MATH7019 — Technological Maths 311

J.P. McCarthy

September 23, 2013

Contents

0.1 Introduction

Lecturer

J.P. McCarthy

Office

Meetings before class by appointment via email only.

Email & Web:

jpmccarthymaths@gmail.com and *http://jpmccarthymaths.com*
This page will comprise the webpage for this module and as such shall be the venue for course announcements including definitive dates for the tests. 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 MATH7019; only those in the category 'MATH7019'. Feel free to use the comment function therein as a point of contact.

Module Objective

This module covers: Taylor series in one and two variables; first and second order differential equations with constant coefficients; probability distributions, statistical inference and control charts.

Module Content

Further Calculus

Maclaurin and Taylor series of a function of a single variable. Review of partial differentiation. Taylor series expansions of functions of two variables. Differentials.

Differential Equations

Review of the solution of first order differential equations using the method of Separable Variables. Euler's method and the Three Term Taylor method for obtaining numerical solutions to first order differential equations. Solution of second order differential equations using the method of Separable Variables and the method of Undetermined Coefficients. Step functions. Solution of differential equations to include those that occur in the theory of beams and beam struts.

Probability

Laws of probability. Probability distributions such as the Binomial, Poisson Distribution and Normal distributions. Applications of these distributions to engineering problems.

Sampling Theory

Sampling from a Normal population. Confidence intervals for the population mean. Hypothesis tests for population means using the z-test and the t-test.

Quality Control

Control charts for sample means and sample ranges. Process capability.

Assessment

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

Continuous Assessment

The Continuous Assessment will be divided between two in-class tests, each worth 15%, in weeks 6 & 11.
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.

Tutorials

The aim of the tutorials will be to help you achieve your best performance in the tests and exam.

Exercises

There are many ways to learn maths. Two methods which arent 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.
There are exercises in the notes for your consumption. 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 MATH7019, they will not detail all there is to know. Further references are to be found in the library. Good references include:

- Glyn James et al, *Advanced Modern Engineering Mathematics*, 3rd Ed. [ISBN: 978-0130454256]
- Douglas C. Montgomery, George C., Runger 2007, *Applied statistics and probability for engineers*, Fourth Ed., John Wiley & Sons Hoboken, NJ.
- J. Bird, 2006, *Higher Engineering Mathematics*, Fifth Ed., Newnes.

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: When is an Approximation Good Enough?

> *Although this may seem a paradox, all exact science is dominated by the idea of approximation.*
>
> Bertrand Russell

Figure 1: A good door closer should close automatically, close in a gentle manner and close as fast as possible.

One possible design would be to put a mass on the door and attach a spring to it (just for ease of explanation we'll only worry about one dimension).

Assuming that the door is swinging freely the only force closing the door is the force of the spring. Now *Hooke's Law* states that the force of a spring is directly proportional to its distance from the equilibrium position. If the door is designed so that the equilibrium position of the spring corresponds to when the door is closed flush, then if $x(t)$ is the position of the door t seconds after release, then the force of the spring at time t is given by:

where $k \in \mathbb{R}$ is known as the spring constant. It can be shown that this system *does* close the door automatically but the balance between closing the door gently and closing the door quickly is lost. Indeed if the door is released from rest at $t = 0$, then the speed of the door will have the following behaviour:

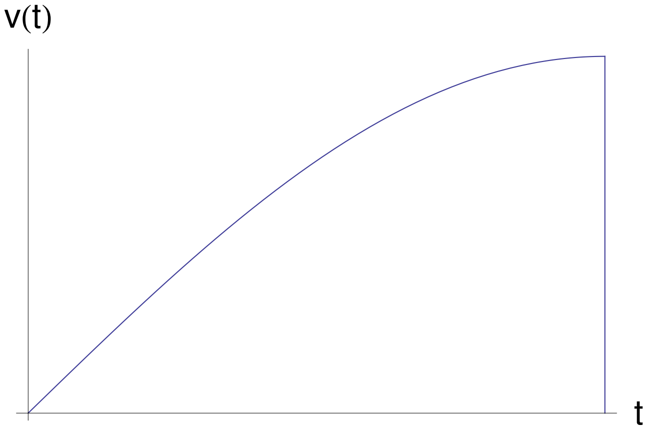


Figure 2: With a spring system alone, the door will quickly pick up speed and slam into the door-frame at maximum speed.

Clearly we need to slow down the door as it approaches the door-frame. A simple model uses a *hydraulic damper*:

Figure 3: A hydraulic damper increases its resistance to motion in direct proportion to speed.

With the force due to the hydraulic damper proportional to speed, the force of the hydraulic damper at time t will be:

for some $\lambda \in \mathbb{R}$. Now by Newton's Second Law:

and the fact that speed is the first derivative of distance, and in turn acceleration is the first derivative of speed, means that the *equation of motion* is given by:

It can be shown that a suitably chosen k and λ will provide us with a system that closes automatically, closes in a gentle manner and closes as fast as possible.

Equations of this form turn up in many branches of physics and engineering. For example, the oscillations of an electric circuit containing an inductance L, resistance R and capacitance C in series are described by

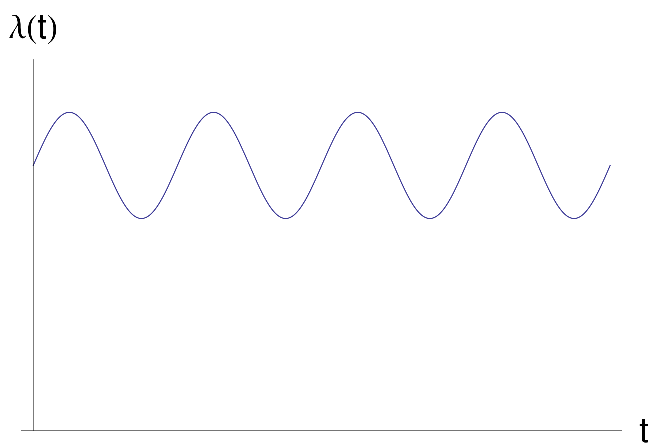
$$L\frac{d^2q}{dt^2} + R\frac{dq}{dt} + \frac{1}{C}q = 0, \tag{1}$$

in which the variable $q(t)$ represents the charge on one plate of the capacitor. These class of equations, *linear differential equations*,

may be solved in various different ways. In this module we will explore one such method — that of *the method of undetermined coefficients.*

Figure 4: Top Gear dropped a VW Beetle from a height of 1 mile and it spun in the air as it fell.

If we are trying to formulate a model for the fall of this car we would have to try and account for the way the roll of the car means that the coefficient of the drag term ($\lambda v(t)$) varies between its maximum and minimum in a wave-like way:



A function with this behaviour is:

$$\lambda(t) = \frac{1}{2}(M + m) + \frac{1}{2}(M - m)\sin\omega t \tag{2}$$

where M and m are the maximum and minimum of $\lambda(t)$ and ω is a constant related to the angular frequency. Then the equation of motion is of the form:

Neither the method of undetermined coefficients nor any other straightforward method I know of solves this differential equation.

Unfortunately this is typical, and for many systems for which a differential equation may be drawn, it may be impossible to solve the equations. There are a number of numerical techniques which can give approximate answers. However if we are participating in some industrial project with millions spent on it we don't want to be chancing our arms on any old estimate or guess. *Approximation Theory* aims to control these errors as follows. Suppose we have a Differential Equation with solution $y(x)$. An approximate solution $A_y(x)$ to the equation can be found using some numerical method. If the approximation method is sufficiently 'nice' we may be able to come up with a measure of the error:

Here $|\cdot|$ is some measure of the *distance* between $y(x)$ and $A_y(x)$. The most common measure here would be maximum error:

We would call the parameter ε here the *control* or the *acceptable error*. Some classes of problem are even nicer in that with increasing computational power we can develop a sequence of approximate solutions $\{A_y^1(x), A_y^2(x), A_y^3(x), \dots\}$ with decreasing errors $\{\varepsilon_1, \varepsilon_2, \varepsilon_3, \dots\}$:

Even nicer still from a mathematical point of view if we can find a sequence of approximations with errors decreasing to zero:

In this case we say that the sequence of approximations *converges*.

In this module we will take a first foray into the approximation theory of numerical methods by estimating the solutions of differential equations.

We will not however be measuring *how* accurate our approximate solutions are. However *statistics*, particularly sampling theory, can tell us when our approximations are good enough. For example, suppose we have a business which constructs a machine component. Suppose the company ordering the component wishes to know what 'stress-level' the component can take. Due to natural variations some samples will have a larger tolerance than others — so how can we approach the business and say that our components can take a stress level of S? In practise we can't, but we can make statements along the line of:

> *On average, our components can withstand a stress-level of S.*

However we can't go around testing every single one of the components produced. So what we do is we take a sample of 100 or 1,000 of these components away and have them tested. In this module we will see that we can be 'quite' confident that the average ability to withstand stress of all the components we produce is very well estimated by the sample average. *Sampling Theory* makes precise this idea.

0.3 Review of Calculus

0.3.1 Functions

The primary objects that we study here are *functions*. For the purposes of this review, a function is an object f that takes as an input a real number x and outputs a real number $f(x)$ *that is given in terms of the input*:

$$f : x \mapsto f(x).$$

For example, $f(x) = x^2$ is a function

$$x \mapsto x^2.$$

The functions that we are interested in are combinations of

— constant functions; e.g. $k(x) = 1$

— lines; e.g. $\ell(x) = 3x - 2$

— quadratics; e.g. $q(x) = x^2 + 2x - 3$

— polynomials; e.g. $p(x) = x^4 + x^3 - 2x$

— powers; e.g. $s(x) = \sqrt{x} = x^{1/2}$

— trigonometric; e.g. $\sin x$, $\cos x$ and $\tan x$

— inverse trigonometric; e.g. $\sin^{-1}(x)$, $\tan^{-1}(x)$

— exponential; $f(x) = e^x = \exp(x)$

— logarithmic; $g(x) = \ln x$

By combinations we mean

— sums; e.g. $f(x) = \sin x - 3$

— scalar multiples; e.g. $f(x) = 5\sqrt[3]{x}$

— differences; e.g. $f(x) = x^2 + 23x - 4 - e^x$

— products; e.g. $f(x) = x^3 \tan x$

— quotients; e.g. $f(x) = \frac{\sin^{-1}(x)}{\sqrt{x}}$

— powers; e.g. $f(x) = (x^4 + x^3)^{-3}$

— roots; e.g. $f(x) = \sqrt{\ln x}$

— compositions; e.g. $f(x) = \sin(x^2 + 3x - 2)$

We can visualise all of these objects by looking at their graph. The graph of $f(x)$ is all of the points of the form $(x, f(x))$. More on this in the function catalogue.

0.3.2 The Derivative of a Function

We developed the idea of the *derivative* of a function that would allow to study the following problems:

1. **Tangents** Most of these elementary functions are *smooth*. This means that *locally* (near a point), they are well-approximated by lines:

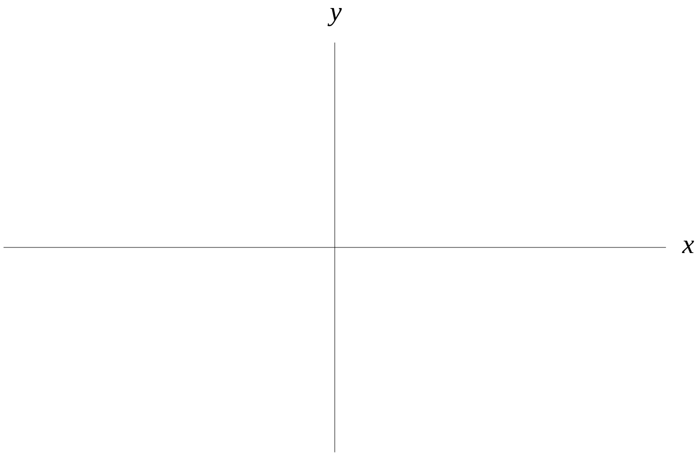


Figure 5: Near the point $x = 1$, the curve $y = \ln x$ is well-approximated by the line $y = x - 1$.

To find the equation of this *tangent* line, which has equation

$$y - y_1 = m(x - x_1)$$

it is necessary to find the slope of the tangent. This led us to the following:

$$\text{slope of tangent to } y = f(x) \text{ at } x = a \ := f'(a)$$

where

$$f'(x) \equiv \frac{dy}{dx} = \lim_{h \to 0} \frac{f(x+h) - f(x)}{h}.$$

This allows us to approximate functions locally by lines; i.e. using their tangents.

2. **Rates of Change** Suppose we have a function that describes the temperature of an object:

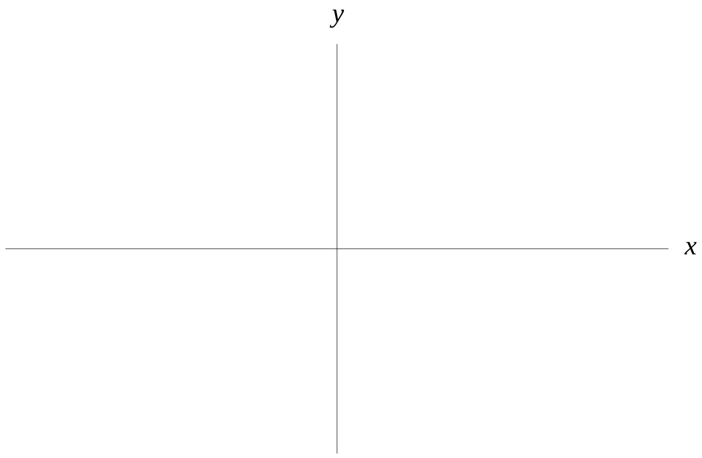


Figure 6: Suppose that we have a formula $y = T(x)$ for the temperature of an object after x seconds. Can we say when the temperature is increasing/decreasing?

It turns out that

$$\text{rate of change of } f(x) \ = f'(x) \equiv \frac{dy}{dx}.$$

and we can analyse the rate of change of a function like this.

3. **Local Maxima/Minima** Suppose we have a function of the form $y = f(x)$. Can we find its (local) maxima and minima? The derivative allows us to do so:

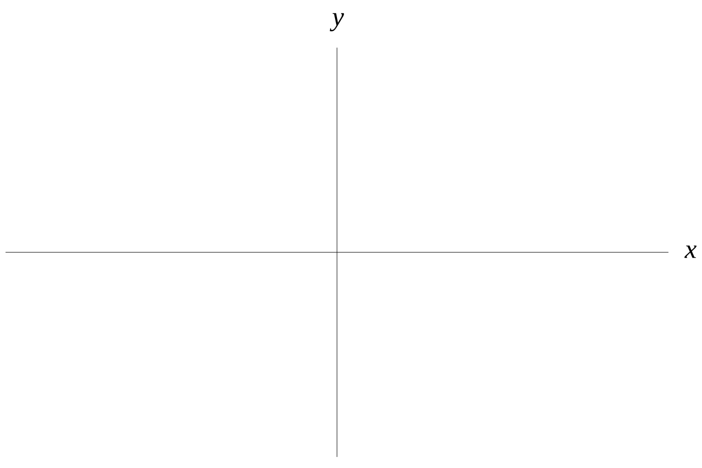


Figure 7: At a turning point, the slope of the tangent to the curve is zero: $f'(x) = 0$. We can use the second derivative to determine whether the turning point is a (local) maximum or minimum.

4. **Differential Equations** Suppose we have a quantity $Q(x)$ which increases in direct proportion to itself so that

$$\begin{aligned}\text{rate of change of } Q \text{ with respect to } x &= kQ \\ \frac{dQ}{dx} &= k \cdot Q(x),\end{aligned}$$

   This is a *differential equation.* Two simple examples include radioactive decay and exponential population growth.

The issue here is that we need to differentiate a function $f(x)$. The derivatives of the functions outlined above are given in the function catalogue. The question is how do we find the derivative of

— sums

— scalar multiples

— differences

— products

— quotients

— powers

— roots

— compositions

0.3.3 "Rules of Differentiation"

The answer is by the Sum, Scalar, Product, Quotient & Chain Rules which need to be well understood to do well in MATH7019. They are not really rules but theorems/facts that describe how we should differentiate sums, products, compositions, etc.

"Rules of Differentiation"

Suppose that $f(x)$ and $g(x)$ are functions, $n \in \mathbb{Q}$ a fraction and $k \in \mathbb{R}$ a real number. Suppose also that $\frac{d}{dx}$ is the *differential operator* (i.e. it means differentiate), and $f'(x)$ and $g'(x)$ are the derivatives of $f(x)$ and $g(x)$ respectively. *Then*

$$\frac{d}{dx}(f(x)+g(x)) = f'(x)+g'(x) \quad \text{[Sum Rule]}$$

$$\frac{d}{dx}(k \cdot f(x)) = k \cdot f'(x) \quad \text{[Scalar Rule]}$$

$$\frac{d}{dx}(f(x)-g(x)) = f'(x)-g'(x) \quad \text{[Difference Rule]}$$

$$\frac{d}{dx}(f(x) \cdot g(x)) = f(x) \cdot g'(x) + g(x) \cdot f'(x) \quad \text{[Product Rule]}$$

$$\frac{d}{dx}\left(\frac{f(x)}{g(x)}\right) = \frac{g(x) \cdot f'(x) - f(x) \cdot g'(x)}{[g(x)]^2} \quad \text{[Quotient Rule]}$$

$$\frac{d}{dx}x^n = nx^{n-1} \quad \text{[Power Rule]}$$

$$\frac{d}{dx}(f(g(x))) = f'(g(x)) \cdot g'(x) \quad \text{[Chain Rule]}$$

Remark

The Difference Rule is a corollary of the Sum and Scalar Rules. The Quotient Rule can be seen as a corollary of the Product and Chain Rules. The Power Rule handles roots. These formulas are all in the tables — the functions are called u and v but they are the same ideas.

0.3.4 Function Catalogue

Constant Functions

1. **Definition** Let $k \in \mathbb{R}$. A *constant* function is of the form

$$f(x) = k.$$

   An example of a constant function is $f(x) = 2$.

2. **Main Idea/Properties** A constant function outputs the same number for all inputs and so has a rate of change of zero.

3. **Derivative** The derivative of a constant (function) is zero:

$$\frac{d}{dx}(k) = 0.$$

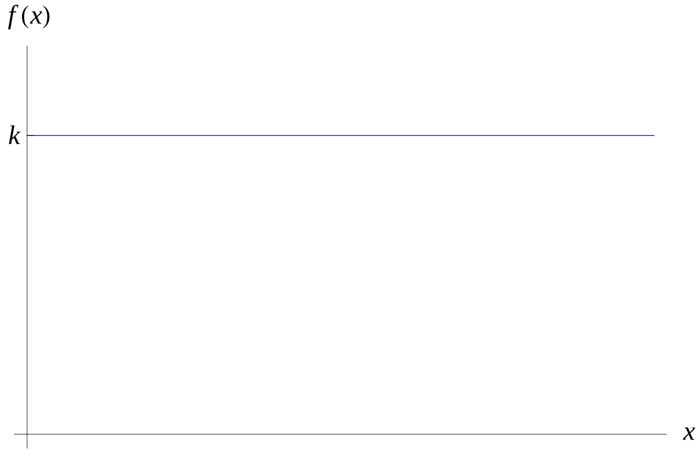


Figure 8: The graph of a constant function. Note that the slope morryah the derivative is zero.

Lines

1. **Definition** Let $m \in \mathbb{R}$ and $c \in \mathbb{R}$. A *line of slope m and y-intercept c* is given by

$$f(x) = mx + c.$$

   An example of a line is $f(x) = 3x - 2$.

2. **Main Idea/Properties** A line does exactly what it says on the tin. The slope/derivative of a line is a constant so the rate of change of a line is constant.

3. **Derivative** The derivative of a line is the slope:

$$\frac{d}{dx}(mx + c) = m.$$

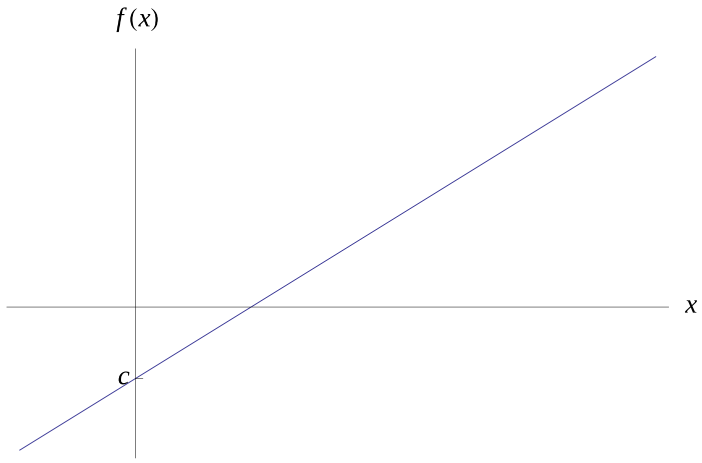


Figure 9: The graph of a line. Note that the slope morryah the derivative is constant.

Quadratics

1. **Definition** Let $a, b, c \in \mathbb{R}$. A quadratic is a function of the form

$$f(x) = ax^2 + bx + c.$$

   An example of a quadratic is $f(x) = x^2 + 1$.

2. **Main Idea/Properties** A quadratic either has a $\bigcup$ shape (when $a > 0$) or a $\bigcap$ shape (when $a < 0$). It has two *roots* given by the $\dfrac{-b \pm \sqrt{\cdots}}{2a}$ formula. If they are both *real* (when $b^2 - 4ac > 0$), then the graph cuts the x-axis at two points. The graph is symmetric about the max/min. Hence the max/min can be found by looking at $f'(x) = 0$ or else be found at the midpoint of the roots. If $b^2 - 4ac < 0$ then the roots contain a $\sqrt{(-)}$ — *complex roots.*

3. **Derivative** The derivative of a quadratic is a line!

$$\frac{d}{dx}(ax^2 + bx + c) = a(2x) + b(1) + 0 = 2ax + b.$$

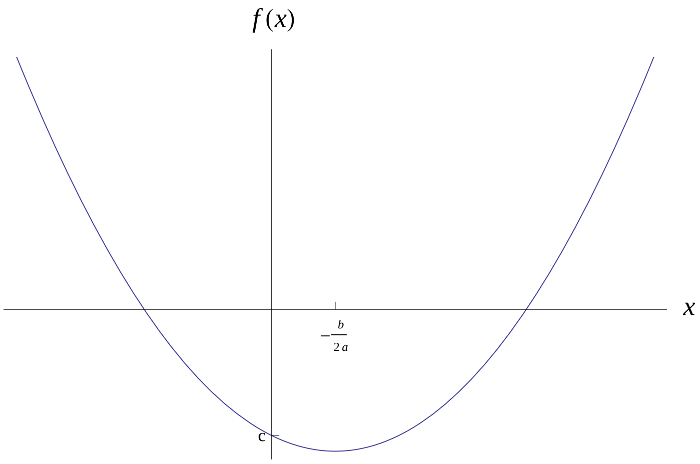


Figure 10: The graph of a quadratic with $a > 0$. Note that the slope goes from negative to zero to positive — like a line. This quadratic has two real roots and the minimum occurs at $-\dfrac{b}{2a}$. At this point the tangent is horizontal. This point can be found by differentiating $ax^2 + bx + c$, e.g. getting the slope, and setting it equal to zero.

Polynomials

1. **Definition** Let $a_n, a_{n-1}, \ldots, a_2, a_1, a_0 \in \mathbb{R}$. A *polynomial of degree* n is a function of the form

$$f(x) = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_2 x^2 + a_1 x + a_0.$$

   As example of a polynomial is $f(x) = x^5 - 3x^3 + 2x^2 + 1$.

2. **Main Idea/Properties** A polynomial of degree n has n roots — some of which may be complex, some of which may be repeated. However if all the roots are real and distinct then the polynomial cuts the x-axis n times.

The derivative of a polynomial of degree n is a polynomial of degree $n-1$;

$$\text{e.g.} \quad \frac{d}{dx}(x^5 - 3x^3 + 2x^2 + 1) = 5x^4 - 3(3x^2) + 2(2x) = 5x^4 - 9x^2 + 4x,$$

which has potentially $n-1$ real roots and hence $n-1$ points where $f'(x) = 0$ — potentially $n-1$ turning points. As an example note that quadratics are degree two polynomials and have one turning point.

3. **Derivative** We differentiate a polynomial using the Sum, Scalar & Power Rules:

$$\frac{d}{dx}(ax^n) = a\frac{d}{dx}x^n = a(nx^{n-1}) = anx^{n-1}.$$

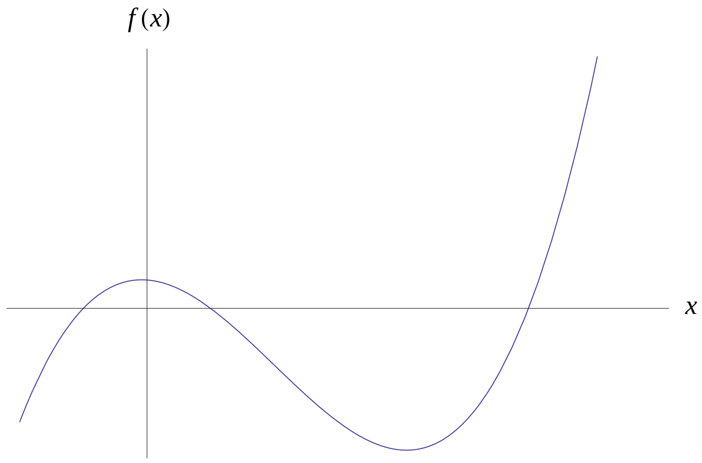


Figure 11: This is an example of a cubic: $ax^3 + bx^2 + cx + d$. Note that it has *three* real roots and *two* turning points. In some sense this is typical behaviour of polynomials.

Roots

1. **Definition** Let $n \in \mathbb{N}$. The *nth root function* is a function of the form:

$$f(x) = \sqrt[n]{x},$$

the *positive n-th root* of x.

2. **Main Idea/Properties** We can show that if we define

$$x^{1/n} = \sqrt[n]{x}$$

then all of the theorems of indices and differentiation work properly with this definition and it turns out that $\sqrt[n]{x} = x^{1/n}$ written as a power can be differentiated using the Power Rule.

3. **Derivative** Using the power rule

$$\frac{d}{dx}\sqrt[n]{x} = \frac{d}{dx}x^{1/n} = \frac{1}{n}x^{1/n-1}.$$

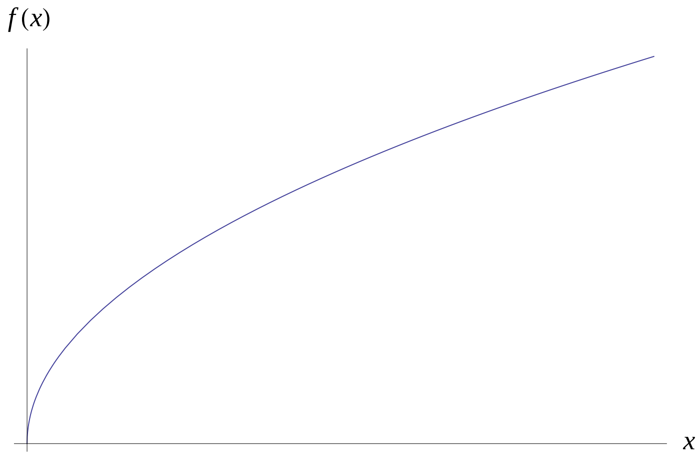


Figure 12: A plot of the square root function $f(x) = \sqrt{x}$. Note that roots are only defined for *positive* values of x.

Trigonometric

1. **Definition** Let $0 \leq \theta \leq 2\pi$ be an angle. Consider now the unit circle

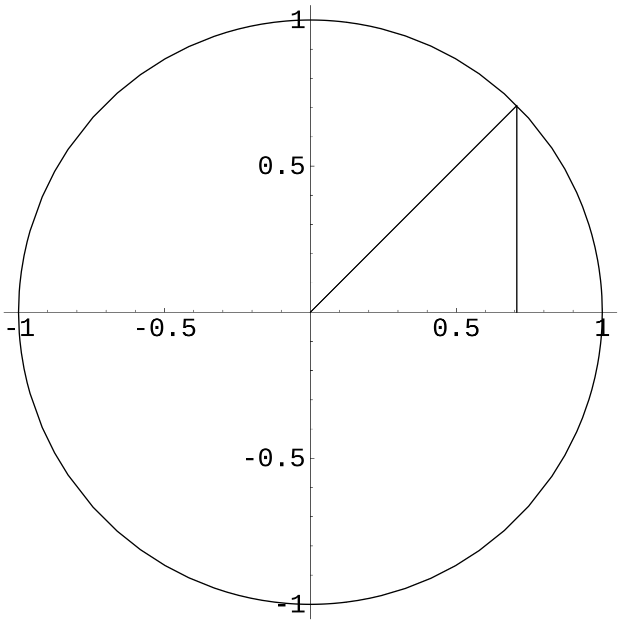


Figure 13: If the angle made with the positive x-axis is θ, then the coordinate of the point on the circle is $(\cos\theta, \sin\theta)$.

This defines cosine and sine for angles between 0 and 2π. The definition is extended by periodicity to the whole of the number line by

$$\cos(\theta + 2\pi) = \cos(\theta)$$
$$\sin(\theta + 2\pi) = \sin\theta$$

We also define

$$\tan x = \frac{\sin\theta}{\cos\theta}$$

2. **Main Idea/Properties** Sine and Cosine are waves that oscillate between ± 1. Sine begins at zero ($\sin 0 = 0$) while cosine begins at one ($\cos 0 = 1$). Apart from this they are very similar: the graph of sine is got by shifting the graph of cosine $\pi/2$ units to the right.

3. **Derivative** We can show that

$$\frac{d}{dx}\sin x = \cos x$$

$$\frac{d}{dx}\cos x = -\sin x.$$

Using the Quotient Rule we have

$$\frac{d}{dx}\tan x = \sec^2 x = (\sec x)^2,$$

where

$$\sec x := \frac{1}{\cos x}. \tag{3}$$

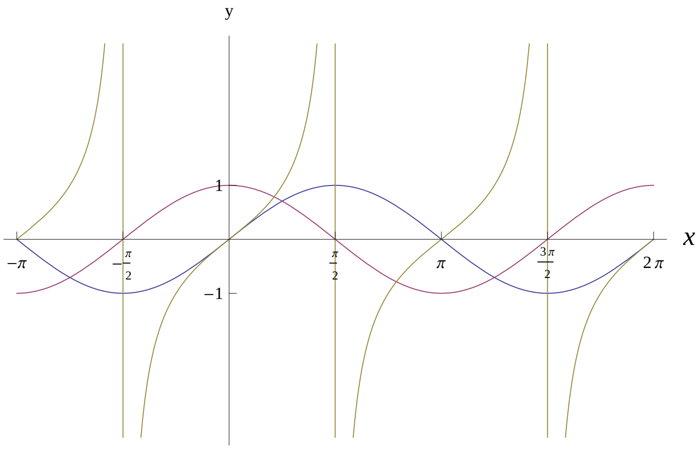


Figure 14: Note that $|\sin x|, |\cos x| \leq 1$ while $\tan x \to \pm\infty$ at $\pi/2$.

Inverse Trigonometric

1. **Definition** These functions are *inverse functions* of the trigonometric functions.

   Let $y \in [-\pi/2, \pi/2]$ and $x \in [-1, 1]$:

   $$y = \sin^{-1}(x) \Leftrightarrow x = \sin y. \tag{4}$$

   Let $y \in [-\pi/2, \pi/2]$:

   $$y = \tan^{-1}(x) \Leftrightarrow x = \tan y. \tag{5}$$

2. **Main Idea/Properties** These are the inverse functions of $\sin x$ and $\tan x$. For example

   $$\sin\theta = \frac{1}{2} \Rightarrow \sin^{-1}(\sin\theta) = \sin^{-1}(1/2) \Rightarrow \theta = \frac{\pi}{6}.$$

   In other words, $\sin^{-1}(x)$ asks for the angle — *between* $\pm\pi/2$ — which has a sine of x.

3. **Derivative** We can show that

   $$\frac{d}{dx}\sin^{-1}\left(\frac{x}{a}\right) = \frac{1}{\sqrt{a^2 - x^2}} \tag{6}$$

   and

   $$\frac{d}{dx}\tan^{-1}\left(\frac{x}{a}\right) = \frac{a}{a^2 + x^2}. \tag{7}$$

Exponential

1. **Definition** The *exponential function* can be defined as a power series

   $$e^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \frac{x^4}{4!} + \cdots \tag{8}$$

2. **Main Idea/Properties** The exponential function is the unique function that is equal to its own derivative.

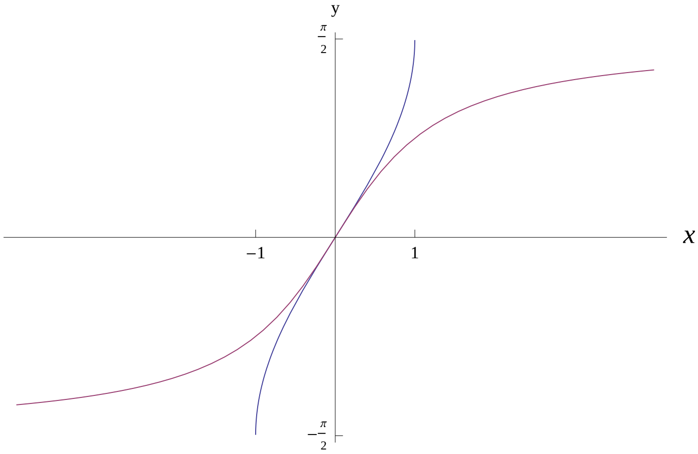


Figure 15: Inverse Sine can only take inputs between ± 1. However we have $\tan^{-1}(x) \to \pi/2$ as $x \to \infty$.

3. **Derivative** We have that

$$\frac{d}{dx}e^x = e^x, \tag{9}$$

and using the Chain Rule

$$\frac{d}{dx}e^{ax} = a \cdot e^{ax}. \tag{10}$$

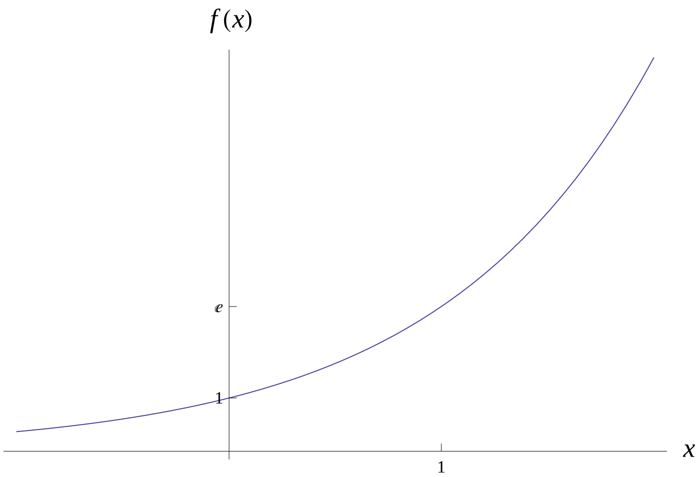


Figure 16: We have that $e^0 = 1$. $e^x \to \infty$ as $x \to +\infty$. When the input is negative, say $x = -N$, then $e^x = e^{-N} = \frac{1}{e^N} \to 0$ as $N \to \infty \Leftrightarrow x \to -\infty$.

Logarithmic

1. **Definition** The natural logarithm is the inverse function of e^x:

$$y = \ln x \Leftrightarrow x = e^y \tag{11}$$

2. **Main Idea/Properties** As $e^y > 0$, the natural logarithm can only take strictly positive inputs. They can be used to solve exponential equations:

$$e^x = 2 \Rightarrow \ln(e^x) = \ln 2 \Rightarrow x = \ln 2.$$

Note that we have

$$\ln 1 = 0$$
$$\ln(xy) = \ln x + \ln y$$
$$\ln(x^n) = n \ln x.$$

3. **Derivative** We can show that

$$\frac{d}{dx} \ln x = \frac{1}{x}.$$

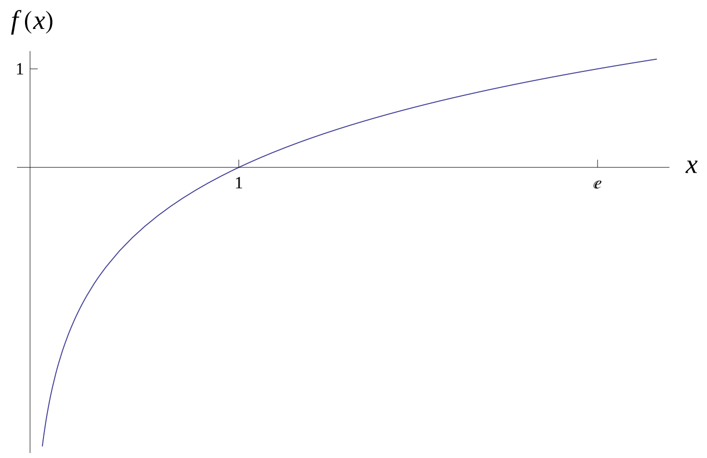


Figure 17: As $x \to 0$, $\ln x \to -\infty$; $\ln 1 = 0$ and $\ln x \to \infty$ — slowly — as $x \to \infty$.

Exercises

(a) Find the first and second derivatives of each of the following functions

$$(a)\ \frac{4}{3x^2}\quad (b)\ (2x+1)^5\quad (c)\ \frac{5}{2t+3}\quad (d)\ \ln(3x+2)$$

$$(f)\ \ln(x^2+2)\quad (g)\ x\ln(2x+1)\quad (h)\ \sin 3t\quad (i)\ \tan x\quad (j)\ \sec x$$

$$(k)\ \ln(\sec x)\quad (l)\ \ln(\cos x)\quad (m)\ \ln(\sec x + \tan x)\quad (n)\ \tan^{-1}(x)\quad (o)\ \arctan(x/2)$$

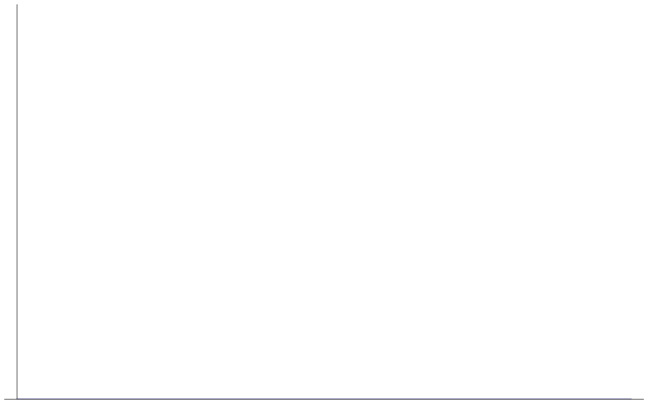
Selected Answers: Some answers to second derivative: (d) $\frac{72}{(3x-2)^3}$, (e) $\frac{-9}{(3x+2)^2}$, (f) $\frac{4-2x^2}{(x^2+2)^2}$, (g) 1st $\ln(2x+1) + \frac{2x}{2x+1}\dots$ 2nd $\frac{2}{2x+1} + \frac{2}{(2x+1)^2}$, (i) $2\tan x \sec^2 x$, (j) $\sec^3 x + \sec x \tan^2 x$, (l) $-\sec^2 x$, (m) 1st $\sec x$ and 2nd $\sec x \tan x$, (n) 1st $\frac{1}{1+x^2}\dots$ and 2nd $\frac{-2x}{(x^2+1)^2}$, (o) 1st $\frac{2}{x^2+4}\dots$ and 2nd $\frac{-4x}{(x^2+4)^2}$.

(b) If $f(x) = \ln(1+\cos 2x)$ show that $f'(x) = -2\tan x$. [Note: $\sin 2A = 2\sin A\cos A$ and $\cos 2A = 2\cos^2 A - 1$]

(c) If $f(x) = \ln(\sec x)$ show that $f'(x) = \sec^2 x$.

0.3.5 Integration

What is the area of the shaded region under the curve $f(x)$?



Start by subdividing the region into n strips S_1, S_2, ..., S_n of equal width as Figure 18.

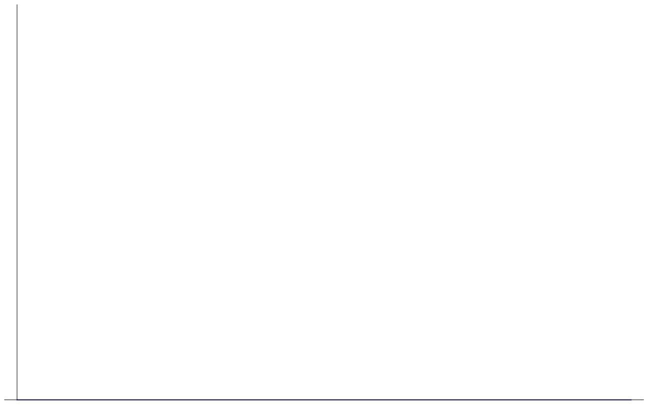
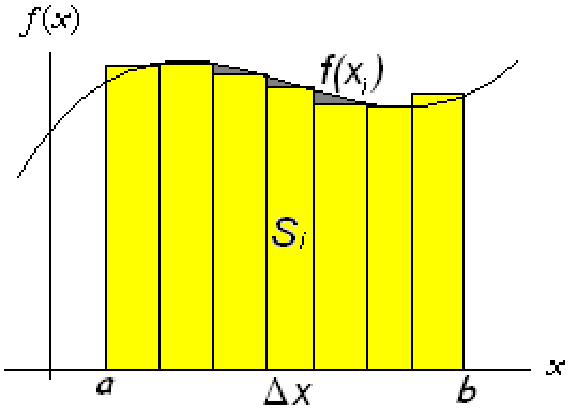


Figure 18:

The width of the interval $[a, b]$ is $b - a$ so the width of each of the n strips is

Approximate the ith strip S_i by a rectangle with width Δx and height $f(x_i)$, which is the value of f at the right endpoint. Then the area of the ith rectangle is $f(x_i)\ \Delta x$:



The area of the original shaded region is approximated by the sum of these rectangles:

This approximation becomes better and better as the number of strips increases, that is, as $n \to \infty$. Therefore the area of the shaded region is given by the limit of the sum of the areas of approximating rectangles:

Definition: The Definite Integral

If $f(x)$ is a *continuous* function defined in $[a, b]$ and x_i, Δx are as defined above, then the *definite integral of f from a to b* is

So an integral is an infinite sum. Associate $\int \cdot\, dx \sim \lim_{n\to\infty} \sum_n \cdot\, \Delta x$.

Fundamental Theorem of Calculus

If f is a function with derivative f' *then*

Examples

(a) Evaluate

$$\int_0^2 3x^2\,dx$$

Solution:

(b) Evaluate

$$\int_1^e \frac{1}{x}\,dx$$

Solution:

(c) Evaluate

$$\int_0^\pi -\sin x\,dx$$

Solution:

Definition: The Indefinite Integral

If $f(x)$ is a function and it derivative with respect to x is $f'(x)$, then

where C is called the *constant of integration.*

The Indefinite Integral $\int f(x)\,dx$ asks the question:

Note the constant of integration. It's inclusion is vital because if $f(x)$ is a function with derivative $f'(x)$ then $f(x) + C$ also has derivative $f'(x)$ as:

Geometrically a curve $f(x)$ with slope $f'(x)$ has the same slope as a curve that is shifted upwards; $f(x)+C$. Note that the constant of integration can be disregarded for the indefinite integral. Suppose the integrand is $f'(x)$ and the anti-derivative is $f(x)+C$. Then:

Finding the derivative of a function f at x is finding the slope of the tangent to the curve at x. Integration meanwhile measures the area between two points $x = a$ and $x = b$. The Fundamental Theorem of Calculus states however that differentiation and integration are intimately related; that is given a function f:

i.e. differentiation and integration are essentially inverse processes.

Examples

Integrate 1-3:

(a) $\int 3x^2\,dx$

Solution:

(b) $\int (1/x)\,dx$

Solution:

(c) $\int -\cos x\,dx$

Solution:

(d)

$$\int_0^{\pi} 4x^3\,dx.$$

Solution:

Straight Integration

From the Fundamental Theorem of Calculus

$$\int f'(x)\,dx = f(x) + C \qquad (12)$$

Thus:

$f(x)$	$\int f(x)$
$x^n\ (n \neq -1)$	$\frac{x^{n+1}}{n+1} + C$
$\cos x$	$\sin x + C$
$\sin x$	$-\cos x + C$
e^x	$e^x + C$
$\sec^2 x$	$\tan x + C$
$\frac{1}{x}$	$\ln \lvert x \rvert + C$

Chapter 1

Further Calculus

The journey for an education starts with a childhood question.

David L. Finn

1.1 Maclaurin & Taylor Series

1.1.1 Power Series

How does a calculator work? Addition & Multiplication are fairly easy to programme but how about finding $\ln 2$? However no calculator can calculate $\ln 2$ exactly — it is an *irrational number* which means that it has a non-repeating decimal expansion. Therefore calculators must use approximations. Here we describe some methods that we might do this.

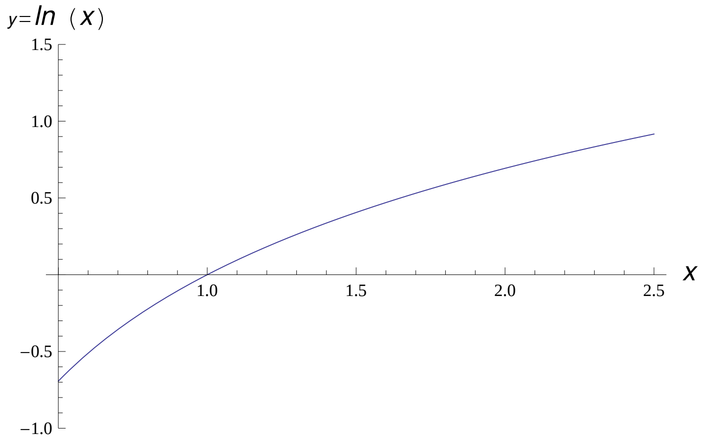


Figure 1.1: Smooth functions such as $y = \ln x$ admit good linear approximations.

To do this we pick a point close to $x = 2$, whose log we *do* know, for example $x = 0$. Now we draw the tangent to the curve at $x = 1$. To get the equation of the tangent to the curve we need a point (x_1, y_1) and a slope. We have $(x_1, y_1) = (1, 0)$ so all we need is a slope

So therefore we have for points near $x = 1$ we have

Now instead of fitting a line to the curve we could fit a quadratic or more complicated *polynomials* to the curve[1] and all are somewhat easy for a calculator as they are just combinations of addition and multiplication:

In general, all of these polynomial approximations are just that... approximations. However would it be possible to have a series of polynomial approximations to a non-polynomial function that converges?

The first clue that this might be possible was the following fact:

Geometric Series

For $|x| < 1$, we have the following

Now this formula is exact in the sense that no matter how close to $\frac{1}{1-x}$ as you want to get, you can get that close by taking enough terms of the 'infinite polynomial'. For example to using only five terms with $x = 0.2$:

$$\frac{1}{1-0.2} = 1.25 \approx 1.2499968 = 1 + 0.2 + 0.2^2 + 0.2^3 + 0.2^4.$$

Formally, such infinite sums of powers of x are called *power series*

In general, power series are only valid for some values of x. The example above only works for $|x| < 1$ (why??). Functions that can be somewhere-well-approximated by power series are called *analytic functions*. Now there are examples functions which are *not* well-behaved, but the ones of any use to ye are either well-enough understood in their original form[2] or else not of much use. We examine in particular two classes of power series.

[1] more on this in MATH7021

[2] see Step Functions later

1.1.2 Maclaurin Series

Above we used a line near $x = 1$ to approximate $\ln x$. The reason we used $x = 1$ was because we knew all about the function at that point: its value and slope. As a rule of thumb, we also tend to know all about functions near $x = 0$ and the Maclaurin Series makes use of this.

Given a function $y = f(x)$ the first thing we assume is that it is analytic near $x = 0$... i.e. a power series is going to work. So we assume that we can write

The problem therefore is to find the values of the coefficients $a_0, a_1, a_2, \dots$. This is actually an infinite number of problems but we can use differentiation to help. If the left-hand function is equal to the right-hand function then they must be equal for all values of x... and their first derivatives, and second derivatives and in fact all of their derivatives. We use this to our advantage as follows. What is happening at $x = 0$:

Now differentiate both functions and examine what happens at $x = 0$:

Now repeat the trick:

Now once more:

With a bit of thought you will see that the coefficients a_n are given by

So we have

Maclaurin Series Formula

If $f(x)$ is analytic near $x = 0$ then we have

$$f(x) = f(0) + f'(0)x + \frac{f''(0)}{2!}x^2 + \frac{f'''(0)}{3!}x^3 + \cdots = \sum_{k=0}^{\infty} \frac{f^{(k)}(0)}{k!}x^k. \tag{1.1}$$

Winter 2012 Question 2 (a)

Consider the function $f(x) = \ln(\sec x)$.

(i) Show that $f''(x) = \sec^2 x$.

(ii) Find the first two non-zero terms of a Maclaurin series of $f(x)$.

Solution:

(i) To differentiate we need the chain rule:

To find $f''(x)$ we differentiate again. We see in the tables that $(\tan x)' = \sec^2 x$ so we are done.

(ii) We know that the Maclaurin series is given by

$$f(x) \approx \sum_{k=0}^{\infty} \frac{f^{(k)}(0)}{k!} x^k = f(0) + f'(0)x + \frac{f''(0)}{2!}x^2 + \frac{f'''(0)}{3!}x^3 + \cdots .$$

We calculate

This means we also need the value of $f'''(0)$. We differentiate $\sec^2 x$ using the chain rule and evaluate at $x = 0$:

this is also equal to zero so we must differentiate again, this time requiring the product rule:

So we have that

$$\ln(\sec x) \approx \frac{1}{2}x^2 + \frac{2}{4!}x^4.$$

Winter 2010 Question 2 (a) (ii)

The maximum deflection $\delta_{\max}$ of a beam of span L is given by

$$\delta_{\max} = \frac{WEI}{P}\left(\sec\left(\frac{mL}{2}\right) - 1\right)\;, \text{ where } m^2 = \frac{P}{EI}.$$

Solution: First we use the Maclaurin Series on $\sec\frac{mL}{2}$

Now we calculate

Exercises:

(a) Find the first three non-vanishing terms of the Maclaurin Series of e^x. Ans: $1+x+\frac{1}{2}x^2$.

(b) Find the first two non-vanishing terms of the Maclaurin Series of $\sin x$. Ans: $x - \frac{x^3}{3!}$.

(c) Find the first three non-vanishing terms of the Maclaurin Series of $\ln(1+x)$. Ans: $x - \frac{x^2}{2} + \frac{x^3}{3}$.

Au '11 Q. 2 (a) (i) Find the first two non-vanishing terms of the Maclaurin Series of $\tan x$. [Note: $\sec x = \dfrac{1}{\cos x}$, $\cos 0 = 1$, $\tan 0 = 0$]. Ans: $\tan x \approx x + \frac{x^3}{3}$.

Wi '12 Q. 2 (a) (i) Find the first two non-vanishing terms of the Maclaurin Series of $\sec x$. [Note: $\sec x = \dfrac{1}{\cos x}$, $\cos 0 = 1$, $\tan 0 = 0$]. Ans: $\sec x \approx 1 + \frac{1}{2}x^2$.

Au '10 Q. 2 (a) (ii) Due to a downward load P applied to a cantilever of length L, the horizontal deflection $\delta_{\max}$ is given by

$$\delta_{\max} = \frac{1-\cos(kL)}{\cos(kL)}\;, \text{ where } k = \sqrt{\frac{P}{EI}}.$$

Using the Maclaurin Series $\sec x \approx 1 + \frac{x^2}{2}$ found above, show that for small values of P the approximation

$$\delta_{\max} \approx \frac{PL^2}{2EI}.$$

1.1.3 Taylor Series

What is the Maclaurin Series of $\ln x$? OK first of all what is $\ln 0$... This example shows that not all functions are analytic near $x = 0$. However, suppose that a function $y = f(x)$ is analytic near a point $x = a$? It turns out that if we write

then an identical analysis to that above yields:

Taylor Series Formula

If $f(x)$ is analytic near $x = a$ then we have

$$f(x) = f(a) + f'(a)(x-a) + \frac{f''(a)}{2!}(x-a)^2 + \frac{f'''(a)}{3!}(x-a)^3 + \cdots = \sum_{k=0}^{\infty} \frac{f^{(k)}(a)}{k!}(x-a)^k. \quad (1.2)$$

Remark

The Maclaurin Series is just the Taylor Series expansion about $x = a = 0$.

Winter 2012 Question 2 (a)

Using the fact that $f''(x) = -\sec^2 x$ when $f(x) = \ln(\cos x)$, find the first three terms of the Taylor Series of $f(x)$.

Solution: We know that the Taylor Series about $x = \pi/4$ is given by

$$f(x) \approx \sum_{k=0}^{\infty} \frac{f^{(k)(\pi/4)}}{k!}(x-\pi/4)^k = f(\pi/4) + f'(\pi/4)(x-\pi/4) + \frac{f''(\pi/4)}{2!}(x-\pi/4)^2 + \frac{f'''(\pi/4)}{3!}(x-\pi/4)^3 + \cdots.$$

We calculate

So we have that

$$\ln(\sec x) \approx \ln\left(\frac{1}{\sqrt{2}}\right) - (x - \pi/4) - \frac{1}{2}(x - \pi/4)^2.$$

Exercises:

Au '12 Q. 2 (a)(i) Find the first three terms of the Taylor Series of $f(x) = \ln(\sec x)$ about the point $x = \dfrac{\pi}{4}$. [NOTE: $\sin(\pi/4) = \cos(\pi/4) = 1/\sqrt{2}$, $\tan(\pi/4) = 1$] Ans: $\dfrac{\ln 2}{2} + \left(x - \dfrac{\pi}{4}\right) + \left(x - \dfrac{\pi}{4}\right)^2$.

Wi '11 Q. 2 (a)(i) Find the first three terms of the Taylor Series of $f(x) = \tan x$ about the point $x = \dfrac{\pi}{4}$. [NOTE: $\sin(\pi/4) = \cos(\pi/4) = 1/\sqrt{2}$, $\tan(\pi/4) = 1$] Ans: $1 + 2\left(x - \dfrac{\pi}{4}\right) + 2\left(x - \dfrac{\pi}{4}\right)^2$.

Au '10 Q. 2 (a)(i) Find the first three terms of the Taylor Series of $f(x) = \sec x$ about the point $x = \dfrac{\pi}{4}$. [NOTE: $\sec x = 1/\cos x$, $\sin(\pi/4) = \cos(\pi/4) = 1/\sqrt{2}$, $\tan(\pi/4) = 1$] Ans: $\sqrt{2} + \sqrt{2}\left(x - \dfrac{\pi}{4}\right) + \dfrac{3}{\sqrt{2}}\left(x - \dfrac{\pi}{4}\right)^2$.

Au '09 Q. 2 (a)(i) Find the first three terms of the Taylor Series of $f(x) = \ln(\cos x)$ about the point $x = \dfrac{\pi}{4}$. [NOTE: $\sec x = 1/\cos x$, $\sin(\pi/4) = \cos(\pi/4) = 1/\sqrt{2}$, $\tan(\pi/4) = 1$] Ans: $-\dfrac{\ln 2}{2} - \left(x - \dfrac{\pi}{4}\right) - \left(x - \dfrac{\pi}{4}\right)^2$.

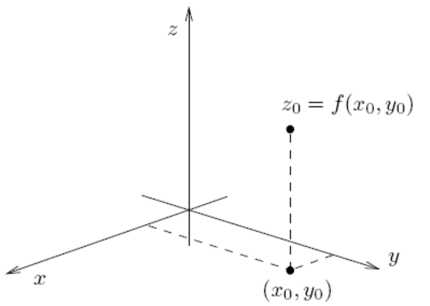
Wi '08 Q. 2 (a)(i) Find the first three terms of the Taylor Series of $f(x) = \ln(\sec x + \tan x)$ about the point $x = \dfrac{\pi}{4}$. [NOTE: $\sec x = 1/\cos x$, $\sin(\pi/4) = \cos(\pi/4) = 1/\sqrt{2}$, $\tan(\pi/4) = 1$] Ans: $\ln(1 + \sqrt{2}) + \sqrt{2}\left(x - \dfrac{\pi}{4}\right) + \dfrac{1}{\sqrt{2}}\left(x - \dfrac{\pi}{4}\right)^2$.

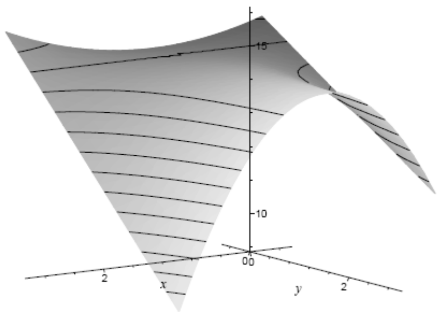
1.2 Partial Differentiation

Functions of Several Variables: Surfaces

Many equations in engineering, physics and mathematics tie together more than two variables. For example Ohm's Law ($V = IR$) and the equation for an ideal gas, $PV = nRT$, which gives the relationship between pressure (P), volume (V) and temperature (T). If we vary any two of these then the behaviour of the third can be calculated:

How P varies as we change T and V is easy to see from the above, but we want to adapt the tools of one-variable calculus to help us investigate functions of more than one variable.
For the most part we shall concentrate on functions of two variables such as $z = x^2 + y^2$ or $z = xsin(y + e^x)$. Graphically $z = f(x, y)$ describes a surface in 3D space — varying the x- and y-coordinates gives the z-coordinate, producing the surface:





As an example, consider the function $z = x^2 + y^2$. If we choose a positive value for z, for example $z = 4$, then the points (x, y) that can give rise to this value are those satisfying $x^2 + y^2 = 4 = 2^2$, i.e. those on the circle centred on the origin of radius 2. Note that at $(x, y) = (0, 0)$, $z = 0$, but if $x \neq 0$ or $y \neq 0$, then $x^2 > 0$ or $y^2 > 0$, and it follows that $z > 0$. Thus the minimum value taken by this function is $z = 0$, at the origin:

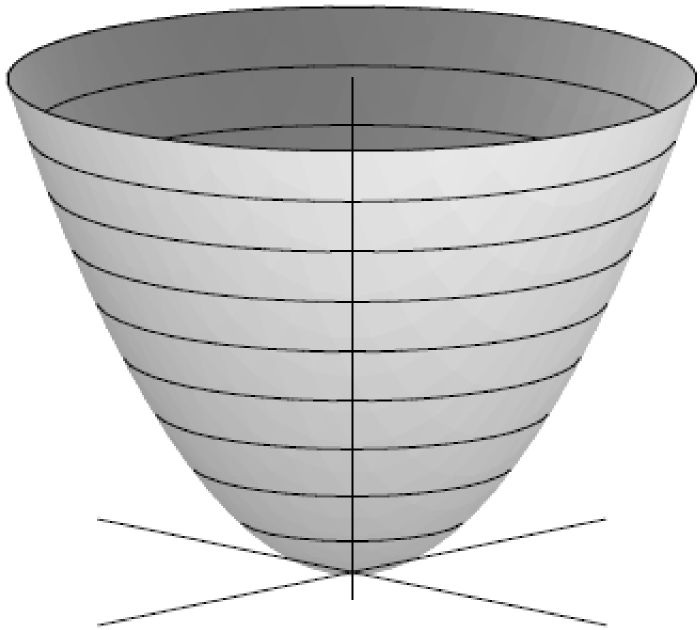
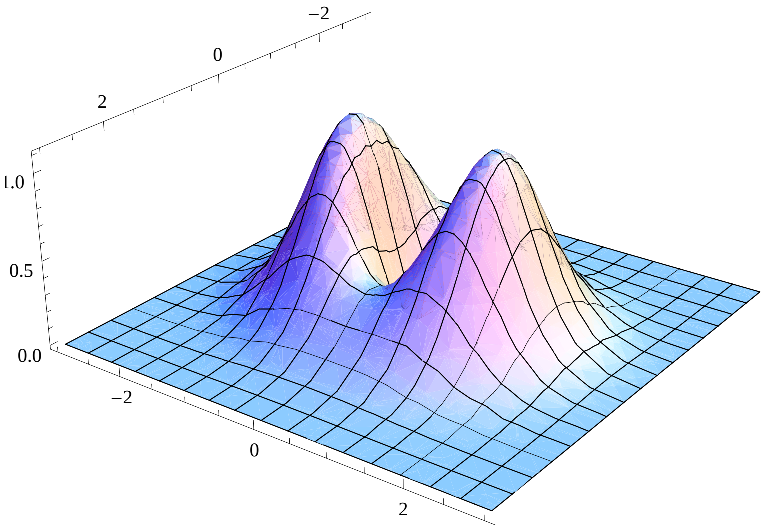
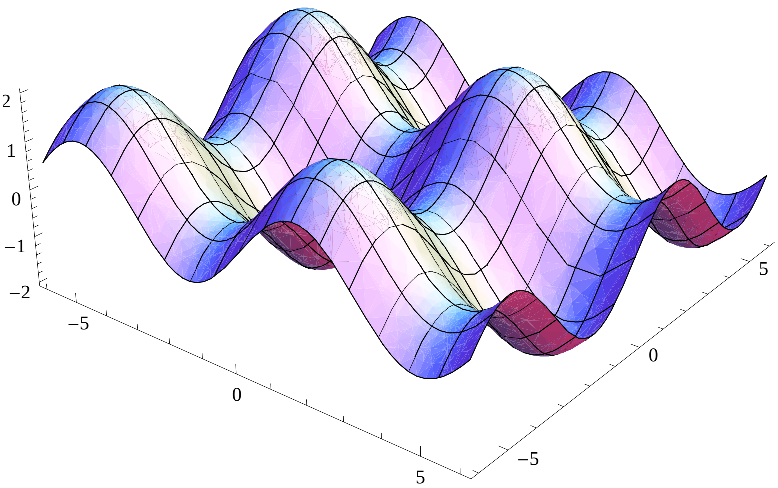


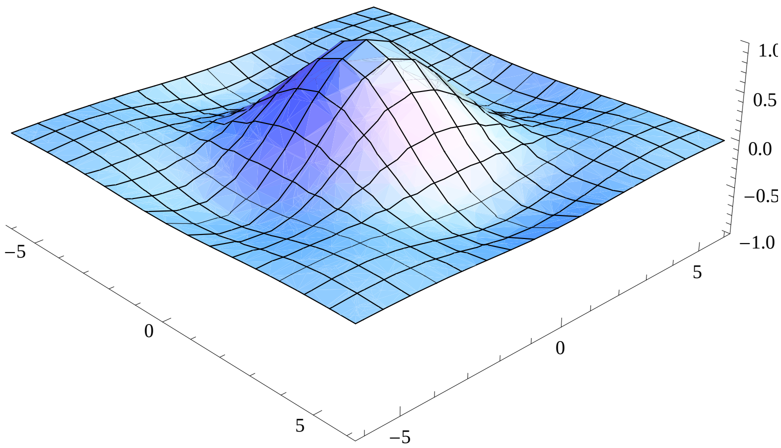
Figure 1.2: The surface defined by the relation $x^2 + y^2 = 2$.

Three examples. Which are which?

$$f(x,y) = (x^2 + 3y^2)e^{-x^2-y^2}$$
$$g(x,y) = \frac{\sin x \sin y}{xy}$$
$$h(x,y) = \sin x + \sin y$$






1.0
0.5
0.0
−0.5
−1.0
−5
0
5
−5
0
5

Partial Derivatives

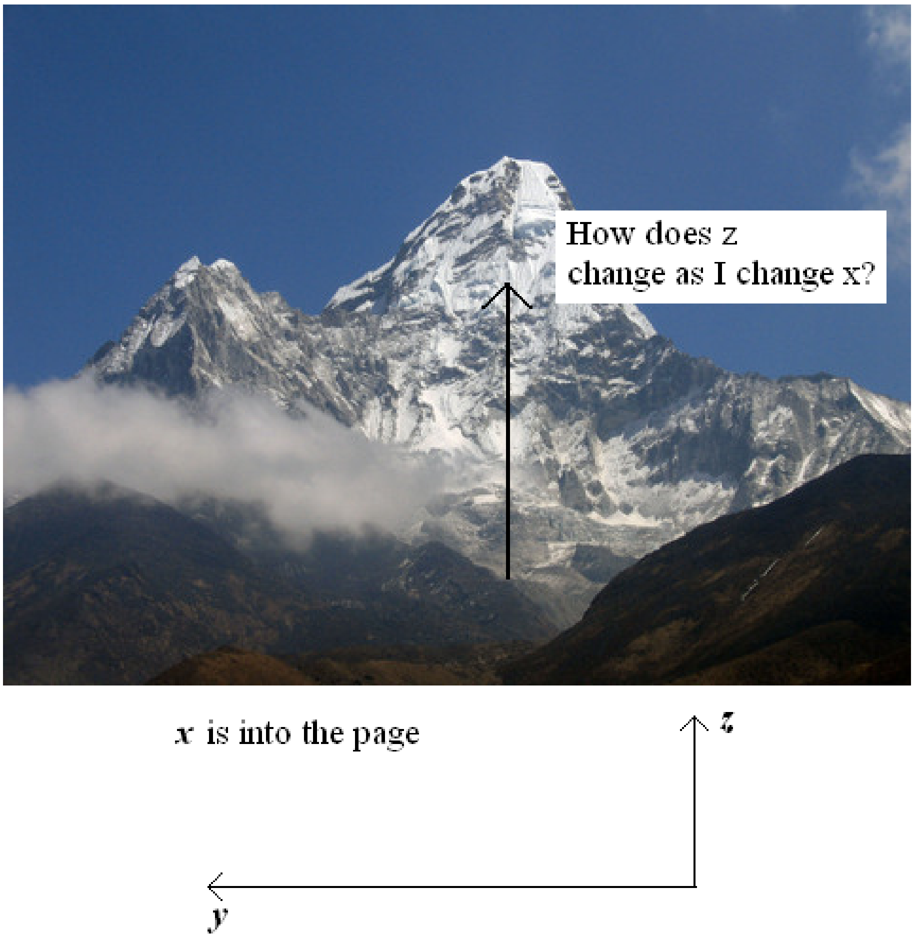


Figure 1.3: What is the rate of change in z as I keep y constant

If we were to look at this from side on:

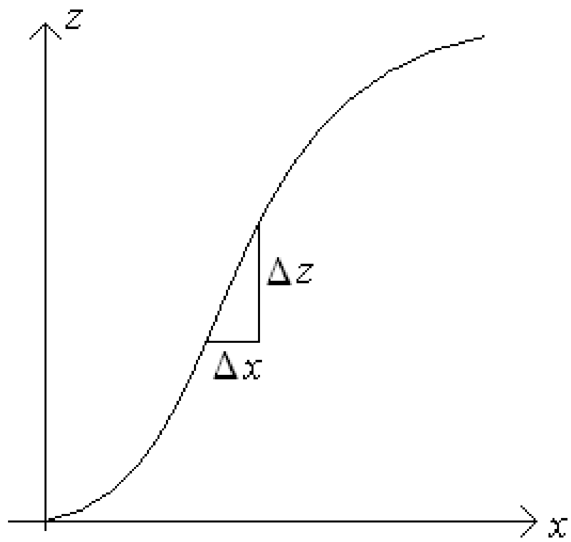


Figure 1.4: When y is a constant z can be considered a function of x only.

In general we have that $z = f(x, y)$; but if $y = b$ is fixed (constant):

We can view $f(x, b)$ as a function of x alone. Now what is the rate of change of a single-variable function $g(x)$ with respect to x:

Which is also the slope of the tangent to f at x. Hence the rate of change of $f(x, y)$ with respect to x at $x = a$ when y is fixed at $y = b$ is the slope of the surface in the x-direction.

Example

Let $z = f(x, y) = x^3 + x^2y^3 - 2y^3$. What is the rate of change of z with respect to x when $y = 2$?

Solution:

Hence the rate of change of z with respect to x, when y is fixed at $y = b$, is given by:

More generally, we fix $y = y$ and define

as *the partial derivative of f with respect to x.*

We define the partial derivative of f with respect to y in exactly the same way.

Example

What are the partial derivatives of

$$z = x^2 + xy^5 - 6x^3y + y^4$$

with respect to x and y respectively?

Solution:

There are many alternative notations for partial derivatives. For instance, instead of $\frac{\partial f}{\partial x}$ we can write f_x or f_1. In fact,

$$\frac{\partial f}{\partial x} \equiv \frac{\partial z}{\partial x} \equiv f_x(x,y) \equiv f_1(x,y)$$
$$\frac{\partial f}{\partial y} \equiv \frac{\partial z}{\partial y} \equiv f_y(x,y) \equiv f_2(x,y)$$

To compute partial derivatives, all we have to do is remember that the partial derivative of a function with respect to x is the same as the *ordinary* derivative of the function g of a single variable that we get by keeping y fixed. Thus we have the following:

1. To find $\frac{\partial f}{\partial x}$, regard y as a constant and differentiate $f(x,y)$ with respect to x.
2. To find $\frac{\partial f}{\partial y}$, regard x as a constant and differentiate $f(x,y)$ with respect to y.

Example

If $f(x,y) = 4 - x^2 - 2y^2$, find $f_x(1,1)$ and $f_y(1,1)$ and interpret these numbers as slopes.

Solution:

Using this technique we can make use of known results from one-variable theory such as the product, quotient and chain rules (Careful — the Chain rule only works if we are differentiating with respect to one of the variables — we may have more to say on this in the next section).

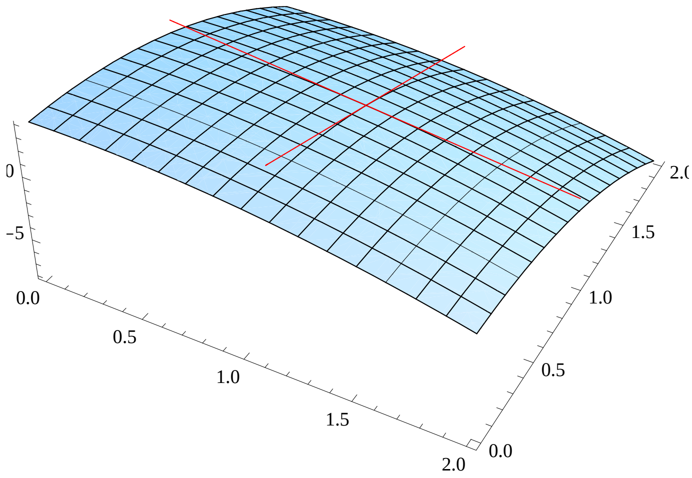


Figure 1.5: The slope in the y-direction at $(1,1)$ is $f_y(1,1) = -2$ whilst the slope in the x-direction at $(1,1)$ is $f_x(1,1) = -4$.

Examples

Find the partial derivative with respect to y of the function

$$f(x,y) = \sin(xy)e^{x+y}$$

Solution:

Compute f_1 and f_2 when $z = x^2y + 3x\sin(x-2y)$.

Solution:

Functions of More Variables

We can extend the notion of partial derivatives to functions of any (finite number) of variables in a natural way. For example if $w = \sin(x + y) + z^2e^x$ then:

Higher Order Derivatives

Suppose $z = x \sin y + x^2y$. Then

Both of these partial derivatives are again functions of x and y, so we can differentiate both of them, either with respect to x, or with respect to y. This gives us a total of four *second order partial derivatives:*

Remark

The mixed partial derivatives in this case are equal:

$$\frac{\partial^2 z}{\partial y \partial x} = \frac{\partial^2 z}{\partial x \partial y}.$$

This is not something special about our particular example — it will be true for all *reasonably well-behaved functions.* This is the *symmetry of second derivatives* and is often known as *Clairaut's Theorem.* Note the confusing notation:

$$\frac{\partial}{\partial x \partial y} = f_{yx} \quad \text{etc.} \tag{1.3}$$

Examples

Compute

$$\frac{\partial z}{\partial x}\,,\ \frac{\partial z}{\partial y}\ ,\ \text{and}\ \frac{\partial^2 z}{\partial x^2}$$

when $z = x^3 y + e^{x+y^2} + y \sin x$.

Solution:

Compute all the second order partial derivatives of the function $f(x, y) = \sin(x + xy)$.

Solution:

Example

Given $s = -5x^3 + 3x^2y - 2y^2$ find

$$\frac{\partial s}{\partial x}, \frac{\partial s}{\partial y} \text{ and } \frac{\partial^2 s}{\partial y \partial x}.$$

[5 Marks]

Solution: We calculate

Example

Given $z = x^3y + y^2$ find $\frac{\partial z}{\partial x}, \frac{\partial^2 z}{\partial x^2}$ and $\frac{\partial^2 z}{\partial x \partial y}$.

Solution: We calculate

Example

Given the function $z = \ln(x^2 + y^2)$ show that

$$\frac{\partial^2 z}{\partial x^2} + \frac{\partial^2 z}{\partial y^2} = 0.$$

[8 Marks]

Solution: We calculate

Exercises

1. Find all the first order derivatives of the following functions:

$$(i)\ f(x,y) = x^3 - 4xy^2 + y^4 \qquad (ii)\ f(x,y) = x^2e^y - 4y$$
$$(iii)\ f(x,y) = x^2 \sin xy - 3y^2 \quad (iv)\ f(x,y,z) = 3x \sin y + 4x^3y^2z$$

2. Find the indicated partial derivatives: (i) $f(x,y) = x^3 - 4xy^2 + 3y$: f_{xx}, f_{yy}, f_{xy}
(ii) $f(x,y) = x^4 - 3x^2y^3 + 5y$: f_{xx}, f_{xy}, f_{xyy}
(iii) $f(x,y,z) = e^{2xy} - \frac{z^2}{y} + xz \sin y$: f_{xx}, f_{yy}, f_{yyzz}

Selected Solutions:

1. (i)

$$\frac{\partial f}{\partial x} = 3x^2 - 4y^2(1) + 0.$$

$$\frac{\partial f}{\partial y} = -4x(2y) + 4y^3 = 4y^4 - 8xy.$$

(ii)

$$\frac{\partial f}{\partial x} = e^y(2x) + 0 = 2xe^y.$$

$$\frac{\partial f}{\partial y} = x^2(e^y) - 4 = x^2e^y - 4.$$

(iii) This one needs a product and a chain rule for f_x and a chain rule for f_y.

$$\begin{aligned}\frac{\partial f}{\partial x} &= x^2 \times \frac{\partial \sin xy}{\partial x} + \sin xy \times \frac{\partial x^2}{x} + 0 \\ &= x^2 \times \cos xy \times \frac{\partial xy}{\partial x} + \sin xy \times 2x \\ &= x^2 \cos xy \times y + 2x \sin xy = x^2 y \cos xy + 2x \sin xy.\end{aligned}$$

$$\begin{aligned}\frac{\partial f}{\partial y} &= x^2 \times \cos xy \times \frac{\partial xy}{\partial y} - 6y \\ &= x^2 \times \cos xy \times x - 6y = x^3 \cos xy - 6y.\end{aligned}$$

(iv)

$$\frac{\partial f}{\partial x} = \sin y(3) + 4y^2z(3x^2) = 3(\sin y + 4x^2y^2z).$$

$$\frac{\partial f}{\partial y} = 3x(\cos y) + 4x^3z(2y) = 3x\cos y + 8x^2yz.$$

$$\frac{\partial f}{\partial z} = 0 + 4x^3y^2(1) = 4x^3y^2.$$

2. (i)

$$f_x = 3x^2 - 4y^2(1) + 0 = 3x^2 - 4y^2$$

$$f_{xx} = 6x + 0 = 6x.$$

$$f_y = 0 - 4x(2y) + 3 = -8xy + 3.$$

$$f_{yy} = -8x(1) + 0 = -8x.$$

$$f_{xy} = (f_x)_y = 0 - 8y = -9y.$$

(ii)

$$f_x = 4x^3 - 3y^3(2x) + 0 = 4x^3 - 6xy^3 = .$$

$$f_{xx} = 12x^2 - 6y^3(1) = 6(2x^2 - y^3).$$

$$f_{xy} = (f_x)_y = 0 - 6x(3y^2) = -18xy^2.$$

$$f_{xyy} = (f_{xy})_y = -18x(2y) = -36y.$$

(iii) This example is going to require the chain rule when we differentiate e^{2xy}. Also to ease differentiation, we use the fact that $1/a^n = a^{-n}$ to write:

$$f(x, y, z) = e^{2xy} - z^2y^{-1} + xz\sin y.$$

$$f_x = e^{2xy} \times \frac{\partial 2xy}{\partial x} + 0 + z\sin y = e^{2xy} \times 2y + z\sin y = 2ye^{2xy} + z\sin y.$$

$$f_{xx} = 2ye^{2xy} \times \frac{\partial 2xy}{\partial y} + 0 = 2ye^{2xy} \times 2y = 4y^2e^{2xy}.$$

$$\begin{aligned} f_y &= e^{2xy} \times \frac{\partial 2xy}{\partial y} - z^2(-1y^{-2}) + xz(\cos y) = e^{2xy} \times 2x + z^2y^{-2} + xz\cos y, \\ &= 2xe^{2xy} + z^2y^{-2} + xz\cos y. \end{aligned}$$

$$f_{yy} = 2xe^{2xy}\frac{\partial 2xy}{\partial y} + z^2((-2)y^{-3}) + xz(-\sin y),$$
$$= 2xe^{2xy} \times 2x - 2\frac{z^2}{y^3} - xz\sin y = 4x^2e^{2xy} - 2\frac{z^2}{y^3} - xz\sin y.$$

Well $f_{yyzz} = (f_{yy})_{zz}$ so first we evaluate:

$$(f_{yy})_z = 0 - \frac{4z}{y^3}(1) - x\sin y(1) = \frac{4z}{y^3} - x\sin y,$$

$$(f_{yy})_{zz} = \frac{4}{y^3}.$$

1.3 Differentials & Applications to Error Analysis

Differentials

For a differentiable function $y = f(x)$ of a single variable x, we define the differential 'dx' to be an independent variable; that is, dx can be given the value of any real number. Differentiable functions are *locally approximately linear*: the tangent at x approximates the function well near x. The differential of y is then defined by:c

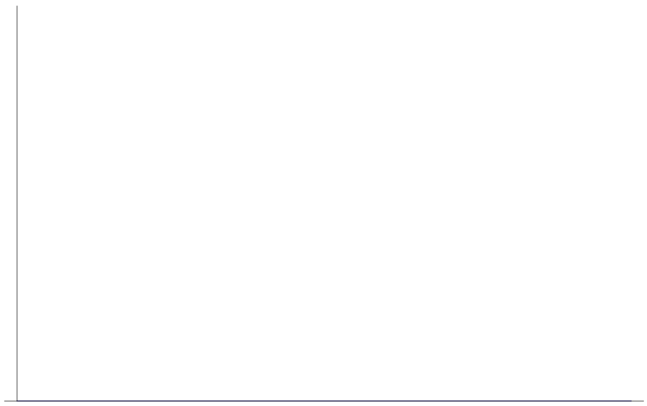


Figure 1.6: The differential estimates the actual change in y, Δy, due to a change in x: $x \to \Delta x$. For small changes in x, the differential is approximately equal to the actual change in y: $dy \approx \Delta y$.

For a differentiable function of two variables $z = f(x, y)$, we define the differentials dx and dy to be independent variables and the differential dz estimates the change in z when x changes to $x + \Delta x$ and y changes to $y + \Delta y$:

Example

If $z = f(x, y) = x^2 + 3xy - y^2$, find the differential dz. If x changes from 2 to 2.05 and y changes from 3 to 2.96, compute the values of dz and Δz (the actual change in z).

Solution:

Example

The pressure, volume and temperature of a mole of an ideal gas are related by the equation $PV = 8.31T$, where P is measured in kilopascals, V in litres and T in kelvins. Use differentials to find the approximate change in the pressure if the volume increases from 12 L to 12.3 L and the temperature decreases from 310 K to 305 K.

Solution:

Exercise: Compare this with the actual change

$$\Delta P = P(305, 12.3) - P(310, 12).$$

Propagation of Errors

Suppose we have a physical property P related to two other properties A and B by:

Now suppose we measure A and B and record values A_0 and B_0 with associated errors ΔA and ΔB. We can now keep track of the errors in P due to errors in A and B by knowing "*how much P will change due to small changes in A (and/ or B) between $A - \Delta A$ and $A + \Delta A$ (and $B - \Delta B$ and $B + \Delta B$)*". The differential of P gives an estimate of this:

Now we don't want errors to cancel each other out so we write:

Example

The base radius and height of a right circular cone are measured as 10 cm and 25 cm, respectively, with a possible error in measurement of as much as 0.1 cm in each. Use differentials to estimate the maximum error in the calculated volume of the cone.

Solution:

This procedure generalises in the obvious way.

Example

The dimensions of a rectangular box are measured to be $h = 75$ cm, $w = 60$ cm, and $l = 40$ cm, and each measurement is correct within 0.2 cm. Use differentials to estimate the largest possible error when the volume of the box is calculated from these measurements.

Solution:

Note that $V(75, 60, 40) = 180,000$ cm^3 so this error is of the order of 1%.

Example

The power P consumed in a resistor is given by

$$P = \frac{V^2}{R},$$

where V is the voltage and R is the resistance across the resistor.

(i) Use partial derivatives and differentials to determine an approximate expression for ΔP, the change in the power P.

(ii) Find the *approximate* change in P when V is changed by 5% and R is decreased by 0.5%.

[8 Marks]

Solution

(i) The differential dP approximates the change in P, ΔP in terms of dV and dR, the changes in V and R respectively:

Hence we write $P(V, R) = V^2 R^{-1}$ and calculate the partial derivatives:

(ii) To find the percentage change in P we look at $\dfrac{\Delta P}{P} \approx \dfrac{dP}{P}$. Hence we divide the differential by P:

Now we have that $\dfrac{dV}{V} = 0.05$ and $\dfrac{dR}{R} = 0.005$. So we have

The answer is 9.5%.

Winter 2012 Question 2 (c)

In calculating the area of a triangular plot of land, the formula

$$A = \frac{1}{2}ab\sin C,$$

is used. Estimate the value of A if a and b were measured as 25 m and 40 m, with maximum errors of 0.01 m and 0.02 m, respectively. The angle C is measured to be $\dfrac{\pi}{6}$ radians.

[5 Marks]

Remark: Reading this question, from a paper I did not set, I would just substitute $a = 25$ m, $b = 40$ m and $C = \dfrac{\pi}{6}$ although it makes sense to present the answer as $A \pm \Delta A$.

Solution: Our best guess of the area is just

Now we use

$$\Delta A \approx dA = \frac{\partial A}{\partial a}\Delta a + \frac{\partial A}{\partial b}\Delta b + \frac{\partial A}{\partial C}\Delta C.$$

where of course we have $\Delta C = 0$. We calculate

so we have

Summer 2012 Question 2(b)(ii)

In measuring a quantity u, the formula

$$u = \frac{2x}{3y}$$

was used. Estimate (a range of values) of u where the values of x and y were measured to be two and one with maximum errors of 0.06 and 0.03, respectively.

[5 Marks]

Solution: Our best guess of u is just

Now we use

$$\Delta u \approx du = \frac{\partial u}{\partial x}\Delta x + \frac{\partial u}{\partial y}\Delta y.$$

We calculate

so we have

Exercises

1. **Autumn 2012 Q. 2 (b)(ii)** In measuring the area of a triangular plot, the formula

$$A = \frac{1}{2}ab\sin C$$

   was used. Estimate a range of values of A where the values of a and b were measured to be 80 m and 40 m with maximum errors of 0.002 m and 0.001 m, respectively. The angle, $C = 30°$. Ans: 800 ± 0.04 m^2.

2. **Winter 2011 Q. 2(b)(ii)** In measuring a quantity u the formula

$$u = \frac{8x}{4x + y}$$

   was used. Estimate a range of values of u where the values of x and y were measured to be one and four with maximum errors of 0.04 and 0.08, respectively. Ans: 1 ± 0.03.

3. **Autumn 2010 Q. 2 (b)(ii)** Find the first order partial derivatives of V with respect to x and y where

$$V = \frac{2x + y}{2x - y}.$$

Estimate a range of values for V where the values of x and y were measured to be two and three with maximum errors of 0.02 and 0.03, respectively. Ans: 7 ± 0.48.

4. **Autumn 2009 Q. 2 (b)(ii)** In calculating a quantity V the formula

$$V = \frac{x^2}{3y}$$

was used. Estimate a range of values for V if x and y were measured as three and one, with maximum errors of 0.02 and 0.01, respectively. Ans: 3 ± 0.07.

5. Use differentials to estimate the amount of tin in a closed tin closed tin can with diameter 8 cm and height 12 cm if the can is 0.04 cm thick.

6. Use differentials to estimate the amount of metal in a closed cylindrical can that is 10 cm high and 4 cm is diameter if the metal in the wall is 0.05 cm thick and the metal in the top and bottom is 0.1 cm thick.

7. If R is the total resistance of three resistors, connected in parallel, with the resistances R_1, R_2 and R_3, then

$$\frac{1}{R} = \frac{1}{R_1} + \frac{1}{R_2} + \frac{1}{R_3}$$

. If the resistances are measured as $R_1 = 25\ \Omega$, $R_2 = 40\ \Omega$ and $R_3 = 50\ \Omega$, with possible errors of 5% in each case, estimate the maximum error in the calculated value of R.

8. The moment of inertia of a body about an axis is given by $I = kbD^3$ where k is a constant and B and D are the dimensions of the body. If B and D are measured as 2 m and 0.8 m respectively, and the measurement errors are 10 cm in B and 8 mm in D, determine the error in the calculated value of the moment of inertia using the measured values, in terms of k.

9. The volume, V, of a liquid of viscosity coefficient η delivered after a time t when passed through a tube of length l and diameter d by a pressure p is given by

$$V = \frac{pd^4t}{128\eta l}.$$

If the errors in V, p and l are 1%, 2% and 3% respectively, determine the error in η. HINT: If the error in A is x% then the error is $xA_0/100$ when $A = A_0$.

Selected Solutions:

1. Assuming that the measurements of 8 cm and 12 cm are taken from the outside of the can, then we could estimate the change in volume of a cylinder if the radius were increased by 0.04 cm to 4 cm and the height increased by 0.08 cm to 12 cm (convince yourself of this with a picture.). Now the tin in the can comprises the difference between a $(r, h) = (3.96, 11.92)$ cylinder and a $(r, h) = (3, 12)$ cylinder. Now the volume of a cylinder is given by

$$V = \pi r^2 h. \tag{1.4}$$

We can use the differential of V, dV (evaluated at $(r, h) = (3.96, 11.92)$ — although the other way around would also be a good estimate) to estimate the change in volume:

$$dv = \frac{\partial V}{\partial r}dr + \frac{\partial V}{\partial h}dh,$$

where $dr = +0.04$ cm and $dh = +0.08$ cm. Now

$$\frac{\partial V}{\partial r} = 2\pi rh\bigg|_{(r,h)=(3.96,11.92)} = 2\pi(3.96)(11.92)\text{ , and}$$

$$\frac{\partial V}{\partial h} = \pi r^2\bigg|_{(r,h)=(3.96,11.92)} = \pi(3.96)^2.$$

$$\Rightarrow dP = 2\pi(3.96)(11.92) \times (+0.04) + \pi(3.96^2) \times (+0.8) \approx 15.805 \text{ cm}^3.$$

2. Using the same method, the differential is a good estimate:

$$dv = \frac{\partial V}{\partial r}dr + \frac{\partial V}{\partial h}dh,$$

where in this case we take $(r, h) = (1.95, 9.8)$, and $dh = +0.2$ and $dr = 0.05$. Now

$$\frac{\partial V}{\partial h} = \pi r^2\bigg|_{(r,h)=(1.95,9.8)} = \pi(1.95)^2,$$

$$\frac{\partial V}{\partial r} = 2\pi rh\bigg|_{(r,h)=(1.95,9.8)} = 2\pi(1.95)(9.8).$$

Hence

$$dP = \pi(1.95^2) \times (+0.2) + 2\pi(1.95)(9.8) \times (0.05) \approx 8.393 \text{ cm}^3.$$

3. Now first we want to get an expression for $R(R_1, R_2, R_3)$:

$$\frac{1}{R} = \frac{1}{R_1}\cdot\frac{R_2R_3}{R_2R_3} + \frac{1}{R_2}\cdot\frac{R_1R_3}{R_1R_3} + \frac{1}{R_3}\cdot\frac{R_1R_2}{R_1R_2},$$

$$\Rightarrow R(R_1, R_2, R_3) = \frac{R_1R_2R_3}{R_1R_2 + R_1R_3 + R_2R_3}.$$

We can approximate the error in R, ΔR by:

$$\Delta R \approx dR = \left|\frac{\partial R}{\partial R_1}\right| dR_1 + \left|\frac{\partial R}{\partial R_2}\right| dR_2 + \left|\frac{\partial R}{\partial R_3}\right| dR_3,$$

where dR_i corresponds to the error in R_i, ΔR_i. Now the errors are 5 % hence:

$$dR_1 = (0.05)(25) = 1.25.$$

$$dR_2 = (0.05)(40) = 2.$$

$$dR_3 = (0.05)(50) = 2.5.$$

Also, using the quotient rule:

$$\begin{aligned}\frac{\partial R}{\partial R_1} &= \frac{(R_1R_2 + R_1R_3 + R_2R_3)(R_2R_3) - (R_1R_2R_3)(R_2 + R_3)}{[R_1R_2 + R_1R_3 + R_2R_3]^2}. \\ &= \frac{R_1R_2^2R_3 + R_1R_2R_3^2 + R_2^2R_3^2 - R_1R_2^2R_3 - R_1R_2R_3^2}{[R_1R_2 + R_2R_3 + R_1R_3]^2}.\end{aligned}$$

Similarly,

$$\frac{\partial R}{\partial R_2} = \frac{R_1^2R_3^2}{[R_1R_2 + R_1R_3 + R_2R_3]^2}.$$

$$\frac{\partial R}{\partial R_3} = \frac{R_1^2R_2^2}{[R_1R_2 + R_1R_3 + R_2R_3]^2}.$$

Note that these will be evaluated at $(R_1, R_2, R_3) = (25, 40, 50)$, using a calculator;

$$\frac{\partial R}{\partial R_1} = \frac{(40)^2(50)^2}{[(25)(40) + (25)(50) + (40)(50)]^2} = \frac{64}{289}.$$

Similarly,

$$\frac{\partial R}{\partial R_2} = \frac{25}{289}.$$

$$\frac{\partial R}{\partial R_3} = \frac{16}{289}.$$

Hence,

$$\Delta R \approx \frac{64}{289} \times (1.25) + \frac{25}{289} \times (2) + \frac{16}{289} \times (2.5) \approx 0.588\,\Omega.$$

4. In class we did an example where we estimated the change in I due to changes in B and D (from the sample test). The only difference between that example and this one is that errors are always positive and we take absolute values; i.e:

$$\Delta I = \left|\frac{\partial I}{\partial B}\right| \Delta B + \left|\frac{\partial I}{\partial D}\right| \Delta D.$$

5. First we solve for a function $\eta(V, P, l, d, t)$:

$$\eta = \frac{Pd^4t}{128Vl} = \frac{d^4t}{128} PV^{-1}l^{-1}.$$

Again we use the differential to estimate the error (as errors in d and t were not mentioned we will assume they don't have errors):

$$\Delta\eta \approx \left|\frac{\partial\eta}{\partial V}\right| \Delta V + \left|\frac{\partial\eta}{\partial P}\right| \Delta P + \left|\frac{\partial\eta}{\partial l}\right| \Delta l.$$

Now we must look at the partial derivatives (verify the last steps yourself):

$$\left|\frac{\partial\eta}{\partial V}\right| = \left|-\frac{d^4t}{128}PV^{-2}l^{-1}\right| = \frac{\eta}{V}.$$

$$\left|\frac{\partial\eta}{\partial P}\right| = \left|\frac{d^4t}{128}V^{-1}l^{-1}\right| = \frac{\eta}{P}.$$

$$\left|\frac{\partial\eta}{\partial l}\right| = \left|\frac{d^4t}{128}PV^{-1}l^{-2}\right| = \frac{\eta}{l}.$$

Now by the hint, the errors in V, P and l $\Delta V = \frac{V}{100}$, $\Delta P = \frac{2P}{100}$, and $\Delta l = \frac{3l}{100}$. Hence,

$$\begin{aligned}\Delta\eta &= \frac{\eta}{V}\times\frac{V}{100} + \frac{\eta}{P}\times\frac{2P}{100} + \frac{\eta}{l}\times\frac{3l}{100} \\ &= \frac{6}{100}\eta.\end{aligned}$$

That is the error in η is 6 %.

1.4 Multivariable Taylor Series

Winter 2012 Question 2 (b)

Consider the function $f(x, y) = \ln|x^2 - 4y|$.

(i) Find all the first and second order partial derivatives of $f(x, y)$.

(ii) Hence, find a Taylor Series expansion of $f(x, y)$ about $x = 3$, $y = 2$.

[7 & 3 Marks]

Solution:

(i) Well we have

Off we go using the Chain Rule:

Now we differentiate these using the Quotient and Chain Rules[3]

[3]if you don't see $f_y = -4(x^2 - 4y)^{-1}$ and $f_x = 2x(x^2 - 4y)^{-1}$ (looking for f_{xy}) just use the quotient rule

(ii) We know that

$$f(x,y) \approx f(a,b) + (x-a)f_x + (y-b)f_y + \frac{(x-a)^2}{2!}f_{xx} + (x-a)(y-b)f_{xy} + \frac{(y-b)^2}{2!}f_{yy}.$$

We calculate the numbers $f(3,2)$, $f_x(3,2)$, $f_y(3,2)$, $f_{xx}(3,2)$, $f_{yy}(3,2)$ and $f_{x,y}(3,2)$:

So we have

Chapter 2

Differential Equations

What is the origin of the urge, the fascination that drives physicists, mathematicians, and presumably other scientists as well? Psychoanalysis suggests that it is sexual curiosity. You start by asking where little babies come from, one thing leads to another, and you find yourself preparing nitroglycerine or solving differential equations. This explanation is somewhat irritating, and therefore probably basically correct.

David Ruelle

2.1 Review of Separable First Order Differential Equations

A differential equation is an equation containing one or more derivatives, e.g.,

$$y' = x^2,$$
$$\frac{d^2y}{dx^2} + x\frac{dy}{dx} = \sin x.$$

Most laws in physics and engineering are differential equations.

The *order* of a differential equation is the order of the highest derivative that appears;

$$y' = x^2 \quad \text{is first order}$$
$$\frac{d^2y}{dx^2} + \left(\frac{dy}{dx}\right)^3 = 1 \quad \text{is second order}$$
$$(\cos x)\left(\frac{d^2y}{dx^2}\right)^3 + \frac{dy}{dx} = y \quad \text{is second order.}$$

A function $y = f(x)$ is a *solution* of a differential equation if when you substitute y and its derivatives into the differential equation, the differential equation is satisfied. For example, $y = \tan x$ is a solution of the differential equation $y' = 1 + y^2$ since if $y = \tan x$, then

A differential equation can have many solutions: $y' = 2$ has a solution $y = 2x+C$ for every constant C. The *general solution* of a differential equation is the set of all possible solutions. The differential equation $y' = 2$ has general solution $y = 2x + C$, where the constant C is arbitrary. It can be very difficult to find the general solution of a differential equation. We shall consider only certain first-order differential equations that can be solved fairly readily.

A *separable* first-order differential equation is one that can be written in the form

In this situation we can *separate the variables*:

Each side can now be integrated:

The point of separating the variables is that we cannot usually integrate expressions like $\int y\,dx$ where both variables appear.

Example

Solve the separable first order differential equation:

$$y' = xy.$$

Solution: First separate the variables and integrate:

Usually we want to solve for y:

Here there are two infinite families of solutions. The solution of a first-order differential equation will always contain an unknown constant — and might have different families of solutions also (e.g. the solution $y^2 = x + C$ has the families $y = +\sqrt{x + C}$ and $-\sqrt{x + C}$). However an extra piece of numerical data such as "$y = 2$ when $x = 1$" sometimes reduces this to a unique solution. Note that this will usually be written as $y(1) = 2$ — for the input $x = 1$, the output is $y = 2$. This extra data is called an *initial condition* or *boundary condition* and the entire problem (differential equation and boundary condition) is often called an *initial-value problem* or *boundary-value problem.*

Example

Solve the initial-value problem

$$\frac{dy}{dx} = \frac{1+x}{xy} \quad \text{for } x > 0, \quad \text{where } y(1) = -4.$$

Solution: First separate the variables and integrate:

Now apply the boundary condition:

Now substitute in the constant and hopefully solve for $y(x)$:

Now the fact that $y = -4$ at $x = 1$ and that $\sqrt{x} > 0$ where defined implies that the solution is $y(x) = -\sqrt{2(\log_e x + x + 7)}$. [Ex:] Show that this solves the differential equation and satisfies the boundary condition.

Further Remarks: Picard's Existence Theorem

There is a theorem in the analysis of differential equations which states that if a differential equation is suitably *nice* in an interval about the boundary condition then not only does a solution exist but it is unique. This allows us to define functions as solutions to differential equations. For example, an alternate definition of the exponential function, e^x, is the unique solution to the differential equation:

$$\frac{dy}{dx} = y\,,\ y(0) = 1.$$

Exercises

1. Solve the following differential equations:

   (a) $y' = 3x^2 + 2x - 7$ Ans: $y = x^3 + x^2 - 7x + C$

(b) $y' = 3xy^2$ Ans: $3x^2y + Cy + 2 = 0$

(c) $\dfrac{dy}{dx} = \dfrac{3x\sqrt{1+y^2}}{y}$ Ans: $2\sqrt{1+y^2} = 3x^2 + C$

(d) $\dfrac{dy}{dx} = \dfrac{x}{4y}$, $y(4) = -2$ Ans: $x^2 = 4y^2$

2. The point $(3, 2)$ is on a curve, and at any point (x, y) on the curve the tangent line has slope $2x - 3$. Find the equation of the curve. Ans: $y = x^2 - 3x + 2$

3. The slope of the tangent line to a curve at any point (x, y) on the curve is equal to $3x^2y^2$. Find the equation of the curve, given that the point $(2, 1)$ lies on the curve. Ans: $-\dfrac{1}{y} = x^3 - 9$

2.2 Numerical Solution of First Order Differential Equations

2.2.1 Direction Fields

Unfortunately, it's impossible to solve most differential equations in the sense of obtaining an explicit formula for the solution. In this section, we show that, despite the absence of an explicit solution, we can still learn a lot about the solution through a graphical approach (direction fields) or a numerical approach (Euler's Method and the Three-Term-Taylor Method).

Suppose we are asked to sketch the graph of the solution of the initial value problem:

$$\frac{dy}{dx} = x + y\ ,\ \ y(0) = 1.$$

We don't know a formula for the solution, so how can we possibly sketch its graph? Let's think about what the differential equation means. The equation $y' = x + y$ tells us that the slope at any point (x, y) on the graph of $y(x)$ is equal to the sum of the x- and y-coordinates at that point. In particular, because the curve passes through the point $(0, 1)$, its slope there must be $0 + 1 = 1$. So a small portion of the solution curve near the point $(0, 1)$ looks like a short line segment through $(0, 1)$ with slope 1:

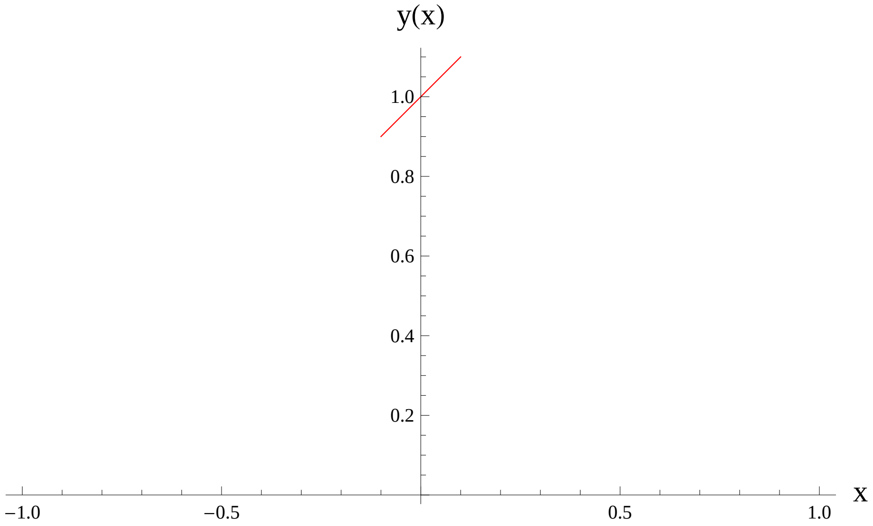


Figure 2.1: Near the point $(0, 1)$, the slope of the solution curve is 1.

As a guide to sketching the rest of the curve, let's draw short line segments at a number of points (x, y) with slope $x + y$. The result is called a *direction field* and is shown below:

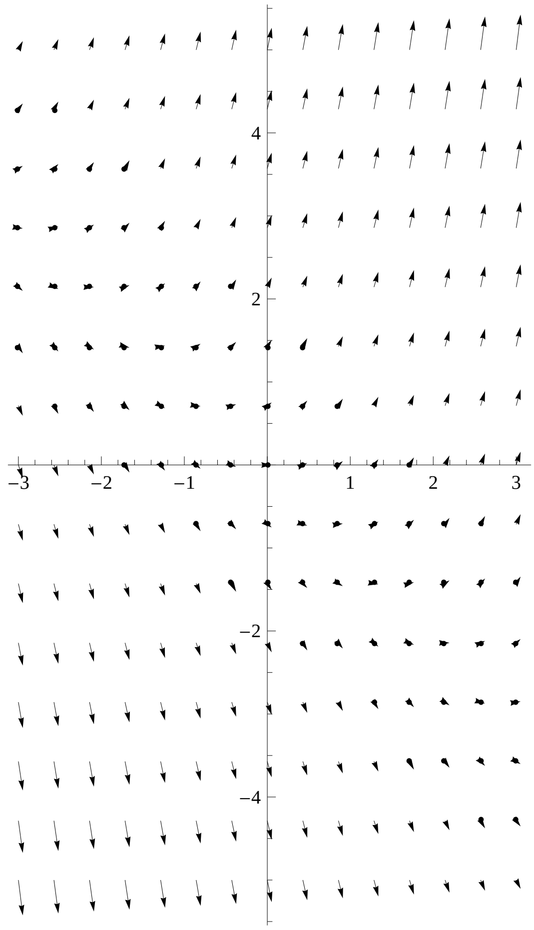


Figure 2.2: For example, the line segment at the point $(1, 2)$ has slope $1 + 2 = 3$. The direction field allows us to visualise the general shape of the solution by indicating the direction in which the curve proceeds at each point.

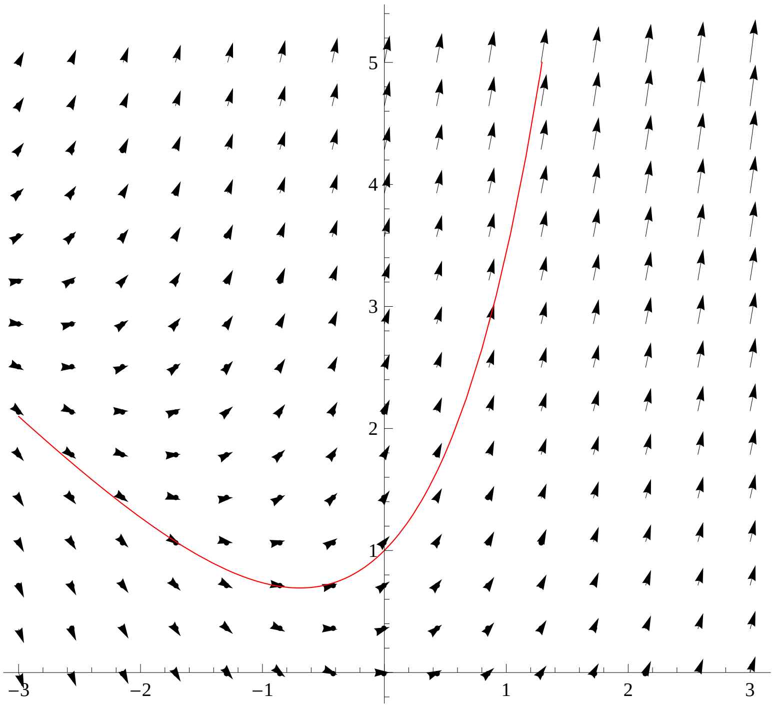


Figure 2.3: We can sketch the solution curve through the point $(0, 1)$ by following the direction field. Notice that we have drawn the curve so that it is parallel to nearby line segments.

2.2.2 Euler's Method

The basic idea behind direction fields can be used to find numerical approximations to solutions of differential equations. We illustrate the methods on the initial-value problem that we used to introduce direction fields:

$$\frac{dy}{dx} = x + y \ , \ \ y(0) = 1.$$

The differential equation tells us that $y'(0) = 0 + 1 = 1$, so the solution curve has slope 1 at the point $(0, 1)$. As a first approximation to the solution we could use the linear approximation $L(x) = 1x + 1$. In other words we could use the tangent line at $(0, 1)$ as a rough approximation to the solution curve.

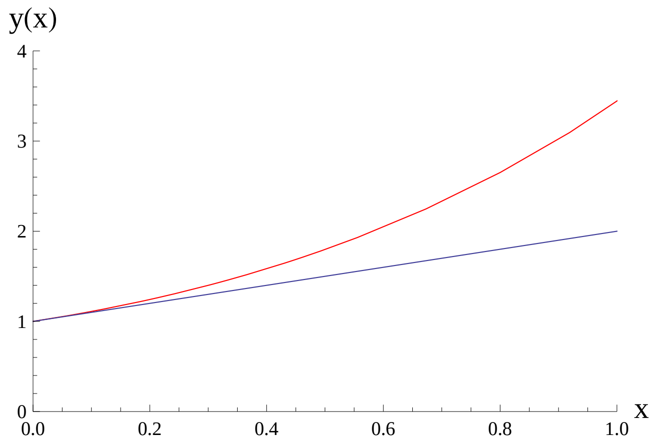


Figure 2.4: The tangent at $(0, 1)$ approximates the solution curve for values near $x = 0$.

Euler's idea was to improve on this approximation by proceeding only a short distance along this tangent line and then making a correction by changing direction according to the direction field:

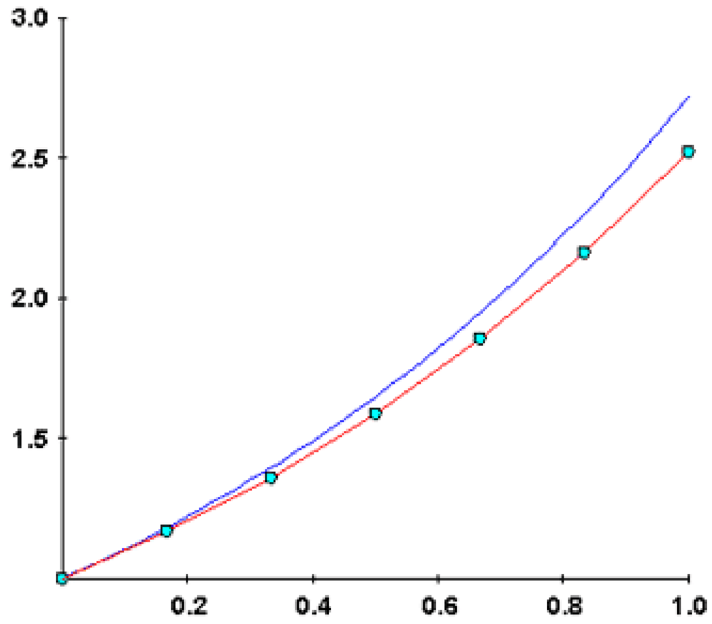


Figure 2.5: Euler's Method starts at some initial point (here $(x_0, y_0) = (0, 1)$), and proceeds for a distance h (in this plot $h = 1/6$.) at a slope that is equal to the slope at that point $y' = x_0 + y_0$. At the point $(x_1, y_1) = (x_0 + h, y_1)$, the slope is changed to what it is at (x_1, y_1), namely $x_1 + y + 1$, and proceeds for another distance h until it changes direction again.

Euler's method says to start at the point given by the initial value and proceed in the direction indicated by the direction field. Stop after a short time, look at the slope at the new location, and proceed in that direction. Keep stopping and changing direction according to the direction field. Euler's method does not produce an exact solution to the initial-value problem — it gives approximations. But by decreasing the step size (and therefore increasing the amount of corrections), we obtain successively better approximations to the correct solution.

For the general first-order initial-value problem $y' = F(x, y)$, $y(x_0) = y_0$, our aim is to find approximate values for the solution at equally spaced numbers x_0, $x_1 = x_0 + h$,$x_2 = x_0 + 2h = x_1 + h$..., where h is the step size. The differential equation tells us that the slope at (x_0, y_0) is $y' = F(x_0, y_0)$:

This shows us that the approximate value of the solution when $x = x_1$ is

$$y_1 = y_0 + hF(x_0, y_0)$$

Similarly,

$$y_1 = y_0 + hF(x_0, y_0)$$

Euler's Method

If

$$\frac{dy}{dx} = F(x, y) \ , \ \ y(x_0) = y_0$$

is an initial value problem. If we are using Euler's method with step size h then

$$y(x_{n+1}) \approx y_{n+1} = y_n + hF(x_n, y_n) \tag{2.1}$$

for $n \geq 0$.

Example

Use Euler's method with step size $h = 0.1$ *to approximate* $y(1)$*, where* $y(1)$ *is the solution of the initial value problem:*

$$\frac{dy}{dx} = x + y \ , \ \ y(0) = 1$$

Solution: We are given that $h = 0.1$, $x_0 = 0$ and $y_0 = 1$, and $F(x, y) = x + y$. So we have

$$\begin{aligned} y_1 &= y_0 + F(x_0, y_0) = 1 + 0.1(0 + 1) = 1.1 \\ y_2 &= y_1 + F(x_1, y_1) = 1.1 + 0.1(0.1 + 1.1) = 1.22 \\ y_3 &= y_2 + F(x_2, y_2) = 1.22 + 0.1(0.2 + 1.22) = 1.362 \end{aligned}$$

Continue this process [Exercise] to get $y_{10} = 3.187485$, which approximates $y(x_{10}) = y(x_0 + 10(0.1)) = y(1)$, as required.

Winter 2012 Question 3 (c)

By using Euler' method with a step $h = 0.1$, estimate the value of y at $x = 0.2$ where

$$\frac{dy}{dx} = x^2 + y^2 \ , \ \ x(0) = 1.$$

[5 Marks]

Solution: We have

so

and

2.3 Three Term Taylor Method

The Three-Term-Taylor Method is a method of approximating the solutions to ordinary differential equations. Given a differential equation, we will be given the initial/boundary condition that $y(x_0) = y_0$. Suppose that we assume that our differential equation has an *analytic solution* at x_0. Then we know that for points *close* to x_0, we have

When we are particularly close to x_0, in particular $|x - x_0| < 1$, then the higher powers of $(x - x_0)$ in the Taylor Series expansion will be getting smaller and smaller and if we want we just take the first three terms as an approximation:

Usually what we want to do here is find the solution at multiples of a step size h here so what we usually do is write $x - x_0 := h$ and so

Given a differential equation

can we find an approximate value of the solution at $x = x_0 + h$? Well we could use the Taylor Series about $x = x_0$

as x is near x_0. Why? Well we know $y(x_0) = y_0$. We can find the value of $y'(x_0)$ because we know that $y'(x_0) = F(x_0, y_0)$. Do we know what $y''(x_0)$ is? Well $y''(x)$ is the derivative with respect to x of $y'(x)$ — morryah the derivative of $F(x, y)$ with respect to x. We assume that y depends on x so we write $y'(x) = F(x, y(x))$ and differentiate implicitly. Therefore we may write

$$y(x_0 + h) \approx y(x_0) + y'(x_0)h + \frac{y''(x_0)}{2}h^2$$

Like Euler's method we can compute an approximate solution from $y(x_0) = y_0$ using step size h using the recursion:

$$y_{k+1} \approx y_k + y'_k h + y''_k \frac{h^2}{2} \tag{2.2}$$

where the notation will make sense when we do an example.

Autumn 2012

Use the Three Term Taylor Method with a step $h = 0.1$ to estimate the value of y at $x = 1.2$ where

$$\frac{dy}{dx} = x + y \text{ , and } y(1) = 0.$$

[NOTE: $y_{k+1} \approx y_k + y'_k h + y''_k \frac{h^2}{2}$]

[4 Marks]

Solution: Note that we have $x_0 = 1$, $y_0 = 0$, $y_0' = x_0 + y_0 = 1 + 0 = 1$. All that remains is to calculate y_0''. First write $y' = F(x, y(x))$:

Now we may estimate $y_1 := y(1.1)$

Now we estimate $y_2 := y(1.2)$. We have $x_1 = 1.1$, $y_1 = 0.11$, $y_1' = x_1 + y_1 = 1.1 + 0.11 = 1.21$. All that remains is to calculate y_1''. We know that $y'' = 1 + y'$ from the first step so we have

1. **Winter 2012 Q. 1 (d)** Use the Three Term Taylor Method with a step of $h = 0.1$ to estimate the value of y at $x = 0.2$ where

   $$\frac{dy}{dx} + 2y(x) = 20 \text{ , and } y(0) = 5.$$

   The Solution of the differential equation above is given by

   $$y(x) = 10 - 5e^{-2x}.$$

   Calculate the error in the approximation above.
   [NOTE: $y_{k+1} = y_k + y_k' h + \frac{y_k''}{2} h^2$.] Ans: 6.638 and 0.0104

2. **Winter 2010 Q. 1 (b) (iii)** Use the Three Taylor Method with a step of $h = 0.1$ to estimate the value of y at $x = 1.2$ where

   $$\frac{dy}{dx} = 2y(x) \text{ , and } y(1) = 2.$$

   [NOTE: $y_{k+1} = y_k + y_k' h + \frac{y_k''}{2} h^2$.] Ans: 2.9768

3. **Autumn 2009 Q. 1(b)** Use the Three Term Taylor Method with a step of $h = 0.1$ to estimate the value of y at $x = 0.2$ where

   $$\frac{dy}{dx} = x^2 + y(x)^2 \text{ , and } y(0) = 1.$$

   [NOTE: $y_{k+1} = y_k + y_k' h + \frac{y_k''}{2} h^2$.] Ans: 1.249

4.

5. **Winter 2009 Q. 1 (c)** Use the Three Term Taylor Method with a step of $h = 0.1$ to estimate the value of y at $x = 0.1$ where

$$\frac{dy}{dx} + 4y(x) = 40 \text{ , and } y(0) = 3.$$

The Solution of the differential equation above is given by

$$y(x) = 10 - 7e^{-4x}.$$

Calculate the error in this approximation.
[NOTE: $y_{k+1} = y_k + y'_k h + \frac{y''_k}{2} h^2$.] Ans: 5.24 and 0.068

Exercises

1. *Use Euler's method with step size 0.5 to compute the approximate y-values* y_1, y_2, y_3 *and* y_4 *of the initial value problem* $y' = y - 2x$, $y(1) = 0$.

2. *Use Euler's method with step size to estimate* $y(1)$, *where* $y(x)$ *is the solution of the initial value problem* $y' = 1 - xy$, $y(0) = 0$.

3. *Use Euler's method with step size* 0.1 *to estimate* $y(0.5)$, *where* $y(x)$ *is the solution of the initial value problem* $y' - y = xy$, $y(0) = 1$.

4. *Use Euler's method with step size* 0.2 *to estimate* $y(1.4)$, *where* $y(x)$ *is the solution of the initial-value problem* $y' - x + xy = 0$, $y(1) = 0$.

2.3.1 The Three Term Taylor Method

2.4 Second Order Linear Differential Equations

There is a certain class of second order differential equations which yield readily to deep analysis. The basic form of these equations is

where $a, b, c \in \mathbb{R}$ are constants and $\phi(x)$ is a smooth function of x. We call the differential equation *homogenous* when $\phi(x) = 0$. For now we will just look at homogenous second order linear differential equation. The heavy analysis can be neatly summarised by the heuristic that a first order linear differential equation

has one solution $y(x)$ which because we are going from derivative to function (integrating) has a constant of integration C

$$y(x) = y(x, C) \sim Cy(x).$$

Now when we have a *second* order differential equation there are *integrations* somehow

$$\frac{d^2y}{dx^2} \xrightarrow{\int dx} \frac{dy}{dx} \xrightarrow{\int dx} y(x),$$

and thus *two* constants of integration, say C_1 and C_2 so we have that the solution depends on x, C_1 and C_2 and some heavy lifting shows that in fact there are *two* linearly independent solutions y_1 and y_2.

$$y(x) = y(x, C_1, C_2) \sim C_1y_1(x) + C_2y_2(x).$$

Linearly independent means that neither y_1 nor y_2 is not zero nor a multiple of the other. Harder work again shows that *all* of the solutions look like $C_1y_1(x) + C_2y_2(x)$. This gives us a strategy for solving second order differential equations... find two solutions and you are done!

Sounds simple doesn't it?! Well as it turned out, someone copped that exponential functions $y_i(x) = Ae^{ax}$ make good candidate solutions as when you differentiate them they stay as multiples of e^{ax}...

Winter 2012 Question 3 (a) & (b)

(a) Find the general solution of the differential equation

$$\frac{d^2y}{dt^2} + 3\frac{dy}{dt} + 2y(t) = 10 \sin t.$$

[7 Marks]

Solution: First we examine the homogenous system

To solve this we first find the roots of the *auxiliary equation*:

Now we must find a particular solution $y_P(t)$. As the right-hand-side is sinucoidal we will try the Ansatz $y_T(t) = C\sin t + D\cos t$. Now we look at the left-hand-side and calculate and compare

Therefore we want $C - 3D = 10$ and $D + 3C = 0$:

Hence the general solution is given by

$$y_G(t) = Ae^{-2t} + Be^{-t} + \sin t - 3\cos t.$$

(b) Solve the differential equation

$$\frac{d^2x}{dt^2} - 4\frac{dx}{dt} + 4x(t) = 16t,\ \ x(0) = 1, x'(0) = 0.$$

[8 Marks]

Solution: First we look at the homogenous system and hence the auxiliary equation

Now we must find a particular solution $x_P(t)$. As the right-hand-side is a linear polynomial we will try the Ansatz $x_T(t) = Ct + D$. Now we look at the left-hand-side and calculate and compare

Therefore we want $4C = 16$ and $4D - 4C = 0$:

Hence the general solution is given by

$$x_G(t) = Ae^{2t} + Bte^{2t} + 4t + 4.$$

Now using the initial conditions we can find A and B. First $x(0) = 1$:

Now we calculate $x'(0)$ and set it equal to 0 to hopefully find B:

Hence we have the solution

$$x(t) = -3e^{2t} + 2te^{2t} + 4t + 4.$$

2.5 Impulse & Step Functions

2.5.1 Introduction

In this section we examine *impulse* and *step functions.* They arise naturally in the theory of beams. They model various *discontinuous* phenomena.

2.5.2 Step Functions

For example, consider a simple switch that has two states: on (1) and off (0). Suppose the switch is turned on at time $t = 0$. We use the *Heaviside Function* to model this:

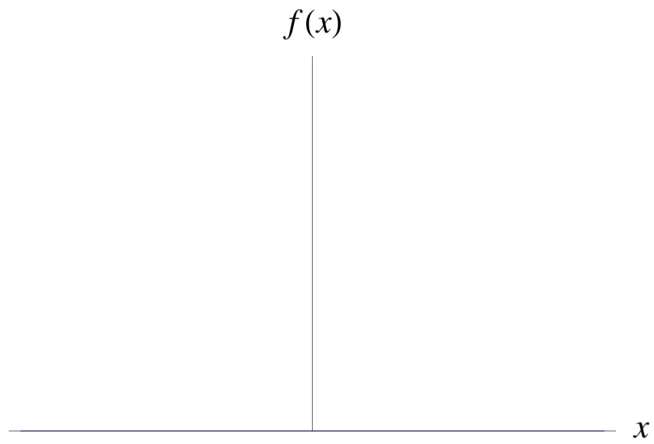


Figure 2.6: The graph of the Heaviside Function.

It is no problem writing down a switch which starts at a time $x = a$:

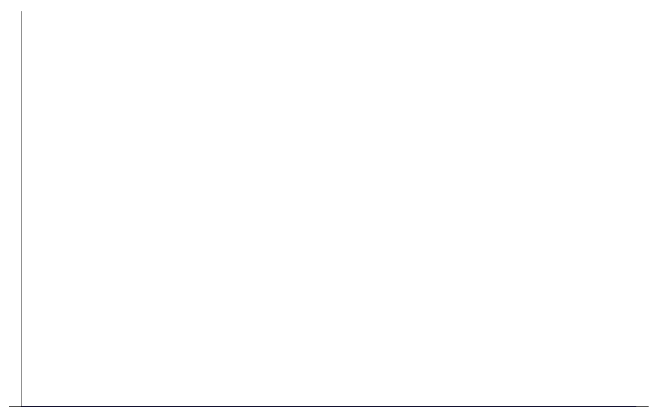


Figure 2.7: The graph of the Heaviside Function $H(x - a)$.

We can also use a combination of Heaviside functions to write an expression for a switch that is on between times $x = a$ and $x = b$:

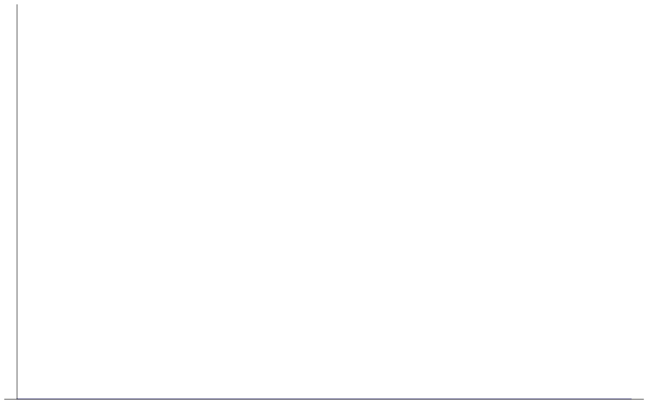


Figure 2.8: This function is equal to $H(x - a) - H(x - b)$. It is an exercise to show that this function is equal to one for $a < x < b$ and zero elsewhere.

Of course the output value doesn't have to be one, it could be 3 say; or we could have a switch equal to 4 between $x = 2$ and $x = 5$:

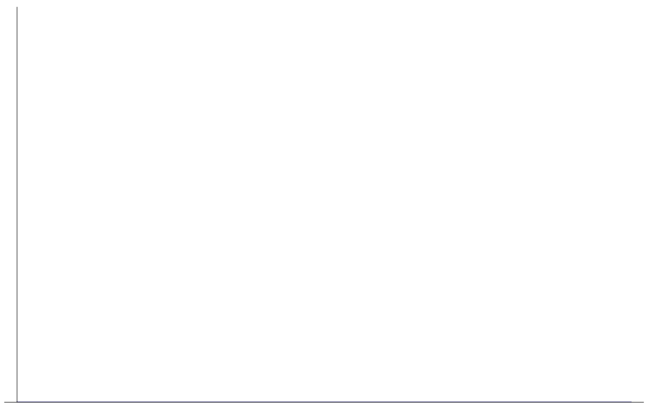


Figure 2.9: The graphs of $3H(x)$ and $4H(x - 2) - 4H(x - 5)$.

Examples

Write down an formula for the following step functions:

1. equal to 7 for $x > -2$ and zero otherwise.
2. equal to 1/2 for $-1 < x < 1$ and zero otherwise.

Solution: 1. $7H(x + 2)$ 2. $\frac{1}{2}H(x + 1) - \frac{1}{2}H(x - 1)$.

What is the derivative of $H(x)$? The slope is 0 except at $x = 0$ where it is infinite. This is the *Dirac Delta Function*:

$$\delta(x) = \begin{cases} \infty & \text{if } x = 0 \\ 0 & \text{otherwise} \end{cases} \tag{2.3}$$

There is a more precise definition that can be done such that the integral of $\delta(x)$ is $H(x)$. Using this function we can model a 'jolt' acting at the time $t = a$ or a point force acting at position $x = a$ of magnitude F is given by $F\delta(x - a)$.

We can also integrate the Heaviside function (this is what we will need to do). First some notation:

2.5.3 Notation

$$[x] := xH(x). \tag{2.4}$$

Note now that $[f(x)]$ is a function whose value if $f(x)$ when $f(x) \geq 0$ and zero otherwise. In essence, $[x]$ is the function which 'chops off' the negative part of the graph of $f(x)$:

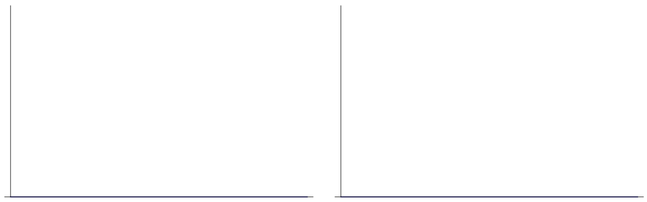


Figure 2.10: The graph of a function $f(x)$ and $[f(x)]$.

Examples

Write down formulas for the following functions:

1. Equal to x^2 for $x > 0$ and zero otherwise.
2. Equal to $x - 4$ for $x > 4$ and zero otherwise.
3. Equal to x for $2 < x < 5$ and zero otherwise.

Solution: 1. $[x^2] = x^2H(x^2)$ 2. $[x - 4] = (x - 4)H(x - 4)$ 3. $[x - 2] - [x - 4]$.

2.5.4 Integration of $H(x)$

The integral of $H(x)$ is equal to $[x] = xH(x)$.

Proof. We have that

Integrating both sides gives the answer.

2.5.5 Integral of $[x - a]$

The integral of $[x - a]$ is equal to $\dfrac{[x - a]^2}{2}$.

Proof. Now

$$[x-a] = \begin{cases} x-a & \text{if } x > a \\ 0 & \text{otherwise} \end{cases}$$

Hence

But note that

$$\frac{[x-a]^2}{2} = \begin{cases} \frac{x^2-2ax+a^2}{2} & \text{if } x > a \\ 0 & \text{otherwise} \end{cases}$$

which is equal to $\int [x-a]\, dx$.

Similarly we can show that

$$\int [x-a]^n \, dx = \frac{[x-a]^{n+1}}{n+1} + C. \tag{2.5}$$

Examples

Solve the following differential equations:

$$\frac{dy}{dx} = 2H(x) - 2H(x-1). \tag{2.6}$$

$$\frac{d^2y}{dx^2} = 3H(x-2) + \delta(x-1). \tag{2.7}$$

$$\frac{d^2y}{dx^2} = [x-2]^2 - [x-4]^2 \tag{2.8}$$

$$\frac{d^3y}{dx^3} = H[x-2]. \tag{2.9}$$

$$\frac{d^4y}{dx^4} = H[x]. \tag{2.10}$$

Solution On board.

Winter 2012 Question 1 (b)

Solve the differential equation

$$EI\frac{d^2y}{dx^2} = -24[x-3] + Rx,$$

where $[x-3]$ is a step function and R is a constant. At the points $x = 1$ m and $x = 4$ m the deflection y is zero. Also, a the point $x = 1$ m the slope of y is zero.

[8 Marks]

Solution: To solve this for $y(x)$ we must in the first instance integrate twice:

Now we have three constants[1] R, C_1 and C_2 to find but luckily we have three equations to find

[1]you will know from context that EI are properties of the material and dimensions of a beam so can be left as

these:

We apply these:

to give us three equations in three unknowns that can be written as:

$$\frac{1}{6}R + C_1 + C_2 = 0 \ \ (*)$$
$$\frac{32}{3}R + 4C_1 + C_2 = 4 \ \ (**)$$
$$\frac{1}{2}R + C_1 = 0 \ \ (***)$$

There are various methods of solving this system of equations[2] and however you do it doesn't really matter as long as you get it correct. What I would do is note that I can write C_1 in terms of R using (***), and hence C_2 in terms of R using (*) and put everything in terms of R in (**) to find R:

Now we can write down that $C_1 = -\frac{2}{9}$ and $C_2 = \frac{4}{27}$ so we have a final answer of

$$y(x) = \frac{1}{EI}\left(-4[x-3]^3 + \frac{4}{54}x^3 - \frac{2}{9}x + \frac{4}{27}\right)$$

they are

[2]the best of which you will learn about in MATH7021 if you pick that elective module

Winter 2010

To find the deflection y at any point on a beam the differential equation below must be solved where $[x-4]$ is a step function and where R is a constant

$$EI\frac{d^2y}{dx^2} = -15[x-4]^2 + Rx.$$

Solve this differential equation where the deflection y and the slope of y are zero at the point $x = 1$. Also at the point $x = 6$ the deflection is zero.
Solution: We start with two integrations:

Now we apply the boundary conditions. First $y(1) = 0$:

Now $y'(1) = 0$:

And finally $y(6) = 0$:

Hence to find R, C_1 and C_2 we must solve the linear system:

$$\begin{aligned} 36R + 6C_1 + C_2 &= 0 \\ \frac{R}{2} + C_1 &= 0 \\ \frac{R}{6} + C_1 + C_2 &= 0 \end{aligned}$$

We can do this either using matrix methods or a substitution method.

Exercises: Find the general solution of the following second order separable differential equations for $M(x)$:

$$\begin{aligned}
\frac{d^2M}{dx^2} &= -\delta(x-1) - \delta(x-5) \\
\frac{d^2M}{dx^2} &= -18 \\
\frac{d^2M}{dx^2} &= 72H(x-4) - 72H(x-1) \\
\frac{d^2M}{dx^2} &= 36H(x-5) - 36H(x-1) - 5\delta(x-4) \\
\frac{d^2M}{dx^2} &= -\delta(x-1) - \delta(x-3) - 144H(x-5) + 144H(x-8) - 3\delta(x-7) \\
\frac{d^2M}{dx^2} &= -x - 2 - 10\delta(x-2) \\
\frac{d^2M}{dx^2} &= -3x - 3 - 72H(x-2) \\
\frac{d^2M}{dx^2} &= -6x - 4 - 72H(x-4) + 72H(x-1) \\
EI\frac{d^2y}{dx^2} &= -144[x-2]^2
\end{aligned}$$

Now find the particular solutions to the first three differential equations under the following initial conditions.

1. $M'(0) = 1$ and $M(0) = 0$.
2. $M'(0) = 54$ and $M(0) = 0$.
3. $M'(0) = 108$ and $M(0) = 0$.

Note after we do the next section we can look back and see that these are the beam equations for:

1. A simply supported beam of length 6 m with point loads of magnitude 1 kN at $x = 1$ and $x = 5$.
2. A simply supported beam of length 6 m with a uniform load of 18 kN m^{-1}.
3. A simply supported beam of length 5 m with a uniform load of 72 kN m^{-1} between $x = 1$ and $x = 4$.

2.6 Applications to Beams & Beam Struts

In this section we learn how to formulate and solve beam equations so that we may calculate/estimate the deflection of a beam due to the loads on it. This theory known variously as the *Euler-Bernoulli Beam Theory*, *Engineer's Beam Theory* or *Classical Beam Theory*. It is a simplification of the theory of elasticity which, after it was used in the design of the Eiffel Tower and Ferris wheels, became a cornerstone of engineering. The theory makes a number of underlying assumptions. First we must show a beam in a deformed state and an undeformed state:

The beams for which we apply the model are assumed to be:

1. slender: their length is much greater than their width
2. isotropic and homogeneous: the material behaves the same at all points in the beam and in all directions
3. constant cross-section

The assumptions of the model (Kirchhoff's Assumptions) are:

1. Normals remain straight: they do not bend
2. Normals remain unstretched: they do not change length
3. Normals remain normal: they stay perpendicular to the neutral plane

Now using these assumptions we can derive an equation relating the *deflection* $y(x)$ at a point x and the *load* $w(x)$. The deflection is given as follows:

This equation (whose derivation is outside the remit of our course) is given by:

where E is Young's Modulus and I is the moment of inertia of the beam. Note that in all cases we use kilo-Newtons (kN) rather than Newtons (N). Solving this equation requires *four* integrations.

This analysis can be broken down somewhat if we look at the *bending moment* of a load. The bending moment is a *moment* due to a vertical loading. The moment of a force is, about an axis O, at a point P, the product of the magnitude of the force in the direction parallel to the axis, F, times the distance from the axis L:

The bending moment and the shearing force at a distance x from the left-hand side of a beam are related to the load per unit length $w(x)$ and the shearing force $V(X)$ by the differential equations

$$\frac{dM}{dx} = V(x), \quad \frac{dV}{dx} = -w(x).$$

Definition

The *shearing force* at the point x is the sum of the forces at or to the left of the point x.

Ye probably know more about shearing than me and we just need this working definition. These can be combined to form the second order differential equation

$$\frac{d^2M}{dx^2} = -w. \tag{2.11}$$

Hence if we know the load we can calculate the bending moment, $M(x)$. Then we can appeal to the master equation to solve for the deflection.

We have three (five) main types of loads:

1. *Uniformly distributed loads, U.D.L. over the entire beam*

   Mathematically we have

$$w_{\text{UDL}}(x) = w \tag{2.12}$$

   where w is the value of the UDL.

Example

Find the bending moment due to a U.D.L. of 36 kN m^{-1} across a beam of length 7 m where $M(0) = 0$ and $M'(0) = 126$.
Solution: First we integrate twice:

Now we apply the boundary conditions. First $M(0) = 0$:

$$M(0) = 0 = C_2.$$

Now $M'(0) = 126$:

$$M'(0) = C_1 = 126.$$

So the answer is

$$M(x) = -18x^2 + 126x.$$

2. *U.D.L. over a segment of the beam.*

Mathematically we have, for a UDL from $x = a$ to $x = b$:

$$w_{\text{UDL }:x=a\to b}(x) = wH(x-a) - wH(x-b) \tag{2.13}$$

where w is the UDL across $a \leq x \leq b$.

Example

Write down the load due to a UDL of 18 kN m^{-1} from $x = 2$ to $x = 3$ on a beam of length 5 m. Now find the bending moment due to this load. Write your answer in terms of M_A, the bending moment at $x = 0$, and R_A the reaction/shearing force at $x = 0$.
Solution: First draw at a picture:

Now write down the bending moment equation and integrate twice:

Now apply the boundary conditions taking into account that $M'(x)$ is the shearing force:

Therefore the answer is

$$M(x) = -9[x-2]^2 + 9[x-3]^2 + R_A x + M_A.$$

3. *Point loads*

Mathematically we have, for a point load of weight w at a point $x = a$:

$$w(x) = w\delta(x-a) \tag{2.14}$$

Example

Write down the load due to a point mass of 13 kN at $x = 3$ on a beam of length 6 m. Now find the bending moment.
Solution: We simply have $w(x) = 13\,\delta(x-3)$ so we have the bending moment equation:

Although this example doesn't ask for it, we can nearly always give our answer in terms of R_A and M_A:

So our best answer is

$$M(x) = -13[x-3] + R_A x + M_A$$

4. *Linear loads over the entire beam*

Here we have $w(0) = w_A$ and $w(L) = W_B$. It is not hard to show that a line has equation

$$\text{OUTPUT} = \text{SLOPE} \times \text{INPUT} + Y\text{-INTERCEPT.} \tag{2.15}$$

The slope is the ratio of how much you go up as you go across (rise/run):

$$\text{SLOPE} = \frac{W_B - W_A}{L} \tag{2.16}$$

Therefore we have

$$w(x) = \frac{W_B - W_A}{L}x + W_A \tag{2.17}$$

Example

Write down the load on a 8 m beam due to a linear load that varies from 10 kN m^{-1} at $x = 0$ to 26 kN m^{-1} at $x = 8$. Hence find the bending moment.
Solution: On board.

5. *Linear load over a segment of the beam**

Pretty much the same as the last section except we utilise step functions:

$$w_{\text{linear } x=a\to b}(x) = \left(\frac{W_b - W_a}{b-a}(x-a) + W_a\right) H(x-a) - \left(\frac{W_b - W_a}{b-a}(x-a) + W_a\right) H(x-b) \tag{2.18}$$

where W_a is the load at a and W_b is the load at b with $a < b$.

Example

Write down an expression for a linear load on a 6 m beam which varies linearly from 4 kN m^{-1} at $x = 2$ to 12 kN m^{-1} at $x = 4$. Hence find the bending moment.
Solution: First we draw a picture:

Now we find the equation of the line. This is preferable to using the formula in my opinion. It will not be so easy to use the $y = mx + c$ formula as the line does not cut at $x = 0$ (i.e. it does not cut at A). Hence we use the following:

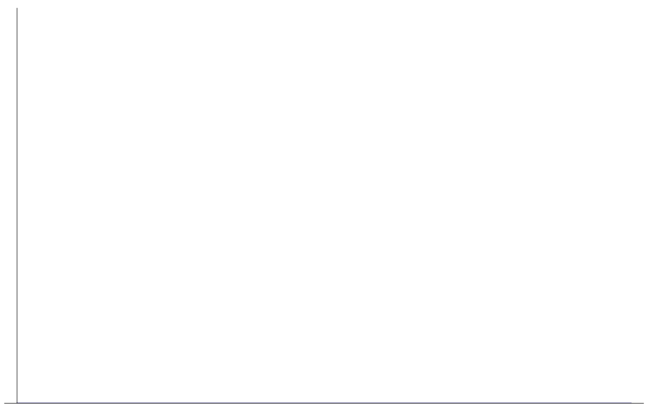


Figure 2.11: The equation of a line can also be written as $y - y_1 = m(x - x_1)$.

Now the only thing is we want this load to be 'switched on' at $x = 2$ and switched off again at $x = 4$ so we use the Heaviside function:

Now to integrate this is not so straightforward but we can take out the Heaviside function rewrite the loading as follows:

Now we want to integrate $-w(x)$ twice to find the bending moment:

That is we have

$$M(x) = \begin{cases} C_1x + C_2 & \text{if } x < 2 \\ -\frac{2}{3}x^3 + 2x^2 + C_1x + C_2 & \text{if } 2 < x < 4 \\ C_1x + C_2 & \text{if } x > 4 \end{cases} \tag{2.19}$$

Now we apply $M(0) = M_A$ and $M'(0) = R_A$:

$$M(x) = \begin{cases} R_A x + M_A & \text{if } x < 2 \\ -\frac{2}{3}x^3 + 2x^2 + R_A x + M_A & \text{if } 2 < x < 4 \\ R_A x + M_A & \text{if } x > 4 \end{cases} \tag{2.20}$$

Exercises What form does the bending moment take for a beam of span 5 m where there is

1. a U.D.L. of 12 kN m^{-1} between $x = 2$ and $x = 5$.
2. a U.D.L. of 10 kN m^{-1} between $x = 0$ and $x = 4$.
3. a U.D.L. of 8 kN m^{-1} between $x = 2$ and $x = 3$.
4. there is a point load of 10 kN at $x = 2$.
5. a linear load varying from 8 kN m^{-1} at $x = 0$ to 18 kN m^{-1} at $x = 5$.
6. a linear load varying from 9 kN m^{-1} at $x = 2$ to 27 kN m^{-1} at $x = 5$.
7. there are point loads of 10 kN and 12 kN at $x = 2$ and $x = 3$ respectively.
8. there is a point load of 12 kN at $x = 3$ and a U.D.L. of 10 kN m^{-1} between $x = 2$ and $x = 5$.
9. there is a point load of 20 kN at $x = 4$ and a U.D.L. of 12 kN m^{-1} between $x = 0$ and $x = 3$.
10. there is a point load of 32 kN at $x = 1$ and a U.D.L. of 16 kN m^{-1} between $x = 2$ and $x = 3$.
11. there is a point load of 10 kN at $x = 2$ and a linear load varying from 5 kN m^{-1} at $x = 3$ to 15 kN m^{-1} at $x = 5$.

2.6.1 A More Pragmatic Method of finding Bending Moments

Note that if you are sure what you are doing you can write down the bending moment straight away. Personally I would prefer to solve the $M'' = -w$ equation but this is an option for you. A very valid question is: why am I showing you how to solve these beam problems from first principles when ye can just look up a table? I have a friend who did Structural Engineering who works for Liebherr. He told me that if he were to solve problems from first principles in his office that he would be told where to go. However, he said that the people who write the computer programmes that they use would know the theory behind all the tables and that these were the people on the big bucks.

If you are going to use Macauley's Method I think we should show the following to be sure about what you are doing:

Macauley's Method

When solving the differential equation

$$\frac{d^2M}{dx^2} = -w(x), \tag{2.21}$$

the following hold:

1. *If the bending moment due to a load* $w_1(x)$ *is* $M_1(x)$, *and the bending moment due to a load* $w_2(x)$ *is* $M_2(x)$, *the bending moment due to the load* $w_1(x)+w_2(x)$ *is given by* $M_1(x)+M_2(x)$.
2. *The bending moment due to a U.D.L. of* w *kN* m^{-1} *is given by*
$$M(x) = -\frac{w^2}{2} + R_A x + M_A, \tag{2.22}$$
where R_A *is the reaction or shearing force at* $x = 0$ *and* M_A *is the bending moment at* $x = 0$.
3. *The bending moment due to a U.D.L. of* w *kN* m^{-1} *applied between points* $x = a$ *and* $x = b$ *(with* $a < b$*) is given by*
$$-\frac{w}{2}[x-a]^2 + \frac{w}{2}[x-b]^2 + R_A x + M_A \tag{2.23}$$
4. *The bending moment due to a point load of magnitude* w *kN at* $x = a$ *is given by*
$$-w[x-a] + R_A x + M_A \tag{2.24}$$
5. *The bending moment due to a linear load varying from* W_A *at* $x = 0$ *to* W_B *at* $x = L$ *is given by*
$$-\frac{(W_B - W_A)x^3}{6L} - \frac{W_A x^2}{2} + R_A x + M_A \tag{2.25}$$

Justification

1. This follows from the fact that integration is linear; e.g. $\int (f(x) + g(x))\,dx = \int f(x)\,dx + \int g(x)\,dx$.

2. The situation looks as follows:

We find by integrating both sides of $\frac{d^2M}{dx^2} = -w$ that the general solution is given by:

$$M(x) = -\frac{w}{2}x^2 + C_1x + C_2. \tag{2.26}$$

Now we apply the following boundary conditions. The bending moment at $x = 0$ is M_A and the shearing force at $x = 0$ is R_A:

3. The situation is as follows:

Now the loading is of the form $W(x) = wH(x - a)$. Therefore we integrate twice to find the bending moment:

Now apply the boundary conditions that shear at $x = 0$ is R_A and the bending moment at $x = 0$ is M_A:

This yields

$$M(x) = -\frac{w}{2}[x-a]^2 + R_A x + M_A. \tag{2.27}$$

4. The situation is as follows:

Now the load function is equal to $w(x) = -W\delta(x-a)$. Integrate twice to find the bending moment:

Apply the boundary conditions to yield:

$$M(x) = -W[x-a] + R_A x + M_A. \tag{2.28}$$

5. Left as an exercise.

Examples

1. Write down the bending moment due to a U.D.L. of 12 kN m^{-1} between $x = 2$ and $x = 5$ on a 5 m beam:

2. Write down the bending moment due to a U.D.L. of 10 kN m^{-1} between $x = 0$ and $x = 4$ on a 5 m beam:

3. Write down the bending moment due to a U.D.L. of 8 kN m^{-1} between $x = 2$ and $x = 3$ and a point load of 12 kN at $x = 1$ on a 5 m beam:

4. a linear load varying from 2 kN m^{-1} at $x = 0$ to 27 kN m^{-1} at $x = 5$ on a 5 m beam:

Exercises Write down the bending moment due to the following loads on a 5 m beam:

1. Point loads of 10 and 12 kN at $x = 2$ and $x = 3$ respectively.
2. A U.D.L. of 10 kN m^{-1} from $x = 2$ to $x = 5$ and a point load of 12 kN at $x = 3$.
3. A U.D.L. of 12 kN m^{-1} from $x = 0$ to $x = 3$ and a point load of 20 kN at $x = 4$.
4. A U.D.L. of 16 kN m^{-1} from $x = 2$ to $x = 3$ and a point load of 32 kN at $x = 1$.
5. a linear load varying from 8 kN m^{-1} at $x = 0$ to 18 kN m^{-1} at $x = 5$.

The three different types of beams that we look at are simply supported beams, fixed end beams and linear loads. The first two differ mathematically only in terms of the initial conditions. The master equation involves four integrations so we must have four initial/boundary conditions.

2.6.2 Simply Supported Beams

A simply supported beam looks as follows:

We have the following boundary conditions:

1. the bending moment at each of the ends is zero: i.e. $M(0) = 0 = M(L)$.
2. the deflection at both ends are zero: i.e. $y(0) = 0 = y(L)$.
3. in addition, if the load is symmetric about the centre $x = L/2$ then we also have $R_A = R_B = W_T/2$, where w_T is the total load.

Example

A light beam of span 5 m is simply supported *at its end points and carries a uniformly distributed load (U.D.L.) of 18 kN* m^{-1} *along the beam. By solving the differential equation*

$$\frac{d^2M}{dx^2} = -w.$$

find the Bending Moment at any point along the beam.
Solution: First we integrate twice to find the general solution:

Now we have to apply some boundary conditions. Here we draw a sketch of the beam.

The (a) uniformly distributed load may be replaced by a point load on the beam at the centre. The total load is the intensity times the length: 18(5) kN = 90 kN. By symmetry the reactions at A and B are given by $R_A = 45$ kN $= R_B$. Also the bending moment about $x = 0$:

$$M(0) = -(90)(2.5) + (45)(5) = 0 \tag{2.29}$$

This is our first boundary condition. From symmetry

$$M(5) = 0. \tag{2.30}$$

Hence we apply these boundary conditions:

This yields the nice solution:

$$M(x) = -9x^2 + 45x \tag{2.31}$$

Here I plot the load, the shearing, and the bending moment on the one graph:

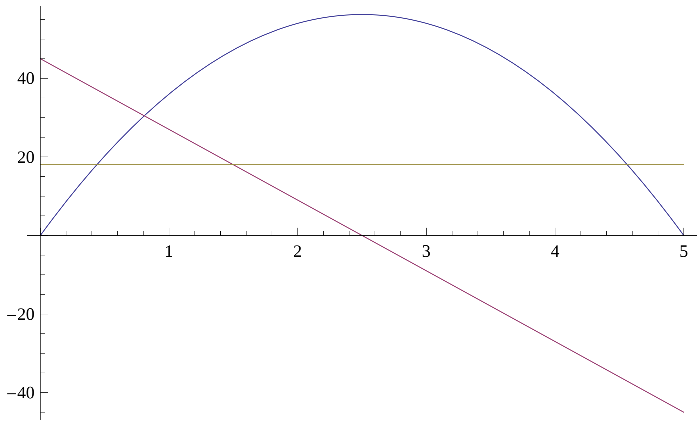


Figure 2.12: A plot of the load, shearing and the bending moment.

Winter 2012: Q. 1 (a)

A light beam of span 4 m is simply supported at its endpoints. At the point $x = 1$ m there is a load of 36 kN. Between the points $x = 2$ m and $x = 3$ m there is a U.D.L. of 30 kN m^{-1}.

(i) Express the Bending Moment M in terms of the step function.

[3 Marks]

(ii) The deflection y at any point on the beam is found by solving the differential equation

$$EI\frac{d^2y}{dx^2} = M.$$

Solve the differential equation where y is zero at both ends of the beam.

[8 Marks]

(iii) What is the deflection of the beam at $x = 3$ m?

[1 Mark]

Solution:

(i) To find the bending moment we solve the differential equation

To solve this we integrate twice

Now because we are simply supported we know that the bending moment at $x = 0$ and $x = 4$ m is zero. The first boundary condition implies that $C_2 = M_A = 0$. Now we apply $M(4) = 0$:

so we have $C_1 = R_A = 51$ so

$$M(x) = -36[x-1] - 15[x-2]^2 + 51x.$$

(ii) We have

$$EI\frac{d^2y}{dx^2} = -36[x-1] - 15[x-2]^2 + 51x.$$

To solve this we integrate twice:

To find C_3 and C_4 we must apply the boundary conditions that $y = 0$ at $x = 0$ and $x = 4$. $y(0) = 0$ implies that $C_4 = 0$. Now we apply $y(4) = 0$:

(iii) This is just $y(3)$:

Winter 2011: Q. 1(c)

A light beam of span 5 m is simply supported at its end points and carries a load that varies uniformly with x the distance from one end of the beam. The load varies from 18 kN m^{-1} at $x = 0$ to 12 kN m^{-1} at $x = 5$. Find a formula for the load per unit length. By solving the differential equation

$$\frac{d^2M}{dx^2} = -w(x), \tag{2.32}$$

find the bending moment M at any point along the beam. Also find the maximum value of the bending moment.

Solution: First a picture:

Now $-w(x) = \frac{6}{5}x - 18$. We integrate twice to find the bending moment:

Now we apply the boundary conditions that $M(0) = 0 = M(5)$. We may *use* R_A equals half the total load as the load is not symmetric:

This yields a bending moment function

$$M(x) = \frac{1}{5}x^3 - 9x^2 + 40x \tag{2.33}$$

The bending moment looks like

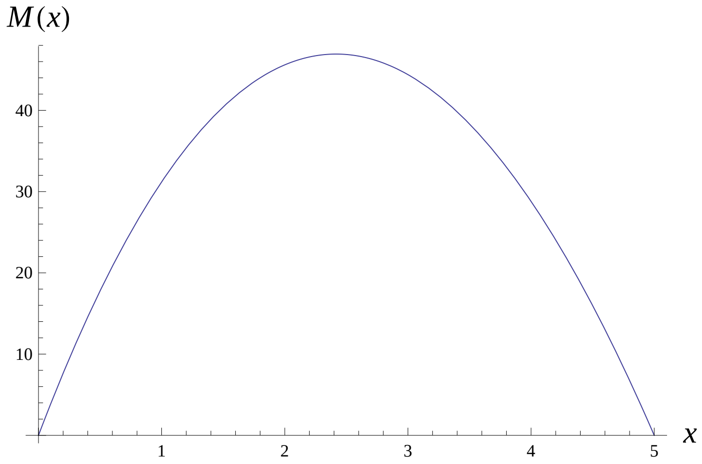


Figure 2.13: A plot of the bending moment.

The maximum occurs when the slope is zero...

Obviously the + here is too big and the maximum is found at $15 - 5\sqrt{57}/3 \approx 2.42$. Throw this into $M(x)$ to find the maximum bending moment:

$$M(2.42) \approx 42.9269 \text{ kN m}.$$

Remark: Normally I would throw the surd into the bending moment function. The use of decimals introduces *rounding errors* into our calculations and pure mathematicians consider unrestricted decimal approximation very gravely indeed. However, at the maximum the bending moment function is quite flat so moving a little bit away from the true value doesn't change $M(x)$ that much:

$$M(x_{\text{max}}) \approx M(x_{\text{max}} + \varepsilon) \tag{2.34}$$

In fact in this example $M(x_{\text{max}})$ and $M(2.42)$ agree to four places of decimals. However doing the same thing around $x = 1$ m is *significant*. In particular $M(1) \approx 31.2$ kN m but $M(1 + 1") \approx 31.789$ kN m — a not insignificant difference — over 500 N m in fact about the torque needed to hold 50 kg stable at an arm's length.

Winter 2011 Q. 1(a)

A light beam of span 6 m is simply supported of its endpoints. Between the points $x = 2$ m and $x = 5$ m there is a U.D.L. of 72 kN m^{-1}. Express the bending moment M in terms of step functions. The deflection y at any point on the beam is found by solving the differential equation

$$EI\frac{d^2y}{dx^2} = -M \tag{2.35}$$

Solve this differential equation where y is zero at both ends of the beam.
Solution: First as always, a picture:

Now I write down the loading and solve $M''(x) = -w$:

Here the loading is not symmetric so the reaction isn't shared equally among the two points. But, as the load is simply supported the bending moment at both of the ends is zero:

$$M(0) = 0 = M(6) \tag{2.36}$$

We can use these two equations to find C_1 and C_2 (R_A and M_A):

Which yields

$$M(x) = 36[x-5]^2 - 36[x-2]^2 + 90x \tag{2.37}$$

Remark: Again, as the picture suggests, most of the reaction force is concentrated at B due to the asymmetry. In fact $M_B = 3(72) - 90 = 126$ kN.

Now integrating twice:

That is we have

$$EIy(x) = 3[x-2]^4 - 3[x-5]^4 - 15x^3 + K_1x + K_2 \tag{2.38}$$

Now we apply the boundary condition that $y(0) = 0 = y(6)$:

Now we can write our final answer:

$$y(x) = \frac{1}{EI}\left(3[x-2]^4 - 3[x-5]^4 - 15x^3 + \frac{825}{2}x\right) \tag{2.39}$$

Winter 2010 Q. 1(a)

A light beam of span 6 m is simply supported of its endpoints. Between the points $x = 4$ m and $x = 6$ m there is a U.D.L. of 72 kN m^{-1}. At the point $x = 2$ there is a load of 72 kN. Express the bending moment M in terms of step functions. Solve the differential equation

$$EI\frac{d^2y}{dx^2} = -M \tag{2.40}$$

to find the deflection at any point on the beam. At both ends the deflection is zero.
Solution: First draw a picture:
We can use Macauley's Method but here I will write down the loading function:

$$w(x) = 72\delta(x-2) + 72H(x-4), \tag{2.41}$$

and integrate $-w(x)$ twice to find $M(x)$:

Now apply the boundary condition $M(0) = 0 = M(6)$:

This yields

$$M(x) = -72[x-2] - 36[x-4]^2 + 72x. \tag{2.42}$$

Now write down the differential equation:

$$EI\frac{d^2y}{dx^2} = -M(x)$$

Now apply the boundary conditions $y(0) = 0 = y(6)$ (as it is simply supported):

Hence we have an answer:

$$y(x) = \frac{1}{EI}\left(36[x-2]^3 + 3[x-4]^4 - 12x^3 + 40x\right) \tag{2.43}$$

2.6.3 Fixed Ends

A fixed end beam of length L looks as follows:

We have the following boundary conditions:

1. the deflection at both ends are zero: i.e. $y(0) = 0 = y(L)$.
2. the slope at both ends is zero: i.e. $y'(0) = 0 = y'(L) = 0$.

Note that $M_A \neq 0 \neq M_B$ necessarily as the wall exerts a bending moment.

Winter 2011: Question 1 (b)

A light beam of span 6 m has both ends embedded in walls. At the point $x = 2$ m there is a load of 36 kN. Between the points $x = 4$ m and $x = 6$ m there is a U.D.L. of 72 kN m^{-1}. Express the bending moment M in terms of step functions. Solve the differential equation

$$EI\frac{d^2y}{dx^2} = -M, \tag{2.44}$$

to find the deflection at any point on the beam.
Solution: First draw a picture:
We can use Macauley's Method but here I will write down the loading function:

$$w(x) = 36\delta(x-2) + 72H(x-4), \tag{2.45}$$

and integrate $-w(x)$ twice to find $M(x)$:

In the case of a simply supported beam we the boundary condition $M(0) = 0 = M(6)$ **but this is not the case when the ends are fixed**. You can carry around an M_A and R_A if you want but I'm just going to use C_1 and C_2.

Now we need to solve

$$\begin{aligned} EI\frac{d^2y}{dx^2} &= -M(x) \\ &= 36[x-2] + 36[x-4]^2 - C_1x - C_2 \end{aligned}$$

Thus

$$EI\frac{d^2y}{dx^2} = 6[x-3]^3 + 3[x-4]^4 - \frac{C_1}{6}x^3 - \frac{C_2}{2}x^2 + C_3x + C_4$$

Four unknowns is tough but we have four boundary conditions which will generate four equations in C_1, C_2, C_3, C_4 which we should then be able to solve. First up let's look at $y(0) = 0 = y'(0)$:

Chapter 3

Probability

The probable is what usually happens.

Aristotle.

3.1 Random Variables

The central concept of probability theory is that of a *random variable*. Examples of random variables include:

1. the outcome of a coin flip
2. the outcome of a dice roll
3. the outcome of a random selection of a card from a deck
4. the number of plants which grow to maturity in a glass house
5. the time spent on hold when calling the tax office
6. the height of an Irish male chosen at random

Associated to a random variable there is a *probability distribution*. The technical definition of a probability distribution is actually quite difficult and for us the following definition will have to do:

3.1.1 Definition

The probability distribution $\mathbb{P}$ of a random variable X is a function

$$\mathbb{P} : \{\text{all possible outcomes concerning } X\} \to [0, 1]. \tag{3.1}$$

Remarks

1. outcomes that *never* happen are assigned a probability of 0 and outcomes that *always* happen are assigned a probability of 1.

Example:

(a) Let $X =$ a dice roll:

(b) Let $Y =$ a five card poker hand:

2. suppose that A and B are mutually exclusive events. Then

$$\mathbb{P}[A \text{ or } B] = \mathbb{P}[A] + \mathbb{P}[B] \tag{3.2}$$

Example: Let $X =$ a random selection from a deck of cards. Now

3. let A be some outcome. Now either A occurs or not-A *always* happens — and also these events are mutually exclusive:

Example: Let $X =$ three coin flips. What is the probability of getting at least one tail? Getting at least one tail is the same as not-(HHH). Therefore

4. when A and B are *independent* outcomes (the occurrence of one outcome makes it neither more nor less probable that the other occurs) we have

$$\mathbb{P}[A \text{ and } B] = \mathbb{P}[A]\mathbb{P}[B]. \tag{3.3}$$

Example: Draw a card X_1 from the deck and replace it at random. Now draw another card X_2. What is the probability of two aces; $X_1 = X_2 = A$? What is the probability of getting two aces if select without replacement?
If we select an ace the first time the chances of selecting an ace again (with replacement) are equal:

However if we select without replacement the probability of a second ace is diminished and instead we must look at the probability:

$$\mathbb{P}[X_1 = A \text{ and } X_2 = A] = \mathbb{P}[X_1 = A]\mathbb{P}[X_2 = A|X_1 = A] \tag{3.4}$$

where $\mathbb{P}[A|B]$ is read *the probability of A given that B is true.* Then

By and large we can describe a random variable X in terms of the probabilities $\mathbb{P}[X = k]$:

3.1.2 Examples

1. Let $X =$ a coin flip. We have

2. Let $Y =$ a dice roll. We have

3. Let $Z =$ a random card from the deck. We have

The other three examples from the start have much more complicated distributions and these are the random variables which we shall pay most attention to.

3.2 Empirical Probability

What is the probability that it will snow in Denver this Christmas Day? In our interpretation of probability we have a random variable $X =$ "it snows on Christmas Day in Denver" and there must be an associated probability distribution. Obviously this is an intractable problem. However if we look at the statistics we find that since 1882 it has snowed on 17 Christmases. We now use this information to *estimate* the true probability distribution as

So we answer the question with $\mathbb{P}[X = \text{snow}] \approx 0.1308$ %.

3.2.1 Empirical Probability Estimate

Suppose that X is a random variable whose probability distribution is unknown. Let $X_1, X_2, \ldots, X_n$ be n independent identically distributed *trials* of X. Then

$$\mathbb{P}[X = k] \approx f_n[X = k] = \frac{\#\text{ instances of } k \in \{X_1, \ldots, X_n\}}{n}. \tag{3.5}$$

How accurate is this kind of an estimate? First it relies on two assumptions:

1. That the trials are all have the probability distribution. *Example*: We assume that the probability of snow in 1882 is equal to the probability of snow in 2012.
2. That the trials are independent. *Example*: Snow in 1901 does not make snow in 1902 more or less likely.

We will not give quantitative bounds on how accurate this method is but we can say that as we take more and more trials that the approximation becomes better and better

Example

Suppose that on three random days that a factory produced 1,800, 2,000 and 2,100 electrical components. Suppose further that on these respective days 52, 64 and 62 faulty devices were produced. Use the data from day one to estimate the probability that a randomly selected device is faulty. Use the data from the three days to give a probabilistically better estimate.
Solution: Firsly we estimate

Now we get a better estimate by viewing the data as 168 faulty devices out of 5,900:

3.3 Binomial Distribution

Consider the case of plants growing in a tunnel. Suppose there are five plants in the tunnel and suppose that we have the probabilities (for each plant)

$$\mathbb{P}[\text{maturity}] = 0.8$$
$$\mathbb{P}[\text{death before maturity}] = \mathbb{P}[\text{not-(maturity)}] = 0.2$$

Now suppose X = number of mature plants in the glasshouse. How do we calculate $\mathbb{P}[X = 3]$? Suppose the plants are labeled I, II, III, IV and V. What about $\mathbb{P}[A] = \mathbb{P}[I, II, III \text{ mature and } IV, V \text{ die}]$? We assume that the mortality statuses of each of the plants are independent so we look at:

$$\begin{aligned}\mathbb{P}[A] &= \mathbb{P}[I \text{ mature}]\mathbb{P}[II \text{ mature}]\mathbb{P}[III \text{ mature}]\mathbb{P}[IV \text{ dead}]\mathbb{P}[V \text{ dead}] \\ &= (0.8)(0.8)(0.8)(0.2)(0.2) \\ &= (0.8)^3(0.2)^2 = 0.02048\end{aligned}$$

However there are many other possibilities for three success stories and two failures:

In total there are ten mutually exclusive ways in which we can get $\mathbb{P}[X = 3]$ so we have

$$\mathbb{P}[x = 3] = 10(0.02048) = 0.2048. \tag{3.6}$$

Data such as this is said to have a *binomial distribution*.

3.3.1 Bernoulli Trials

Consider a random variable T with only two outcomes: success or failure. Suppose that $\mathbb{P}[\text{success}] = p$ then necessarily $\mathbb{P}[\text{failure}] = 1-p$. Such a random variable is said to be a *Bernoulli Trial*. Usually $T = 1$ for success and $T = 0$ for failure.

Examples

1. $T =$ flip of a coin with success defined as a head. Then $\mathbb{P}[\text{success}] = 1/2$.
2. $T =$ roll of a dice with success defined as a six. Then $\mathbb{P}[\text{success}] = 1/6$.
3. $T =$ random card from a deck of cards with success defined as an Ace. Then $\mathbb{P}[\text{success}] = 1/13$.
4. $T =$ the outcome from a door-to-door sales approach with success defined as a sale. Perhaps $\mathbb{P}[\text{success}] = 1/15$?
5. $T =$ the time spent by a random shopper in a shopping queue with success defined as being served in less than two minutes. Perhaps $\mathbb{P}[\text{success}] = 0.8$?
6. $T =$ the number of people a random Irish male murders with success defined as having murdered someone (success in this context can be 'good', indifferent or even evil)! Perhaps $\mathbb{P}[\text{success}] = 0.00001$?!

3.3.2 Binomial Probability Distribution

Suppose that X is the number of successes in a series of n independent and identically distributed Bernoulli Trials, where $\mathbb{P}[\text{success}] = p$:

$$X = T_1 + T_2 + \cdots + T_n \tag{3.7}$$

Then X has the following probability distribution:

$$\mathbb{P}[X = k] = \binom{n}{k} p^k (1-p)^k. \tag{3.8}$$

where $\binom{n}{k} = nCk$ the number of ways of selecting k objects from n distinct objects.

Proof. We can represent the Bernoulli Trials by a random variable W_i with two outcomes — up or down — with $\mathbb{P}[W_i = \uparrow] = p$ and $\mathbb{P}[W_i = \downarrow] = 1-p$. We are thus looking at the probability of k up arrows out of n. There are

possible ways that this can happen and all occur with equal probability, namely

Example: Mature Plants

Use the Binomial Distribution to find the probability that one or more out of ten plants die when the probability of death is 0.1.
Solution: We want to find $\mathbb{P}[X \geq 1] = \mathbb{P}[\text{not-}X = 0] = 1 - \mathbb{P}[X = 0]$:

Winter 2011: Question 4(a)

Concrete blocks are required to satisfy a certain target crushing strength. Any block which fails to satisfy this target strength is considered to be faulty and is rejected. Ten sample of blocks were taken at random. Each sample contained forty blocks and the number of faulty blocks in the samples were counted and recorded:

$$2,\ 3,\ 2,\ 1,\ 2,\ 0,\ 1,\ 1,\ 1,\ 3,\ 1$$

If a single block is taken at random, estimate the probability that it will be rejected. In a sample of 120 of these blocks estimate the probability that more than two will be rejected using the Binomial Distribution.
Solution: Firstly we have 400 blocks in total, $2 + 3 + \cdots + 1 = 16$ of which were faulty:

Now let X be the number of faulty blocks. As X has a binomial distribution with parameters $n = 120$ and $p \approx 0.04$ we can use the binomial distribution formulae to calculate the probabilities $\mathbb{P}[X = 0]$, $\mathbb{P}[X = 1]$ and $\mathbb{P}[X = 2]$:

$$\begin{aligned}\mathbb{P}[X > 2] &= \mathbb{P}[\text{not} - (X = 0, 1, 2)] \\ &= 1 - \mathbb{P}[X = 0 \text{ or } 1 \text{ or } 2] = 1 - (\mathbb{P}[X = 0] + \mathbb{P}[X = 1] + \mathbb{P}[X = 2]) \\ &= 1 - \left(\binom{120}{0}(0.04)^0(0.96)^{120} + \binom{120}{1}(0.04)^1(0.96)^{119} + \binom{120}{2}(0.04)^2(0.96)^{118}\right) \\ &= 1 - 0.137173 = 0.8628.\end{aligned}$$

Autumn 2010: Question 4(a)

Samples of batches of forty items were taken at random from the output of a machine and for each sample the number of defective items were counted and recorded

Number of Defectives	0	1	2	3	4
Number of Batches	51	25	18	5	

If an item is picked at random estimate the probability that it is defective. In a sample of one hundred of these items, estimate the probability that three or more will be defective using the Binomial distribution. Solution: There are, in total,

or 4000 individual items in the sample. We estimate the probability of a defective by calculating the number of defectives :

So we have

$$\mathbb{P}[\text{defective}] \approx 0.02. \tag{3.9}$$

Now we have a random variable X the number of defectives. Here success is defined as a defective and we have $p = 0.02$. We calculate

Winter 2009: Question 4 (a)

On average 0.2 % of machines in a plant need to be replaced every year. In a sample of 100 of these machines calculate the probability that less than three will need to be replaced in a particular year using the Binomial distribution.

Solution: Let X be the number of machine replacements in a year. Under the binomial distribution success is defined as having to replace a machine. What is the probability of success? We show that it must be 0.002 (=0.2 %). Suppose there are in total S machines in the factory, 0.2 % which need to be replaced — i.e. $0.002S$. Now we have

Hence we must calculate the probability that X is less than three with $p = 0.002$.

$$\begin{aligned}\mathbb{P}[X < 3] &= \mathbb{P}[X = 0, 1 \text{ or } 2] = \mathbb{P}[X = 0] + \mathbb{P}[X = 1] + \mathbb{P}[X = 2] \\ &= \binom{100}{0}(0.002)^0(0.998)^{100} + \binom{100}{1}(0.002)^1(0.998)^{99} + \binom{100}{2}(0.002)^2(0.998)^{98} \\ &= 0.99881\end{aligned}$$

Autumn 2009: Question 4 (a)

Items are sold in batches of two hundred. From past history three items in every one thousand fail to satisfy tolerance requirements and are deemed to be defective. Use the Binomial distribution to find the probability that a sample of a batch of these items contains two or more defectives.
Solution: Let X be the number of faulty items in a batch and success is defined as a faulty item. Hence $p = 3/1000 = 0.003$. We thus calculate

$$\begin{aligned}\mathbb{P}[X > 2] = \mathbb{P}[\text{not} - (X \leq 1)] &= 1 - \mathbb{P}[X = 0 \text{ or } 1] \\ & 1 - (\mathbb{P}[X = 0] + \mathbb{P}[X = 1]) \\ & 1 - \left(\binom{200}{0}(0.003)^0(0.997)^{200} + \binom{200}{1}(0.003)^1(0.997)^{199}\right) \\ &= 1 - 0.878297 = 0.121703\end{aligned}$$

Exercises:

1. 20 % of the items produced by a machine are defective. Four items are chosen at random. Find the probability that none of the chosen items is defective.
2. Five unbiased coins are tossed.
   (a) Find the probability of getting three heads and two tails.
   (b) The five coins are tossed eight times. Find the probability of getting three tails exactly four times.
3. During a match John takes a number of penalty shots. The shots are independent of each other and his probability of scoring with each shot is 0.8.
   (a) Find the probability that John misses each of his first four penalty shots.
   (b) Find the probability that John missed exactly three of his first four penalty shots.
   (c) If John takes ten penalty shots during the match, find the probability that he scores at least eight of them.
4. Whenever Anne's mobile phone rings, the probability that she answers the call is 3/4. A friend phones Anne six times.
   (a) What is the probability that she misses all the calls.
   (b) What is the probability that she misses the first two calls and answers the others?
   (c) What is the probability that she answers exactly one of the calls?
   (d) What is the probability that she answers at least two of the calls?
5. The probability of passing a driving test is 2/3. Six students take the test. Use the binomial distribution to find
   (a) the probability that none of the students passes.

(b) the probability that half the students pass the test.

6. A manufacturer of electronic components employs the following quality control plan. Out of each batch of components, 15 are randomly selected and tested. If 3 or more of the 15 components are found to be defective, the entire batch is rejected. Find, correct to three decimal places, the probability that a batch will be rejected if 1/20 of its components are defective?

Wi '12 Q. 4 (a) A production manager is required to monitor the number of defective items output by a machine. Over a period of time he finds that 22% of the output are defective. The production manager decides to examine a sample of 25 items. Find the probability that

(i) five items are defective. Ans: 0.19

(ii) at least three items are defective. Ans: 0.936

3.4 Poisson Distribution

Consider the random variable that arises from counting the number of occurrences of a rare event. Such a random variable is termed a *Poisson* variable, and its probability distribution is termed the Poisson distribution.

Example

1. Number of typographical errors per page
2. Number of industrial accidents per month in a factory
3. Number of deaths per year for an 'age cohort'
4. Number of outbreaks of disease per year

The Poisson distribution has one parameter, λ ('lambda'), the mean average number of occurrences per unit measurement (time,space,page,etc).

The formula for generating the probability distribution is given by

$$\mathbb{P}[X = k] = \frac{e^{-\lambda}\lambda^k}{k!} \tag{3.10}$$

where k is any integer value from 0 upwards.

Theoretically there is no upper limit on possible values for k. However, depending on the value of λ, the probability of higher values for k can quickly become very small.

Example

The number of outbreaks of Ebola is a Poisson variable with mean $\lambda = 0.7$ outbreaks per year. In the next year what is the probability of

1. no outbreaks?
2. at most two outbreaks?
3. at least two outbreaks?

In the next decade what is the probability of five outbreaks?
Solution:

1. Using the formula:

2. $\mathbb{P}[X \text{ at most } 2] = \mathbb{P}[X \leq 2] = \mathbb{P}[X = 0 \text{ or } 1 \text{ or } 2]$:

3. $\mathbb{P}[X \text{ at least } 2] = \mathbb{P}[X \geq 2] = \mathbb{P}[\text{not} - (X = 0 \text{ or } 1)]$:

The time period is now ten years rather than one year. We need to adjust the mean accordingly. If the mean per year is 0.7, then the mean per decade is $10(0.7) = 7$.

$$\mathbb{P}[X = 5] = \frac{e^{-7}7^5}{5!} = 0.1277.$$

Winter 2011: Question 4(c)

On average 2880 vehicles arrive at the Jack Lynch tunnel is any single hour. If four or more vehicles arrive at a particular point during a five second period a queue forms. Calculate the probability of a queue forming in a given five second period.
Solution: Let $X =$ the number of cars that arrive at the Jack Lynch tunnel in a five second period. We have

2880 cars per hour
48 cars per minute
4 cars per five seconds

Hence $\lambda = 4$.

Now we must calculate

$$\begin{aligned}\mathbb{P}[X \geq 4] &= \mathbb{P}[\text{not} - (X < 4)] \\ &= 1 - \mathbb{P}[X = 0 \text{ or } 1 \text{ or } 2 \text{ or } 3] :\end{aligned}$$

Winter 2012 Question 4 (b)

A production manager is required to monitor the number of defective items output by a machine. Defective items are suspected to be output randomly at an average of 1.2 per hour.

(i) Calculate the probability of the machine outputting three defective items in any given two hour period.

(ii) Calculate the probability of having more than 10 defective items over and 8 hour period. (Assume a Poisson Distribution).

[2 & 2 Marks]

Solution:

(i) We have that the number of defectives per hour X is a Poisson random variable with $\lambda = 1.2$ and so we want $\mathbb{P}[X = 3]$:

(ii)

Winter 2009 Question 4(c)

In a computer laboratory the number of files sent by students to a printer during a period of one hour were counted and are recorded below

Number of Files	0	1	2	3	4
Number of students	4	7	10	3	1

Calculate the average number of files sent to this printer during any single hour. By using the Poisson Distribution calculate the probability that a student will send two or three files to the

printer during any single hour.
Solution: In total there are $4 + 7 + 10 + 3 + 1 = 25$ students and altogether they sent

$$4(0) + 7(1) + 10(2) + 3(3) + 1(4) = 40$$

files to the computer. Hence the average student sent

Now let X = the number of files a student sends to the printer in an hour. We have $\lambda = 1.6$:

$$\mathbb{P}[X = 2 \text{ or } 3] = \mathbb{P}[X = 2] + \mathbb{P}[X = 3]$$

Winter 2010: Question 4(c)

The average number of vehicles arriving at a particular interchange is 1080 per hour. Using the Poisson Distribution calculate the probability of four or more vehicles arriving at this interchange in any five second period.
Solution: Once again

$$\begin{aligned}
&1080 \text{ cars per hour} \\
&18 \text{ cars per minute} \\
&1.5 \text{ cars per five seconds}
\end{aligned}$$

Hence $\lambda = 1.5$. Now we must calculate

$$\begin{aligned}
\mathbb{P}[X \geq 4] &= \mathbb{P}[\text{not} - (X < 4)] \\
&= 1 - \mathbb{P}[X = 0 \text{ or } 1 \text{ or } 2 \text{ or } 3] :
\end{aligned}$$

Autumn 2009 Question 4(c)

The number of calls to a helpdesk were counted during a one hour period.

Number of Calls	2	3	4	5	6	7
Number of Hours	1	3	7	6	5	3

Calculate the mean number of calls per hour. By using the Poisson Distribution calculate the probability that there will be less than three calls in any particular ten minute period.
Solution: The total number of hours is

The total number of calls is

Hence the mean number of calls per hour is 4.8.

Let $X=$ the number of calls in any ten minute spell. If there are 4.8 calls per hour there are 4.8/6=0.8 calls per ten minute period. Hence we have $\lambda = 0.8$. We wish to calculate $\mathbb{P}[X < 3] = \mathbb{P}[X = 0 \text{ or } 1 \text{ or } 2]$:

Exercises:

1. Do Winter 2011 Question 4(a), Winter 2009 4(a) and Autumn 2009 4(a) using the Poisson Distribution.
2. The IT manager of a large company reports 3 computer failures during the past 10 days. From this fact he assumes the number of computer failures follows a Poisson distribution with a rate of $\lambda = 3/10 = 0.3$ failures per day.
   (a) What is the probability that no failures occur on any given day?

   (b) What is the probability that at least 3 failures occur on any given day?
   (c) What is the probability that at most 2 failures occur on any given day?

3. A recent audit of the telesales department of this multinational software company indicates the team sells 3.1 copies of its leading product per hour. Assume that the distribution of the number of copies sold per hour can be approximated by the Poisson probability distribution.
   (a) What is the probability that at most 3 copies are sold in an hour?
   (b) What is the probability that at least 2 copies are sold in any given hour?
   (c) What is the probability that no copies of the software are sold in 2 consecutive hours?
4. The number of invoices issued per month containing mistakes is a Poisson random variable with mean 1.7.
   (a) What is the probability that no invoice contains a mistake in any given month?
   (b) What is the probability that the firm issues at most 2 invoices with mistakes in any given month?
   (c) What is the probability that the firm issues at least 2 invoices with mistakes in any given month?
5. The number of new clients at a solicitors practice follows a Poisson distribution with a mean of 2.8 new clients per week. For any given week, calculated the following probabilities:
   (a) There are 4 new clients.
   (b) There at least 3 new clients.
   (c) Find the probability of exactly 10 new clients in a four-week period.
6. The number of new computer viruses detected by an anti-virus software can be modelled as a Poisson distribution a mean of 8 per week.
   (a) What is the probability that exactly 4 new viruses will be detected in a given week?
   (b) What is the probability that at least 6 new viruses will be detected in a given week?
   (c) What is the probability that exactly 40 new viruses will be detected in a 4- week period?
7. The number of parking tickets issued by a traffic warden follows a Poisson process with a mean of 4 per hour.
   (a) What is the probability that exactly three tickets are issued during a particular hour?
   (b) What is the probability that at least three tickets are issued during a particular hour?

3.5 Normal Distribution

A machine is filling a very fine powder into containers. Let $X =$ the full put into the next one. Due to the variability of the machine, the value of X is not certain. A *continuous random variable* is a random variable that can take on any value in a continuous range. This is different to what we have looked at before as before we examined some random variable X which takes on values $k = 0, 1, 2$, etc. and we examined the probabilities $\mathbb{P}[X = k]$. For a continuous random variable there are just too 'many' possible outcomes and all the probabilities $\mathbb{P}[X = k]$ are zero. Hence we must change tack.

Suppose that in a particular population of males the following distribution of heights are found:

Height	<150	150-160	160-170	170-180	180-190	190-200	>200
Relative Frequency	0.03	0.1	0.2	0.33	0.2	0.1	0.03

This data could be summarised in a relative frequency distribution:

Figure 3.1: The heights of males in a population.

What we speculate is that there is a probability density function that can tell us not only how what proportion of people have a height between 170 and 180 cm but between 170 and 172 cm or any other interval:

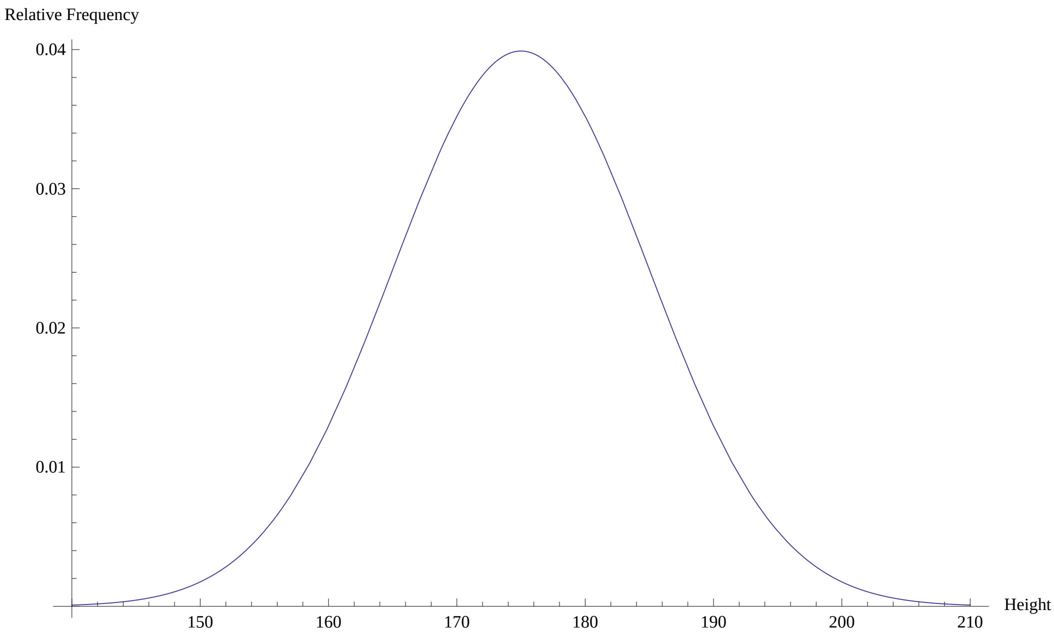


Figure 3.2: The *probability density function* of the heights.

Many examples of continuous data are 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.

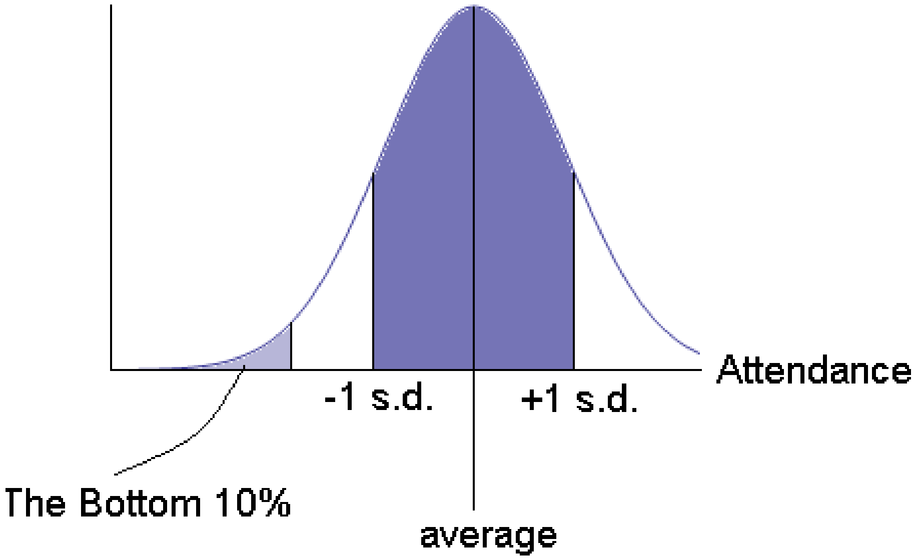


Figure 3.3: Note that for a normal distribution approximately 68% of the data is found within a distance of one standard deviation from the average.

3.5.1 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*

$$z = \frac{X - \mu}{\sigma} \tag{3.11}$$

This is a transform that converts the X-distribution to a z-distribution (Note that $\mu \to 0$ by (3.11) as you would hope.):

We can also go from z- back to X- via

$$X = \mu + z\sigma \tag{3.12}$$

It is not clear why we are doing this but the following fact makes it all clear:

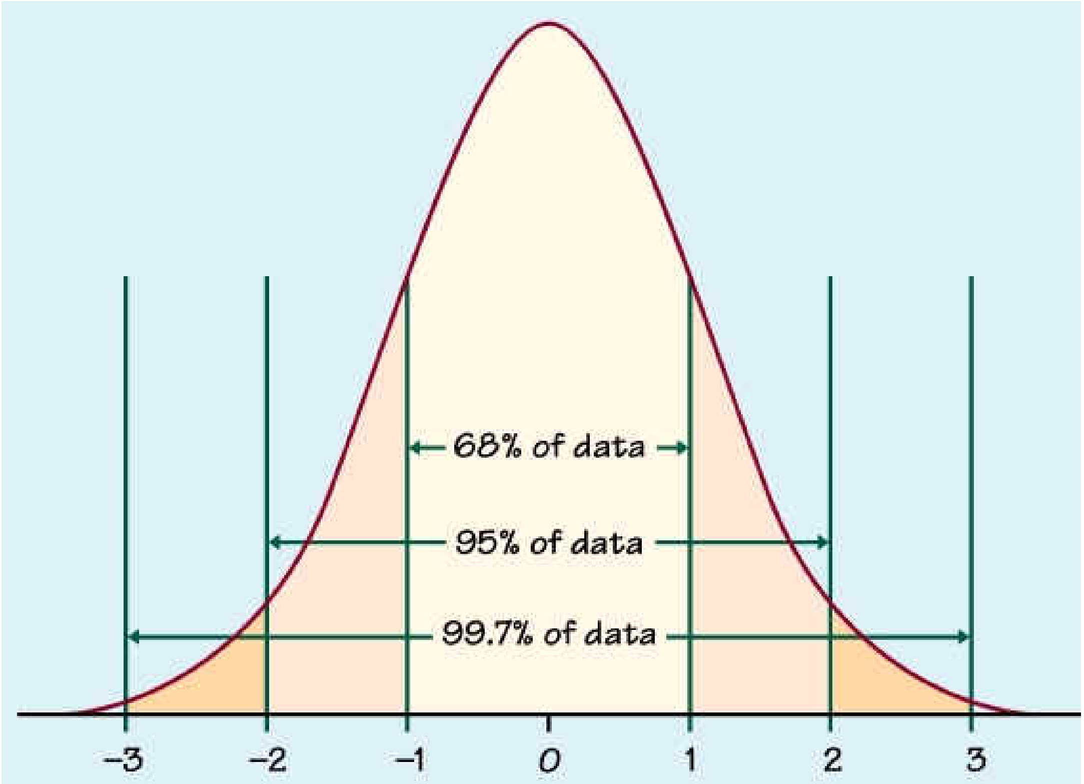


Figure 3.4: All normal distributions are just transforms of $z = N[0, 1]$.

Fundamental Calculation of Normal Distributions

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

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

Then we can transform $x_1 \to z_1$ *and* $x_2 \to z_2$ *using (3.11); and in this case*

$$\mathbb{P}[x_1 \leq X \leq x_2] = \mathbb{P}[z_1 \leq z \leq z_2] \tag{3.13}$$

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 $\mathbb{P}[z \geq z_1] = 0.2$:

Alternatively we can look at $\mathbb{P}[z \leq z_1] = 0.8$ and read off the table and get $z = 0.84$.

Winter 2011: Question 4(b)

In a normally distributed population the mean value and standard deviation are given by 20 kg and 0.07 kg, respectively.

1. *Find the probability that a reading is less than 20.1 kg.*
2. *Find the percentage of readings that lie between 19.8 kg and 20.2 kg.*
3. *If 99.9% of items have a weight less than some critical weight W find the value of W.*

Solution: On board.

Winter 2009: Question 4(b)

The weights of blocks are assumed to be Normally distributed with a mean value of 1 kg and with a standard deviation of 0.06 kg.

1. *Calculate the probability that a block will weigh less than 1.075 kg.*
2. *Also calculate the probability that a block will weigh between 0.92 and 1.14 kg.*
3. *If 99% of weights have a weight less than some critical weight W find the value of W.*

Solution: On board.

Winter 2010: Question 4(b)

The lengths of pipes are assumed to be normally distributed with a mean value of 5 m and with a standard deviation of 0.003 m. What percentage of pipes have a length (i) less than 5.002 m and (ii) between 4.99 m and 5.01 m. If 99.8% of diameters are less than some value α, find the value of α. Solution: On board.

Autumn 2010: Question 4(b)

The weights have bricks are assumed to be normally distributed with a mean value of 1 kg and with a standard deviation of 0.006 kg. Calculate the percentage of bricks that weigh (i) less than 1.02 kg and (ii) between 0.99 kg and 1.01 kg. (iii) If 0.1% of bricks have a weight greater than some critical weight W find the value of W. Solution: On board.

Exercises:

1. z is a random variable with standard normal distribution. Find the value of z_1 for which $\mathbb{P}[z > z_1] = 0.0808$.
2. z is a random variable with standard normal distribution. Find $\mathbb{P}[1 < z < 2]$.
3. z is a random variable with standard normal distribution. Calculate $\mathbb{P}[-2.13 < z \leq 1.46]$.
4. z is a random variable with standard normal distribution. Find $\mathbb{P}[z < -0.46]$.
5. The heights of students in a certain college are normally distributed with mean 165 cm and standard deviation 10 cm. If a student is chosen at random find the probability that the student's height is
   (a) less than 170 cm
   (b) between 160 cm and 180 cm.

6. The amounts due on monthly mobile phone bills are normally distributed with mean E53 and standard deviation E15. If a bill is chosen at random, find the probability that the amount due is between E47 and E74.

7. The lifetime of a particular type of electric bulb is normally distributed with a mean of 1500 hours and a standard deviation of 120 hours. If 140 bulbs are purchased, how many can be expected to have a lifetime between 1400 hours and 1730 hours inclusive? Give your answer correct to the nearest whole number.

8. 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.

9. 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

   (a) less than 8.07 kg

   (b) greater than 8.08 kg

   (c) 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.

10. 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?

3.6 Sampling Theory

Consider the problem of finding the mean-average length, μ, of all the metallic rods produced in a factory. Plainly this is impossible. However we could *approximate* this population mean-average by taking a random sample of say n rods from 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. Suppose we measure these rods to have lengths

$$l_1, l_2, \ldots, l_n.$$

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

$$\bar{l} = \frac{l_1 + \cdots + l_n}{n} = \frac{1}{n}\sum_{i=1}^{n} l_i,$$

Now we could take $\bar{l}$ as an estimate of μ; $\bar{l} \approx \mu$. How accurate is this? Now consider all the possible samples we could have taken from the population:

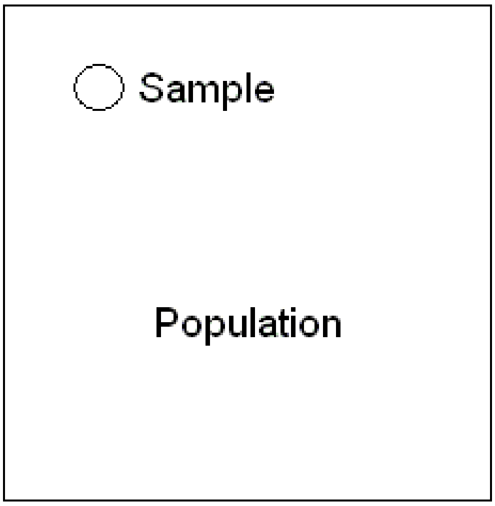


Figure 3.5: There are many, many ways of choosing n rods from a population of perhaps millions. If we look at the sample mean-average as a random variable, then the mean-average of the sample mean-averages is equal to the population mean-average.

We can show that the standard deviation of the sample means from the population mean is given by $\sigma/\sqrt{n}$, where σ is the standard deviation of the population. For the standard deviation we use

$$\sigma = \sqrt{\frac{\sum(x-\bar{x})^2}{n-1}} = \sqrt{\frac{n}{n-1}\left(\frac{\sum x^2}{n} - (\bar{x})^2\right)} \tag{3.14}$$

For a small sample $n < 30$, the sample means have what is know as a t-distribution. You may also come across the term *variance*. The variance is the square of the standard deviation.

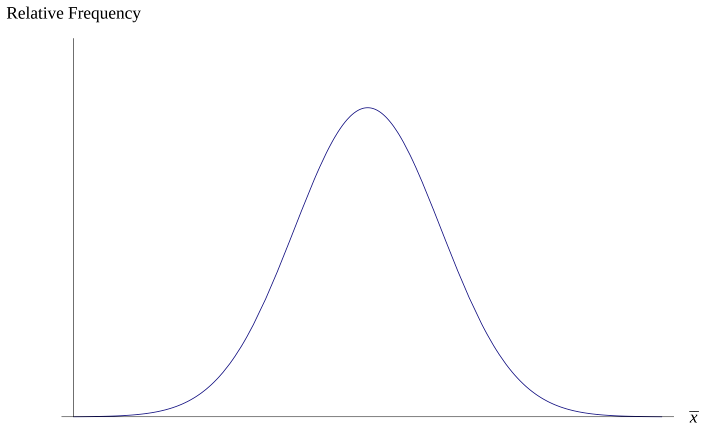


Figure 3.6: For small sample sizes from normally distributed data, the most likely sample mean is equal to the population mean. However, represented by the top and bottom tails, there are possibilities of choosing a sample with a particularly large or particularly small sample mean.

Similarly to the normal distribution, there is a table which, depending on the sample size, gives the percentage of the data which lies within t standard deviations of the mean. For example, suppose a sample of size $n = 10$ is taken from a population of standard deviation σ. Then 95% of the sample means $\bar{x}$ lie within 2.262 standard deviations of the (population mean):

$$\mu - 2.262\frac{\sigma}{\sqrt{10}} \leq \bar{x} \leq \mu + 2.262\frac{\sigma}{\sqrt{n}}.$$

Now we can be 95% confident that our sample mean is in this range:

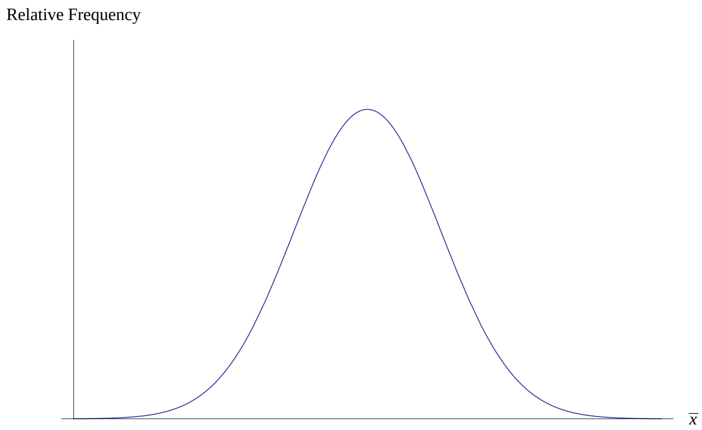


Figure 3.7: If the sample mean is within this distance of the population mean then, in turn, the population mean will be within this distance of the sample mean.

Note that in reality we have no idea what the standard deviation, σ, of the population is but it can be shown that the sample standard deviation provides a good estimate: $s \approx \sigma$.

Formula: Confidence Interval for Small Sample Sizes

Suppose that a sample of size $n < 30$ is taken from a population with a standard deviation of σ. If it is found that the sample mean is $\bar{x}$, then the following is the p% Confidence Interval for the population mean:

$$\bar{x} - t\frac{\sigma}{\sqrt{n}} \leq \mu \leq \bar{x} + t\frac{\sigma}{\sqrt{n}} \tag{3.15}$$

Here t depends on n and the level of confidence p, and is to be read from the t-distribution table.

Note the interpretation:

On the p% confidence level, the population mean lies within the confidence interval.

Winter 2011: Question 5 (a) (i)

Six determinations were made about the density of a particular material

$$3.09,\ 3.11,\ 3.11,\ 3.11,\ 3.12,\ 3.12.$$

Find 99% confidence limits for the density.

[HINT: For the data above $s^2 = 0.012$].

Solution: We must calculate the interval

$$\bar{x} - t\frac{\sigma}{\sqrt{n}} \leq \mu \leq \bar{x} - t\frac{\sigma}{\sqrt{n}} \tag{3.16}$$

We calculate the sample mean, $\bar{x}$:

$$\bar{x} = \frac{3.09 + 3.11 + 3.11 + 3.11 + 3.12 + 3.12}{6} = 3.11.$$

We approximate $\sigma \approx s = \sqrt{0.012}$. Now the t-value. Here we have $n = 6$ so the degrees of freedom are $n - 1 = 5$. For a 99% confidence interval we will want 2.5% or 0.025 in the tail... this yields the t-value $t = 2.571$:

$$\bar{x} - t\frac{\sigma}{\sqrt{n}} \leq \mu \leq \bar{x} - t\frac{\sigma}{\sqrt{n}}$$
$$\Rightarrow 3.11 - 2.571\frac{\sqrt{0.012}}{\sqrt{6}} \leq \mu \leq 3.11 - 2.571\frac{\sqrt{0.012}}{\sqrt{6}}$$
$$\Rightarrow 2.99502 \leq \mu \leq 3.22498$$

3.6.1 Confidence Intervals for Large Sample Sizes

In contrast to these confidence intervals we can generate far superior confidence intervals for larger sample sizes $n > 30$. Essentially we have the same analysis as before except the distribution of sample means is not a t-distribution but rather a normal distribution:

$$\bar{x} \sim N[\mu, \sigma/\sqrt{n}]. \qquad (3.17)$$

This means that the sample means have a normal distribution with mean of μ, the population mean; and a standard deviation of $\sigma/\sqrt{n}$, where σ is the population standard deviation. When $n > 30$ and the data is not grouped, we cannot easily calculate the sample standard deviation and instead probably use a statistical software package. Following the analysis through yields:

Formula: Confidence Interval for Large Sample Sizes

Suppose that a sample of size $n > 30$ is taken from a population with a standard deviation of σ. If it is found that the sample mean is $\bar{x}$, then the following is the p% Confidence Interval for the population mean:

$$\bar{x} - z\frac{s}{\sqrt{n}} \leq \mu \leq \bar{x} + t\frac{s}{\sqrt{n}} \qquad (3.18)$$

Here z depends on the level of confidence p (but not the sample size), and can be read from the z-distribution tables.

The z values for 90%, 95% and 99% are 1.645, 1.96 and 2.56 (why)?

Winter 2011: Question 5 (a) (ii)

If a group of 50 people measured the density of this material and their data yielded a mean value of 3.11 and a standard deviation of 0.2 find 99% confidence limits for the density.
Solution: We must calculate the interval

$$\bar{x} - z\frac{\sigma}{\sqrt{n}} \leq \mu \leq \bar{x} - z\frac{\sigma}{\sqrt{n}} \qquad (3.19)$$

We have $\bar{x} = 3.11$ and $\sigma \approx 0.2$. As this is a large sample we must find the z-value for 95%. This is 1.96:

$$\bar{x} - z\frac{\sigma}{\sqrt{n}} \leq \mu \leq \bar{x} - z\frac{\sigma}{\sqrt{n}}$$
$$\Rightarrow 3.11 - 1.96\frac{0.2}{\sqrt{50}} \leq \mu \leq 3.11 + 1.96\frac{0.2}{\sqrt{50}}$$
$$\Rightarrow 2.94997 \leq \mu \leq 3.27003$$

Exercises:

W '10 (i) Ten measurements were made of a particular distance between two points (m) and the following results were obtained:

$$12.45,\ 12.50,\ 12.50,\ 12.55,\ 12.45,\ 12.50,\ 12.50,\ 12.45,\ 12.55,\ 12.55.$$

Find 99% confidence limits for the actual distance.

[Hint: $\sum x = 1250$, $\sum x^2 = 1562.525$]

(ii) If 60 students measured this distance and their measurements yielded a mean value of 12.52 m with a variance of 0.0020 m^2, find 99% confidence limits for the actual distance.

W '09 (i) A group of six people measured a particular height and these measurements are recorded below. Find the sample mean $\bar{x}$ and the sample variance s^2. Hence find 95% confidence limits for the actual height.

$$6.49,\ 6.49,\ 6.50,\ 6.50,\ 6.51,\ 6.51$$

[Hint: $\sum x = 39$, $\sum x^2 = 253.5004$]

(ii) In calculating this height, 50 measurements gave a mean value of 6.5 m with a standard deviation of 0.008 m. Find 95% confidence limits for the actual height.

A '09 (i) Five students measured a certain area and the following values were obtained

$$1.05,\ 1.07,\ 1.07,\ 1.05,\ 1.06\,,\text{hectares.}$$

Find a 98% confidence interval for the actual area.

(ii) If the measurements of this area by 50 students gave a mean value of 1.06 hectares with a standard deviation of 0.09 hectares, find a 98% confidence interval for the actual area.

W '08 The compressive strength of concrete is subject to variation.

(i) Five samples of a concrete mix were taken and the strength of concrete (N mm^{-2}) provided for these samples are recorded:

$$29.9,\ 29.9,\ 30.0,\ 30.1,\ 30.1.$$

Set up a 95% confidence interval for the mean strength.

(ii) If a sample of 50 measurements gave a mean value of 30 N mm^{-2} with a standard deviation of 0.12 N mm^{-2}, set up a 95% confidence interval for the mean strength.

3.7 Hypothesis Testing

Consider a manufacturing company which claims to produce metal rods that have a mean length of 2 m. This is an assertion that has been made by the company about the population of produced goods. In this section this shall be referred to as H_0, the *null hypothesis* — the *status quo*:

The mean length of our rods is 2 m (H_0)

After using these rods for a period of time a construction company might say hang on; these rods aren't made to these specifications. They might put forward an *alternative hypothesis*, H_A. For example,

The mean length of these rods is not 2 m, OR
The mean length of these rods is bigger than 2 m.

To settle this dispute the manufacturing company may take a random sample of n rods. From this they will find a sample mean $\bar{x}$ and standard deviation s. The question is: do the sample observations support H_0 or H_A? To decide this carry out the following steps:

1. State the H_0 and H_A for the situation under investigation.
2. On the assumption that H_0 is true, depending on the level of significance required, find a 'confidence interval' for the sample means.
3. If the sample mean falls outside this 'confidence interval', reject H_0. Otherwise we cannot reject H_0.

Suppose for example, we are working with the 5% level of significance. Assuming H_0 is true, the probability that a sample mean lies outside the 'confidence interval' is 0.05. If in a random sample of that size, we find that the sample mean lies outside these limits, this is good evidence in favour of rejecting H_0. Alternatively, if the sample mean lies inside these limits (but maybe quite far from the assumed mean), then we can say that 'rare events' happen and there is no strong reason to reject H_0 at this point.

3.7.1 Two types of Errors in Hypothesis Testing

Whenever we have to choose between two alternatives, there are two distinct types of error that may occur:

I: Reject H_0 when H_0 is true.

II: Cannot reject H_0 when H_A is true.

Type I errors are considered more serious then Type II. On the one hand, Type I errors can be seen as 'throwing the baby out with the bath water' but an analogy with criminal law shows why. In this context, the level of significance can be seen as the risk at which the decision-maker is willing to make a Type I error.

Example: Murder Trial

In a murder trial, the null hypothesis (innocent until proven guilty) is given by:

The defendant did not murder the deceased (H_0)

The alternate hypothesis is

The defendant did murder the deceased (H_A)

Implicit in H_0 is a range of possibilities that are consistent with H_0. It is up to the prosecuting council to produce evidence which lies outside this range of possibilities (e.g. the probability of the defendant being innocent AND owning the gun is small). If the prosecuting council succeed in producing such evidence (similar to a sample mean outside the 'confidence interval'), the jury reject H_0 and the man is convicted. A Type I error here is convicting an innocent man — which is considered worse than a Type II: acquitting a guilty man.

For small sample sizes we should use a t-distribution but your exam will have $n > 30$ and hence we use a z-distribution to construct our 'confidence interval'.

Winter 2011: Question 5(b) (i)

Concrete is claimed to have a specific characteristic strength of 25 N mm^{-2}. If 40 samples gave a mean value of 25.05 N mm^{-2} with a standard deviation of 0.2 N mm^{-2}, test at the 0.05 *level of significance that the mean is not equal to 25 N mm^{-2}.*

Solution: First we state the null and alternate hypothesis:

H_0: *the mean is equal to 25*

H_A: *the mean is not equal to 25*

Now we set up the 95% 'confidence interval' for the sample means:

$$\mu - z\frac{\sigma}{\sqrt{n}} \leq \bar{x} \leq \mu + z\frac{\sigma}{\sqrt{n}}. \tag{3.20}$$

Here we want a 95% 'confidence interval' so we have $z = 1.96$. We take $\mu = 25$, $\sigma = 0.2$ and $n = 40$:

$$25 - 1.96\frac{0.2}{\sqrt{40}} \leq \bar{x} \leq 25 + 1.96\frac{0.2}{\sqrt{40}}$$
$$\Rightarrow 24.938 \leq \bar{x} \leq 25.062$$

Finally 25.05 lies in this 'confidence interval' so we cannot reject H_0 at the 0.05 level of significance.

Winter 2009: Question 5(b) (i)

The breaking strength of a particular metal is estimated to be 43.5 N mm^{-2}. A sample of 100 specimens gave a mean value of 43.4 N mm^{-2}, with a standard deviation of 0.9 N mm^{-2}. Test at the 0.01 level of significance that the mean value is not equal to 43.5 N mm^{-2}.

Solution: First we state the null and alternate hypothesis:

H_0: *the mean is equal to 43.5*

H_A: *the mean is not equal to 43.5*

Now we set up the 99% 'confidence interval' for the sample means:

$$\mu - z\frac{\sigma}{\sqrt{n}} \leq \bar{x} \leq \mu + z\frac{\sigma}{\sqrt{n}}. \tag{3.21}$$

Here we want a 99% 'confidence interval' so we have $z = 2.57$. We take $\mu = 43.5$, $\sigma = 0.9$ and $n = 100$:

$$43.5 - 2.57\frac{0.9}{\sqrt{100}} \leq \bar{x} \leq 43.5 + 2.57\frac{0.9}{\sqrt{100}}$$
$$\Rightarrow 43.269 \leq \bar{x} \leq 43.731$$

Finally 43.4 lies in this 'confidence interval' so we cannot reject H_0 at the 0.01 level of significance.

3.7.2 One Sided 'Confidence Intervals'

The above is basically how we do these questions but what about alternative hypotheses such as

The mean length is larger than 2 m.

These need a one-sided 'confidence interval'. In this case we are not interested in the case where the sample mean is small — we are looking for it to be big. However, again, we would need the sample mean to be abnormally large to change our mind about μ:

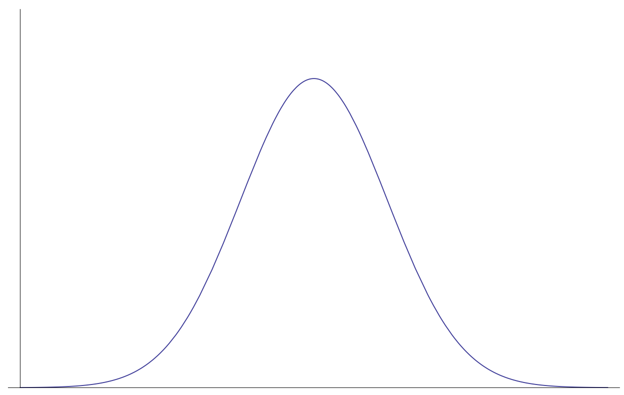


Figure 3.8: If the sample mean is exceptionally large, at a certain level of significance, then we can reject H_0 at that level of significance.

In fact, these 'confidence intervals' are even easier to construct:

Formula: One-Sided 'Confidence Limits'

Let X be a random variable with mean μ and standard deviation σ. Then for samples of size $n > 30$, $P\%$ of the sample means lie in the range:

$$\bar{x} \leq \mu + z\frac{\sigma}{\sqrt{n}} \tag{3.22}$$

where z is such that $\mathbb{P}[Z \geq z] = P\%$.

Also $P\%$ of the sample means lie in the range:

$$\bar{x} \geq \mu - z\frac{\sigma}{\sqrt{n}} \tag{3.23}$$

where z is such that $\mathbb{P}[Z \geq z] = P\%$.

Winter 2011: Question 5 (b) (ii)

(Concrete is claimed to have a specific characteristic strength of 25 N mm^{-2}. If 40 samples gave a mean value of 25.05 N mm^{-2} with a standard deviation of 0.2 N mm^{-2}, test at the 0.05 level of significance that the mean is not equal to 25 N mm^{-2}.)

Also test the hypothesis that the mean is greater than 25 N mm^{-2}.

Solution: First we state the null and alternate hypothesis:

H_0: *the mean is equal to 25*

H_A: *the mean is greater than 25*

Now we set up the 95% 'confidence interval' for the sample means:

$$\bar{x} \leq \bar{\mu} + z\frac{\sigma}{\sqrt{n}}. \tag{3.24}$$

Here we want a 95% 'confidence interval' so we have $z = 1.65$. We take $\mu = 25$, $\sigma = 0.2$ and $n = 40$:

$$\bar{x} \leq 25 + 1.65\frac{0.2}{\sqrt{40}}$$
$$\Rightarrow \leq \bar{x} \leq 25.0522$$

Finally 25.05 lies in this 'confidence interval' so we cannot reject H_0 at the 0.05 level of significance.

Winter 2009: Question 5(b) (ii)

(The breaking strength of a particular metal is estimated to be 43.5 N mm^{-2}. A sample of 100 specimens gave a mean value of 43.4 N mm^{-2}, with a standard deviation of 0.9 N mm^{-2}. Test at the 0.01 level of significance that the mean value is not equal to 43.5 N mm^{-2}.)

Also test that the mean value is less than 43.5 N mm^{-2}.

Solution: First we state the null and alternate hypothesis:

H_0: *the mean is equal to 43.5*

H_A: *the mean is less 43.5*

Now we set up the 99% 'confidence interval' for the sample means:

$$\bar{x} \geq \bar{\mu} - z\frac{\sigma}{\sqrt{n}}. \tag{3.25}$$

Here we want a 99% 'confidence interval' so we have $z = 2.33$. We take $\mu = 43.5$, $\sigma = 0.9$ and $n = 100$:

$$\bar{x} \geq 43.5 + 2.33\frac{0.9}{\sqrt{100}}$$
$$\Rightarrow \geq 43.2903$$

Finally 43.4 lies in this 'confidence interval' so we cannot reject H_0 at the 0.01 level of significance.

Chapter Checklist

1. What range of values can $\mathbb{P}[A]$ take?
2. Under what condition is $\mathbb{P}[A \text{ or } B] = \mathbb{P}[A] + \mathbb{P}[B]$?
3. Under what condition is $\mathbb{P}[A \text{ and } B] = \mathbb{P}[A] \cdot \mathbb{P}[B]$?
4. Under what condition is $\mathbb{P}[\text{not } A] = 1 - \mathbb{P}[A]$?!
5. What kind of data/random variable has a *binomial distribution*?
6. What kind of data/random variable has a *Poisson distribution*?
7. What kind of data/random variable has a *normal distribution*?
8. In what sense is there only one kind of normal data?
9. Why do we sample?
10. What is a *confidence interval* for a population mean?
11. Small samples use t-values and large samples use z-values.
12. Can you describe hypothesis testing to the man on the street?

Exercises:

W '10 Concrete lintels are produced by a machine and it is claimed that the mean length is 2.5 m with a variance of 0.0004 m^2. A sample of one hundred of these lintels gave a mean length of 2.52 m. At the 0.02 level of significance, test that the claim above that the mean is equal to 2.5 m and also test the hypothesis that the mean value is greater than the claimed mean of 2.5 m.

W '09 Concrete lintels are produced by a machine and it is claimed that the mean length is 2.5m with a variance of 0.004 m^2. A sample of fifty of these lintels gave a mean length of 2.48 m. At the 0.01 level of significance test that the claim above is not true. Also test the hypothesis that the mean value is less than the claimed mean of 2.5 m.

A '10 A new type of mould was introduced and a sample of 100 blocks gave a mean length of 450.3 mm. At the 0.01 level of significance test that the mean length has not changed. Also at this level of significance test that the mean length has increased.

W '08 In the past it has been found that the mean life of components was 150 hours. The process of production was changed and a sample of 50 items yielded a mean value of 155 hours with a standard deviation of 12 hours. At the 5% level of significance, test that the mean has changed and test that the mean has increased.

A '09 The lengths of blocks produced were assumed to be normally distributed with a mean value of 450 mm and with a standard deviation of 1.4 mm. A new type of mould was introduced and a sample of 100 blocks gave a mean length of 450.3 mm. At the 0.05 level of significance, test that the mean length has not changed and that the mean has increased.

Chapter 4

Quality Control

4.1 Quality Control Charts

Suppose that $X(t)$ is a temporal random variable that describes some aspect of a system or process that we might be interested in. Examples include the number of absentees on a production line, the pH of a gas, the mass of a unit or the temperature of a room. Many systems display a dominant 'average'-behavior. In terms of $X(t)$, this is what is called the *expected value* or, my own term, the *expected average* of $X(t)$. Most systems are not perfect and there are natural variations about the expected value:

Figure 4.1: On average, the random variable $X(t)$ is equal to μ — but there are natural variations.

There are many reasons why we would like to reduce these natural variations. this might not always be possible. By studying the process a little more we are more likely to identify sources of variation that we *can* control.

If there are some items that are outside of some preset limits or tolerance limits the process is considered to be unstable. Many factors can cause a process to become unstable such as human errors, inferior materials or malfunctioning machinery. If the process is unstable the process is said to be "out of control". If a process is unstable early detection is important. One method of detection would be to use a control chart. One such chart is a Control Chart for sample means. A control chart for sample mean contains two sets of limits (lines). One pair of lines form the outer control limits (O.C.L.) and a second pair form the inner control limits (I.C.L.). A number of samples are taken and for each sample the means are calculated. On this chart we plot the sample means against time:

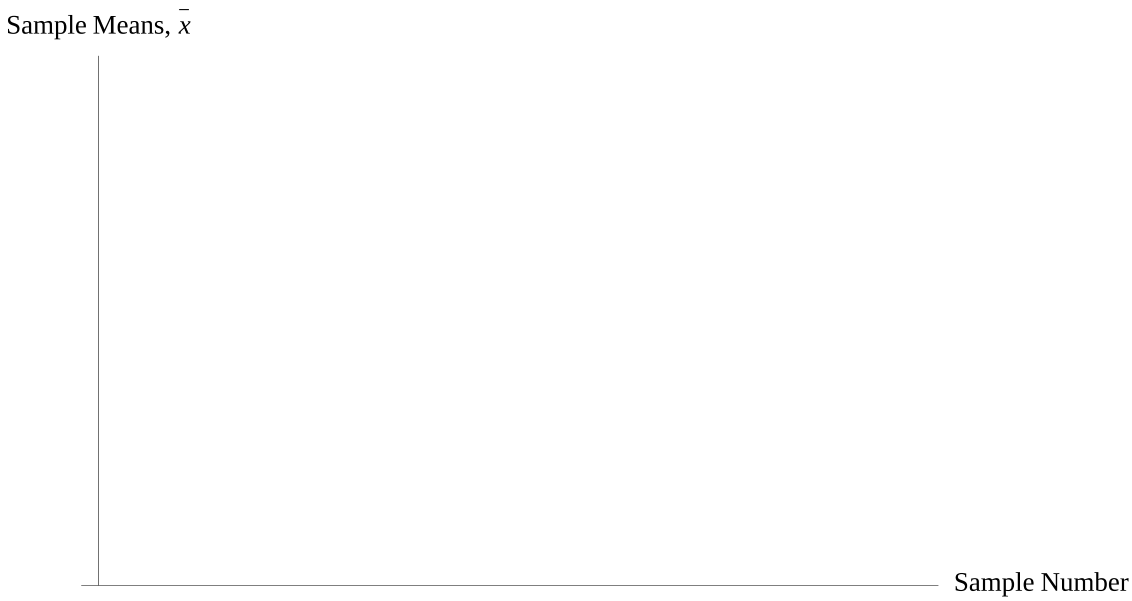


Figure 4.2: We estimate the population mean, μ, by the grand mean $\bar{\bar{x}}$. This is the expected average behaviour. Symmetrically above and below this line lies the *inner control limits* and the *outer control limits*. We then plot the sample means.

99.8% of sample means lie within the outer control limits (O.C.Ls). The chances of a sample mean lying outside of these limits is less than 0.2%, that is 0.1% (one in a thousand) above and 0.1% below the outer control limits. If a sample mean lies outside the outer control limits this is an indication that it is likely that the process is out of control. When this happens it would be necessary to stop the process and adjustments would need to be made. These are also called *action lines.*

In the case of the inner control limits 95% of sample means lie within these limits and the chances of a sample mean lying above these limits is less than 2.5% above and less than 2.5% (1 in 40) below these inner control limits. If two consecutive sample means lie outside of the inner control limits the chances of this occurring is less than 1 in 1600 and it would be necessary to stop the process. These are also called *warning lines.* More on these later but first we must talk about how we might calculate these values.
The limits are based on the 99.8% and 95% confidence limits, respectively. The formulae used are

$$\bar{x} - 3.09\frac{\sigma}{\sqrt{n}} \leq \mu \leq \bar{x} + 3.09\frac{\sigma}{\sqrt{n}} \tag{4.1}$$

for the 99.8% (outer) control limits and

$$\bar{x} - 1.96\frac{\sigma}{\sqrt{n}} \leq \mu \leq \bar{x} + 1.96\frac{\sigma}{\sqrt{n}} \tag{4.2}$$

for the inner (95%) control limits. The value of $\bar{x}$ that is used is the grand mean $\bar{\bar{x}}$ which is the mean value of all samples taken. The values of $3.09\sigma/\sqrt{n}$ and $1.96\sigma/\sqrt{n}$ are approximated by multiplying the mean range by the appropriate value of $A_{0.025}$ and $A_{0.001}$ which are read from a set of tables — as it would be too time consuming to calculate the sample standard deviation.

Formula: Inner and Outer Control Limits

Suppose that a series of samples of size n are made of a process. Suppose further that these samples have mean $\bar{x}_i$ and ranges w_i. If $\bar{\bar{x}}$ is the mean of the means, and $\bar{w}$ the mean of the ranges, then the inner control limits are

$$\bar{\bar{x}} \pm A_{0.025}\bar{w} \tag{4.3}$$

where $A_{0.025}$ is a constant, depending on n, taken from a table.

Suppose that a series of samples of size n are made of a process. Suppose further that these samples have mean $\bar{x}_i$ and ranges w_i. If $\bar{\bar{x}}$ is the mean of the means, and $\bar{w}$ the mean of the ranges, then the inner control limits are

$$\bar{\bar{x}} \pm A_{0.001}\bar{w} \tag{4.4}$$

where $A_{0.001}$ is a constant, depending on n, taken from a table.

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

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: Before we do anything we must fill in the missing data:

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, $\overline{x_i}$	90.50	91.25		89.8				
Ranges, w_i	2.5	2.1		1.6				

Exercises:

A '11 To monitor the breaking strength of precast concrete 6 samples of size 3 were taken and the measurements are recorded below. The table is incomplete

Sample	**1**	**2**	**3**	**4**	**5**	**6**
Strength N mm^{-2}	54.9	55.1	54.8	54.5	54.8	55.0
	54.8	54.9	54.6	54.9	54.6	55.1
	54.4	54.7	54.7	54.7	55.0	54.9
Means, $\overline{x_i}$	54.7	54.9	54.7			55.0
Ranges, w_i	0.5	0.4	0.2			0.2

Set up and plot a control chart for samples means. Include on the chart the mean values of the samples above and comment on the control of the process.

W '10 To monitor the pH of a certain chemical a quality control chart is used. Samples of size 4 were taken and the mass of chemical present in these samples are recorded below.

Sample	**1**	**2**	**3**	**4**	**5**	**6**
Means, x_i	5.35	5.37	5.33	5.34	5.36	5.35
Ranges, w_i	0.04	0.06	0.05	0.03	0.07	0.05

For the data above set up a control chart for sample means. Plot the chart and comment on whether or not the process is under control.

A '10 To monitor the breaking strength of precast concrete 6 samples of size 5 were taken and the measurements are recorded below.

Sample	**1**	**2**	**3**	**4**	**5**	**6**
Means/ N mm^{-2}	54.6	54.5	54.6	54.5	54.6	54.8
Ranges/ N mm^{-2}	0.4	0.3	0.5	0.5	0.4	0.3

Set up and plot a control chart for samples means. Include on the chart the mean values of the samples above and comment on the control of the process.

W '09 To monitor the presence of a certain chemical a quality control chart is used. Samples of size four were taken and the mass of chemical present in these samples are recorded

Sample	**1**	**2**	**3**	**4**	**5**	**6**
Means/ g	3.47	3.53	3.51	3.48	3.50	3.51
Ranges/ g	0.05	0.09	0.12	0.11	0.14	0.09

For the data above set up a control chart for sample means. Plot the chart and comment on whether or not that the process is under control.

W '08 To monitor the breaking strength of precast concrete 5 samples of size 3were taken and the measurements are recorded below.

Sample	**1**	**2**	**3**	**4**	**5**
Means/ N mm^{-2}	54.5	54.7	54.7	54.6	54.5
Ranges/ N mm^{-2}	0.2	0.3	0.2	0.2	0.1

Set up and plot a control chart for sample means. Comment on the process.