9890
2.3	0.9893	0.9896	0.9898	0.9901	0.9904	0.9906	0.9909	0.9911	0.9913	0.9916
2.4	0.9918	0.9920	0.9922	0.9924	0.9927	0.9929	0.9931	0.9932	0.9934	0.9936
2.5	0.9938	0.9940	0.9941	0.9943	0.9945	0.9946	0.9948	0.9949	0.9951	0.9952
2.6	0.9953	0.9955	0.9956	0.9957	0.9959	0.9960	0.9961	0.9962	0.9963	0.9964
2.7	0.9965	0.9966	0.9967	0.9968	0.9969	0.9970	0.9971	0.9972	0.9973	0.9974
2.8	0.9974	0.9975	0.9976	0.9977	0.9977	0.9978	0.9979	0.9980	0.9980	0.9981
2.9	0.9981	0.9982	0.9982	0.9983	0.9984	0.9984	0.9985	0.9985	0.9986	0.9986

z	3.00	3.10	3.20	3.30	3.40	3.50	3.60	3.70	3.80	3.90
P	0.9986	0.9990	0.9993	0.9995	0.9997	0.9998	0.9998	0.9999	0.9999	1.0000

Figure 3.3: This tables gives the probabilities $\mathbb{P}[z \leq z_1]$ for $z_1 \geq 0$. We will have to use the symmetry of $N[0,1]$ to find more exotic probabilities such as $\mathbb{P}[z_1 \leq z \leq z_2]$.

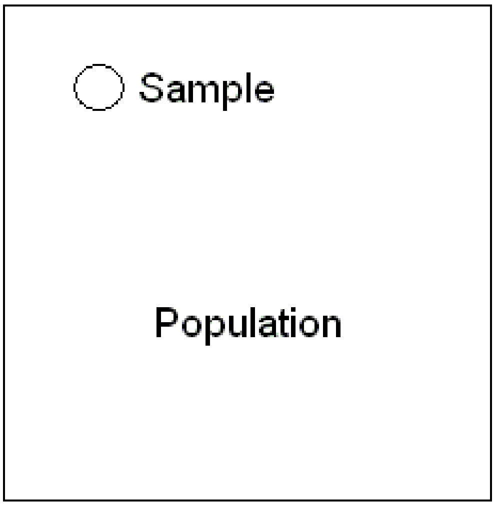


Figure 3.4: There are about 10^{3733} possible samples of 1,000 males from the population of about 2,000. If we look at the sample mean-average as a random variable, then the mean-average of the sample mean-averages is equal to the population mean-average.