SCIE1000

Theory and Practice in Science

Name: _____

Contact details: _____

Eleventh edition, 2018

SCIE1000

The core SCIE1000 content falls into four key areas: thinking, computing, science and mathematics. The content is taught in a highly ``interleaved'' manner rather than in separate blocks, but the figure may help you to appreciate the course goals and content.

Science

- blood alcohol
- fluid flow
- heart disease
- atmospheric CO_2
- temperature
- species abundance
- climate change
- biodiversity
- UV light
- wind chill
- breathing
- daytimes/seasons
- radioactive decay
- pH scale
- breast cancer
- pharmacokinetics
- recreational drugs
- contraception
- alcohol
- diabetes
- glycaemic index
- bioavailability
- bioequivalence
- populations
- unconstrained growth
- constrained growth
- resource management
- cancer
- lifecycle models
- turtles
- predator/prey models
- epidemics
- rubella
- catastrophes

Mathematics

- straight lines
- quadratics
- power functions
- periodic functions
- exponentials
- logarithms
- average rates of change
- derivatives
- Newton's method
- integrals
- areas under curves
- differential equations
- Euler's method
- exponential DE
- logistic DE
- systems of DEs
- Lotka–Volterra model
- SIR models

Computing

- software design
- errors
- input and output
- variables
- calling functions
- conditionals
- loops
- arrays
- plotting
- writing functions

Thinking

- scientific thinking
- modelling
- units
- hypotheses
- history of scientific thinking
- inductive reasoning
- Popperian science
- quantitative reasoning
- medical science
- science in the media

Course rationale

SCIE1000 covers a range of topics in science and mathematics, including

- thinking:
  - how to design, formulate and test models of real-world phenomena;
  - quantitative reasoning and critical evaluation; and
  - the nature of science and scientific thinking
- computer programming in the language Python;
- science, including problems and issues in a range of discipline areas; and
- mathematical techniques.

Rather than requiring you to memorise specific facts, SCIE1000 aims to help you learn how to integrate conceptual, scientific, mathematical and computational techniques, and how to apply these skills to a range of scientific disciplines.

You may find some concepts harder or more interesting than others. Due to time constraints, it is not possible to illustrate every concept with an example from every field of science. Instead, we cover five broad topics: the nature of science and scientific modelling; climate and climate change; scientific thinking; drugs; and life, death and populations. There are also numerous examples from other areas of science: the techniques covered in SCIE1000 are important in *all* areas of science!

Almost every example and case study is based on a research paper, or is a reasonably accurate model of a real situation. Unlike many courses, examples are generally not contrived or "made up". For example, when we derive an equation to model blood alcohol content, the equation genuinely models *real* data. You can estimate **your own** blood alcohol content using the equation.

Throughout your studies, different courses will develop different aspects of your science skills, which together will allow you to graduate with a range of skills and the knowledge necessary to understand and do science. Make sure you appreciate what each of your courses aims to achieve, and hence how your courses fit together and how they differ.

Relationships

We believe that students and lecturers in a course should adhere to the obligations outlined below. Each party should inform the other if they believe that these obligations are not being met.

We will do our best to deliver a course that:

1. contains modern, interesting content from a range of science areas;
2. is relevant to your studies and future professions;
3. is intellectually challenging, accurate and correct;
4. is well-taught, by a team of engaging, professional experts;
5. respects your diverse backgrounds, aspirations and abilities;
6. helps you to improve both your technical knowledge and your generic learning skills;
7. includes assessment that is appropriate, challenging and identifies your level of skills, without being excessive; and
8. provides you with useful, appropriately timed feedback.

We expect that you will do your best to:

1. commit an appropriate amount of time, effort and intellectual engagement to your studies, and submit assessment on time;
2. attend lectures, tutorials and computer laboratory classes, and remain quiet and attentive in class;
3. respect your classmates, the teaching staff and the course content;
4. complete necessary pre-readings before lectures;
5. accept that at times we will cover content that you will find difficult. or of which you may not immediately see the relevance;
6. actively study all components of the course, including science, mathematics, computing and philosophy;
7. not plagiarise from classmates or other sources; and
8. seek help and advice in a timely manner.

Table of Contents

Part 1: Understanding

"Before me there were no created things,
Only eterne, and I eternal last.
All hope abandon, ye who enter in!"

These words in sombre colour I beheld
Written upon the summit of a gate;
Whence I: "Their sense is, Master, hard to me!"

And he to me, as one experienced:
"Here all suspicion needs must be abandoned,
All cowardice must needs be here extinct.

We to the place have come, where I have told thee
Thou shalt behold the people dolorous
Who have foregone the good of intellect."

And after he had laid his hand on mine
With joyful mien, whence I was comforted,
He led me in among the secret things.

Divine Comedy (1308 – 1321), Dante Alighieri (c.1265 – 1321).
(Translation: Henry Wadsworth Longfellow.)

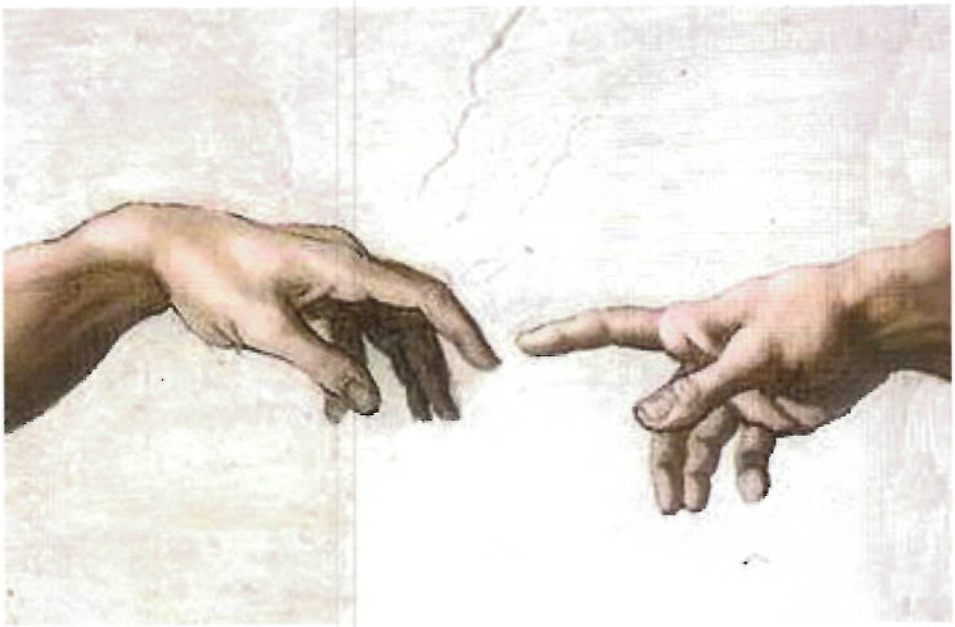


Image 0.1: *The Hands of God and Adam* (1508 – 1512), Michelangelo (1475 – 1564), Sistine Chapel ceiling, Apostolic Palace, Vatican. (Source: en.wikipedia.org)

SCIE1000 is divided into five broad parts.

1. "Understanding" covers what models are, and why they are useful in science;
2. "Getting hot" examines a range of mathematical modelling tools in the context of climate and climate change;
3. "Thinking" develops an understanding of scientific thinking and knowledge;
4. "Drugs" studies rates of change and areas under curves, in a pharmacological context; and
5. "Life and death" explores differential equations in the context of organisms and populations.

Each part focuses on a specific important scientific context, and covers some of the tools needed to develop mathematical and computing models of phenomena within that context. This is Part 1, "Understanding".

The purpose of science is to **understand**, **explain**, **predict** and **influence** phenomena. SCIE1000 aims to help you to develop a range of relevant skills, integrating aspects of science, philosophy, mathematics and computing.

The first part of the notes comprises two chapters. The first chapter gives a general introduction to SCIE1000, some expectations that we have of you and you (should) have of us, then a quick discussion of some key skills required to do science, a course rationale, and then some important background information about the importance of units when measuring something, programming and errors.

The second chapter introduces the concept of *modelling* in science, and five different types of model. We will develop and apply some models, primarily in a medical context. We will conclude with a quick discussion of units, programming and errors.

Do not attempt to memorise specific details associated with any of the contexts we cover. Instead, focus on applying the underlying concepts to novel situations, understand how to think critically, develop and apply models, and interpret your results.

Chapter 1: **A short discussion of nearly everything**

Gaudeamus igitur, Iuvenes dum sumus
Post iucundam iuventutem, Post molestam senectutem
Nos habebit humus, Nos habebit humus.
Vivat Academia, Vivant professores
Vivat membrum quodlibet, Vivant membra quaelibet
Semper sint in flore, Semper sint in flore.

Artist: *Traditional.* Song: *Gaudeamus.*

Image 1.1: *The School of Athens* (1510 – 1511), Raphael (1483 – 1520), Stanze di Raffaello, Apostolic Palace, Vatican. (Source: en.wikipedia.org.)

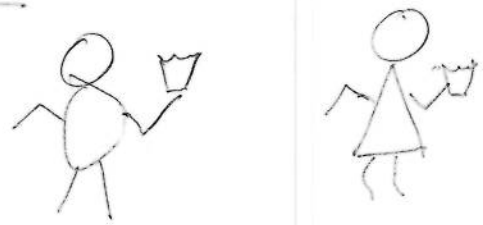
* To emphasise that science and knowledge play fundamental roles in human history, culture and society, the notes include scientifically relevant cultural experiences. *The School of Athens* depicts some famous scientists, mathematicians and philosophers, including Plato, Aristotle, Euclid, Socrates and Pythagoras.

1.1 A taste for SCIE1000: Bloody alcohol

Question 1.1.1

How could you model the blood alcohol content of a person after he or she consumes a quantity of alcohol?

Visualise system



Dependent variable, BAC = Blood Alcohol Concentration

	Independent variable	Dependent
#drinks	n ↑	BAC ↑
Mass	M ↑	BAC ↓ ♂ = 70% H_2O, ♀ = 60% H_2O
Time	t ↑	BAC ↓

So, $BAC = n - M - t$ or $BAC = \frac{n}{M}t$

or $BAC = \frac{n}{M} - t$

$BAC = n - \frac{t}{M}$ (as M ↑ BAC ↑) ✗

1.2 Units

- Every physical quantity must have units unless it is a pure number (such as 2 or π). For example, if $x = 3$ m then x is a length, but if $y = 3$ then y is just a number. **These two things are different.**
- The consequences of using inconsistent units can be severe.

Example 1.2.1

The Mars Climate Orbiter was launched in 1998 as part of a $USD330 million project, but in September 1999 it crashed into Mars. Here is an extract from the report into the accident [34]:

> "During the 9-month journey from Earth to Mars, propulsion maneuvers were periodically performed... coupled with the fact that the angular momentum (impulse) data was in English, rather than metric, units, resulted in small errors being introduced in the trajectory estimate over the course of the 9-month journey. At the time of Mars insertion, the spacecraft trajectory was approximately 170 km lower than planned...
>
> ...it was discovered that the small forces ΔVs reported by the spacecraft engineers for use in orbit determination solutions was low by a factor of 4.45 (1 pound force = 4.45 Newtons) because the impulse bit data contained in the AMD file was delivered in lb-sec instead of the specified and expected units of Newton-sec."

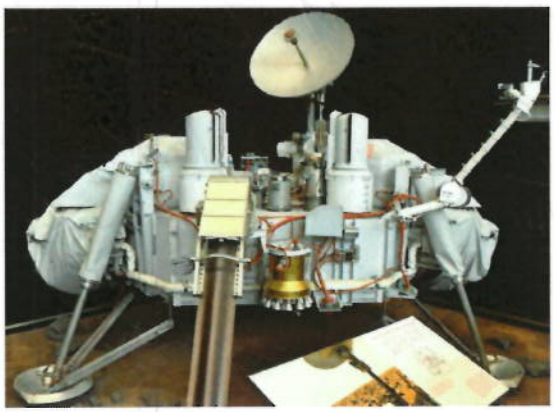


Photo 1.1: Mars Lander (proof test model) from the Viking program, launched 1975. (Source: PA.)

SI Units and prefixes

The most commonly used units of measurement are defined by the **International System of units**, and are called **SI** units. There are seven **SI base units**; their standard names and symbols are shown in Figure 1.1. Figure 1.2 shows the **SI prefixes** that denote multiples of the SI units.

Base quantity	SI unit name	Symbol
length	metre	m
mass	kilogram	kg
time	second	s
electric current	ampere	A
thermodynamic temperature	kelvin	K
amount of substance	mole	mol
luminous intensity	candela	cd

Figure 1.1: Names and symbols of the seven SI base units.

Multiple	Prefix	Symbol	Multiple	Prefix	Symbol
10^{1}	deca	da	10^{-1}	deci	d
10^{2}	hecto	h	10^{-2}	centi	c
10^{3}	kilo	k	10^{-3}	milli	m
10^{6}	mega	M	10^{-6}	micro	μ
10^{9}	giga	G	10^{-9}	nano	n
10^{12}	tera	T	10^{-12}	pico	p
10^{15}	peta	P	10^{-15}	femto	f
10^{18}	exa	E	10^{-18}	atto	a
10^{21}	zetta	Z	10^{-21}	zepto	z
10^{24}	yotta	Y	10^{-24}	yocto	y

Figure 1.2: The 20 SI prefixes.

Derived units

Many natural and scientific quantities require more complex units than SI base units. These **can always be defined** in terms of the seven base units, and are called **SI derived units**.

Mathematical notation for SI derived units

There is a convenient standard mathematical notation for SI derived units, based on the following principles:

- if the quantity involves the mathematical "product" of two SI units, then we separate their SI symbols by a space or dot;
- if the same SI unit occurs in a "product" more than once then we use mathematical power notation; and
- if the quantity involves the "quotients" of SI units then for the derived unit we use either a quotient sign /, or (more often) mathematical power notation with a negative power.

Example 1.2.2

Some examples of quantities with their SI derived units are:

- Volume, measured in m^3 (or L, where 1 L is defined to be 10^{-3} m^3).
- Concentration, measured in mol/L or mol L^{-1} or mol $\cdot$ L^{-1}.

Example 1.2.3

Some frequently-used SI derived units have been given special names and symbols. Figure 1.3 shows some well-known examples.

Quantity	Name	Symbol	SI units	SI base units
frequency	hertz	Hz	-	s^{-1}
force	newton	N	-	$m \cdot kg \cdot s^{-2}$
pressure, stress	pascal	Pa	$N \cdot m^{-2}$	$m^{-1} \cdot kg \cdot s^{-2}$
energy, work, quantity of heat	joule	J	$N \cdot m$	$m^2 \cdot kg \cdot s^{-2}$
power, radiant flux	watt	W	$J \cdot s^{-1}$	$m^2 \cdot kg \cdot s^{-3}$
electric potential difference, electromotive force	volt	V	$W \cdot A^{-1}$	$m^2 \cdot kg \cdot s^{-3} \cdot A^{-1}$
Celsius temperature	degree Celsius	°C	-	K

Figure 1.3: Some well-known units and their SI base units.

Dimensional analysis

Dimensional analysis is an important technique in science:

- An equation describing a physical situation can be true *only* if it is **dimensionally homogeneous**; that is, both sides of the equation have the same units.
- Units can be multiplied and cancelled (divided).
- If quantities have the same units then they can be added or subtracted.

Dimensional analysis allows a quick check of whether a calculation is 'plausible': if the dimensions do not match, then there **must** be an error.

The importance of units

In scientific work, you should always deal correctly with units. Sometimes, when learning new concepts, it can make things seem more complicated if units are included. Sometimes in SCIE1000 we define variables so as to not require units. For example, consider two correct definitions.

- "Let t be the time since the rocket was launched"; and
- "Let t be the number of seconds since the rocket was launched".

If the first definition is used, then t requires a unit (such as seconds) in all calculations. If the second is used then t does not require a unit.

Question 1.2.4

The *cardiac output* CO of a heart is the volume of blood ejected by the heart during a particular time period, and equals the *stroke volume* SV multiplied by the *heart rate* HR. If $SV = 70$ mL and the heart beats 1.5 times each second, calculate CO in L min^{-1}.

$CO = SV \times HR = 70\ \text{mL} \times 1.5\ \text{beats/s}$

$= 105\ \text{mL/s}$

$= 105\ \text{mL/s} \times \frac{1\ \text{L}}{1000\ \text{mL}} \times \frac{60\ \text{s}}{1\ \text{min}} = 6.3\ \text{L/min}$

1.3 Programming

- Computation is important when formulating and applying models, particularly when dealing with complex phenomena.
- A *program* is a set of instructions that make a computer do something. Web browsers (for example, *Internet Explorer*), word processors (*Microsoft Word*) and "apps" for a mobile phone are all familiar examples of programs. Even the Python programming language itself is a program.
- In SCIE1000, we will only write short, relatively simple programs to model some phenomena. However, even when programs are not complex, you should always be guided by a number of "good programming" principles.

Good programming

In general, "good" computer programs are:

- correct
- simple
- well-documented
- easy to read
- efficient
- easy to use
- easy to understand
- thoroughly tested

- To assist with achieving these goals, programs should:
  - include blank lines and spacing to assist readability;
  - have extensive comments to explain what is happening; and
  - use *meaningful names* for variables and functions.
- The first step in programming is to understand exactly what problem is being solved, and hence specify **exactly** what the program should do; specifications should be precise, accurate and complete.
- Once the problem has been understood, the programmer needs to write a sequence of commands that together solve the problem.
- Programming requires technical skill, experience and creativity.

1.4 Errors in programs

- Even the best and most experienced computer programmers will sometimes (even often) write programs with errors in them. The consequences of software errors (*bugs*) can be very serious.

Example 1.4.1

In 2010, Toyota recalled around 500,000 hybrid-fuel cars to repair software errors that could cause the braking system to fail. There were fears that the error could lead to a "diplomatic incident" between the USA and Japan.

- Types of programming error include: incomplete (or incorrect) problem description, design faults in the software, unanticipated 'special cases', coding errors, logic errors and miscommunication between programmers.

Testing and debugging

Most newly written programs include errors, and it is important to adopt a systematic approach to minimising the number of errors, then identifying and fixing any that occur. This process is called *testing* and *debugging*.

Some programming errors are easy to find (such as missing brackets), some will result in runtime error messages (for example, trying to divide by zero), but in other cases a program may produce incorrect output without an error message. To find such errors, you will need to test your program with different input values, and check the output "by hand".

When writing programs, make sure that you:

- Think about the clearest and most logical way to solve the problem;
- Write your program in an organised, systematic manner.
- Include useful comments in the program.
- Test your programs on a range of data;
- Check some output carefully to make sure it is correct; and
- Pay attention to any error messages!

Chapter 2: **A career in modelling**

I'm very well acquainted, too, with matters mathematical,
I understand equations, both the simple and quadratical,
About binomial theorem I'm teeming with a lot o' news,
With many cheerful facts about the square of the hypotenuse.
I'm very good at integral and differential calculus;
I know the scientific names of beings animalculous:
In short, in matters vegetable, animal, and mineral,
I am the very model of a modern Major-General.

Artist: *Gilbert and Sullivan.* Song: *Major general song.*

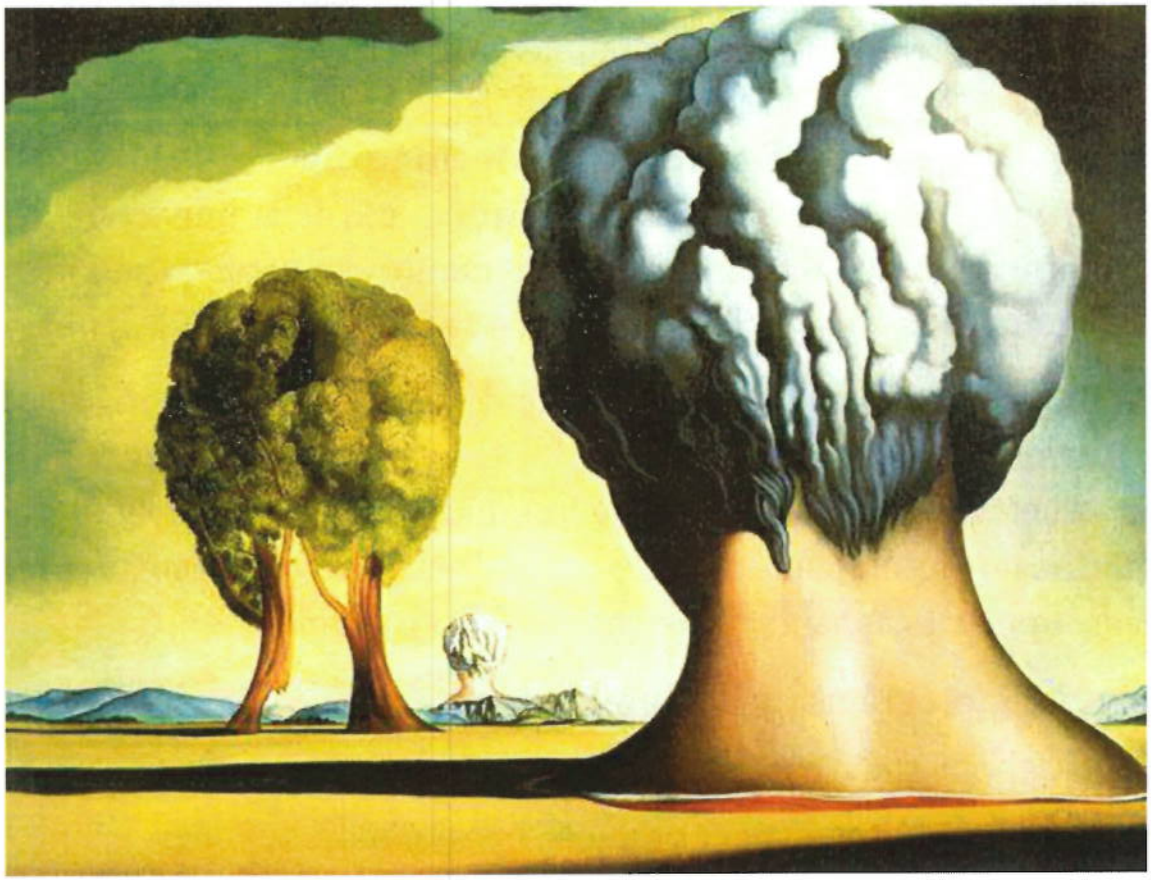


Image 2.1: *The Three Sphinxes of Bikini* (1947), Salvadore Dali (1904 – 1989), Morohashi Museum of Modern Art. (Source: Museum publication.)

2.1 Science's next top model

Science

Science aims to **understand**, **explain**, **predict** and **influence** phenomena. Understanding science, and thinking in a 'scientific manner', requires:

- *discipline knowledge and content* – the language, information, knowledge and skills specific to a discipline;
- *scientific thinking and logic* – the conceptual process of performing systematic investigations, hypothesising, thinking critically and defensibly, and making valid deductions and inferences;
- *communication and collaboration* – the process of working with others, sharing information and resources;
- *curiosity, creativity and persistence* – the relatively intangible characteristics that include the ability to ask and answer 'interesting' questions, and solve difficult problems in novel ways;
- *observation and data collection* – the processes and techniques used to collect useful data about particular phenomena;
- *modelling and analysis* – the process of developing conceptual representations of phenomena, then using approximation, mathematics, statistics and computation in order to allow predictions to be made.

- Most science is intrinsically quantitative, because quantifying phenomena allows us to measure, describe and compare them efficiently and precisely.
- Science often proceeds by:
  - observing and measuring values, such as the amount, frequency, magnitude, duration or rate of some phenomenon; and
  - Using these data and scientific thinking to answer predictive questions about that phenomenon, such as
    - "What will happen if ...?"
    - "How can ...?"
    - "What causes ...?"
    - "Why does ...?"

Models

Models are simplifications and approximations of reality, usually based on measured data, that allow us to understand phenomena, make predictions and evaluate possible impacts of interventions. All models need to strike an appropriate balance between accuracy and complexity.

- Models can be *physical* or *conceptual*. Examples of physical models include building scale models of bridges or dams, and subjecting the model to tests. In SCIE1000, we will focus on conceptual models.

- *Statistics* is fundamental to modelling, allowing models to be developed from uncertain, imprecise data, even when there is not a strong body of underlying theory (for example, physics has many theoretical models, but much of biology does not).

- Ways of developing 'appropriate' models include:

  - using "common sense" and logical deduction;
  - using existing knowledge of similar phenomena; and
  - observing measured data and seeing what they "look like" (many phenomena change according to simple underlying patterns, such as at a constant rate or at a rate proportional to the current value).

- The five common ways of developing and presenting (conceptual) quantitative models are: words (descriptive text); values (for example, weight/height/ag tables); pictures (such as graphs or flow diagrams); mathematics (using equations); and computer programs.

- Note that there is nothing "right" or "wrong" about each approach – each is suited to different uses and/or target audiences. Most models can be developed and presented in **all** of these ways.

- In SCIE1000 we will use all five methods, but will focus on the final two.

$$CO = SV \times HR$$

$$= 70\ ml/beat \times 1.5\ beats/s$$

$$= 105\ ml/s$$

$$= 105\ ml/s \times \frac{1\ L}{1000\ ml} \times \frac{60\ s}{min}$$

$$= 6.3\ L/min$$

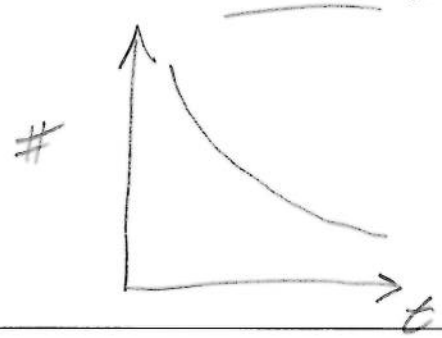
2.2 Mathematics and models

- Some people view mathematics as being an abstract process, unlike disciplines such as biology or chemistry that relate directly to the 'real' world.
- This view ignores the many links between mathematics and science :
  - Scientists use discipline knowledge and a special language to describe nature and the real world (for example, biologists use taxonomic categories, anatomical descriptions and medical terminology).
  - Mathematicians also use discipline knowledge and a special language to describe nature and the real world (for example, exponential and linear functions all describe relationships between natural phenomena).
- Science and mathematics are not separate: often, they are so closely interlinked that they are indistinguishable!

Question 2.2.1

Later we will see that some elements undergo *radioactive decay*. Consider an experiment in which the quantity of undecayed radioactive material is measured, a graph is drawn, the rate of decay is analysed, and predictions are made about the total quantity of material that will decay during a given time period. Is this phenomenon an example of mathematics or science? What is the difference?

- Maths & science are interconnected
- Use maths to quantify the phenomenon & make calculations/predictions
- Maths is part of science



$A = A_0 e^{-kt}$

Mathematics

Mathematics is a standardised formal language, way of thinking and body of knowledge that allow us to:

- develop models to represent reality;
- increase our understanding of phenomena when we can describe them mathematically;
- perform correct, logical deductions;
- communicate unambiguously; and
- draw conclusions and make predictions.

- Because all areas of personal and professional life include quantitative concepts, *everyone* needs to learn some mathematical language and thinking.
- Proficiency in science requires a higher level of mathematical knowledge and sophistication in its use.
- SCIE1000 aims to develop skills in using mathematics, but we do not study mathematics for its own sake, or develop new mathematical knowledge; if you wish to do that, enrol in discipline-based mathematics courses.
- Instead, we study mathematics **solely** to understand and model the 'real' world. For example, *mathematical functions* are the formal descriptions of patterns in data, *derivatives* describe the change in a value at a point in time, and *integrals* describe behaviour over a period of time.

Question 2.2.2

The following graph shows the measured blood flow rate in coronary artery bypass vein grafts during a single heart beat, after a patient underwent coronary bypass surgery (see [46]).

(continued over)

Question 2.2.2 (continued)

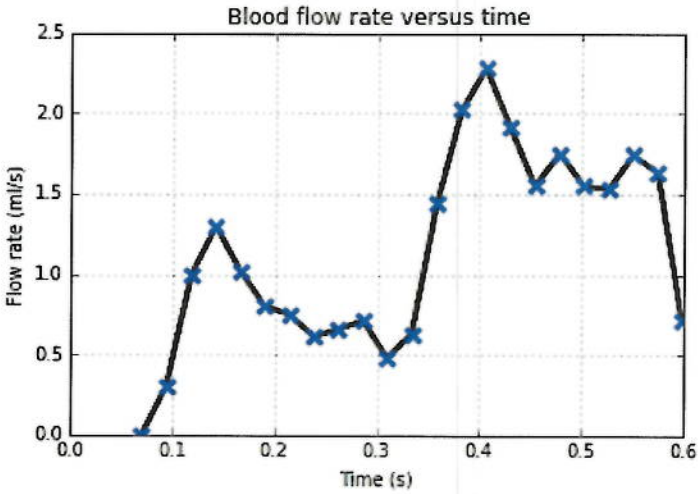


Figure 2.1: Blood flow rate in coronary artery bypass vein grafts, during a single heart beat.

Briefly discuss the meaning and significance of each of:

(a) The *height* or *value* of the graph at a given time.

y-value = flow rate = ml/s = velocity = volume of blood flowing through an artery per second

(b) The *slope* or *derivative* of the graph at a given time.

$\text{Slope} = \frac{\text{rise}}{\text{run}} = \frac{\text{ml/s}}{\text{s}} = \text{ml/s}^2$ = acceleration

∴ Δ in flow rate over time ⇒ diagnose a problem

(c) The *area under the curve* or *integral* of the graph over a time interval.

AUC = ml/s × s = ml = volume.

Total volume of blood thru artery on a heart beat.

(d) Estimate the total blood flow through the vein graft in a heart beat.

AUC ≈ Height × Width = 1.25 × 0.5 = 0.625 ml

- All three quantities identified in Question 2.2.2 have important physical meanings. This occurs for many phenomena and practical applications.
- Mathematics provides a range of logically valid techniques that allow us to deduce information that we cannot measure or obtain in other ways (due to physical, ethical or financial limitations).

2.3 Modelling in action

Case Study 1: **Let it flow**

- *Fluid dynamics* involves studying liquids and gases that are moving, which is important in many branches of science (particularly geology, environmental science and biomedical science) and engineering.

Question 2.3.1

Develop a model for the flow rate of blood through a given blood vessel. (Hint: which factors are important; do they increase or decrease the rate?)

Q flow rate.

X (indept. variable)	Y (dept. variable)	Why?
$L \uparrow$	$Q \downarrow$	More resistance
$r \uparrow$	$Q \uparrow$	Less " "
Viscosity $\mu \uparrow$	$Q \downarrow$	Flows less
ΔP pressure $\uparrow$	$Q \uparrow$	Greater force per unit area

So $Q = \Delta P + r - L - \mu$

OR $Q = \frac{\Delta P r}{\mu L}$ OR $Q = \frac{\Delta P \pi r^2}{\mu L}$

The following formula (called the *Hagen-Poiseuille equation*) is often used to estimate such flows:

$$Q = \frac{\Delta P \pi r^4}{8 \mu L}$$

r^4 ← related to distance from wall + turbulence

Compare your formula with the Hagen-Poiseuille equation.

Very similar — congratulations!

- High levels of certain types of cholesterol in the blood can lead to blockages in coronary arteries, which can eventually lead to a heart attack.
- During a heart attack, a lack of blood supply causes heart muscle tissue to die and the dead tissue is replaced with scar tissue.

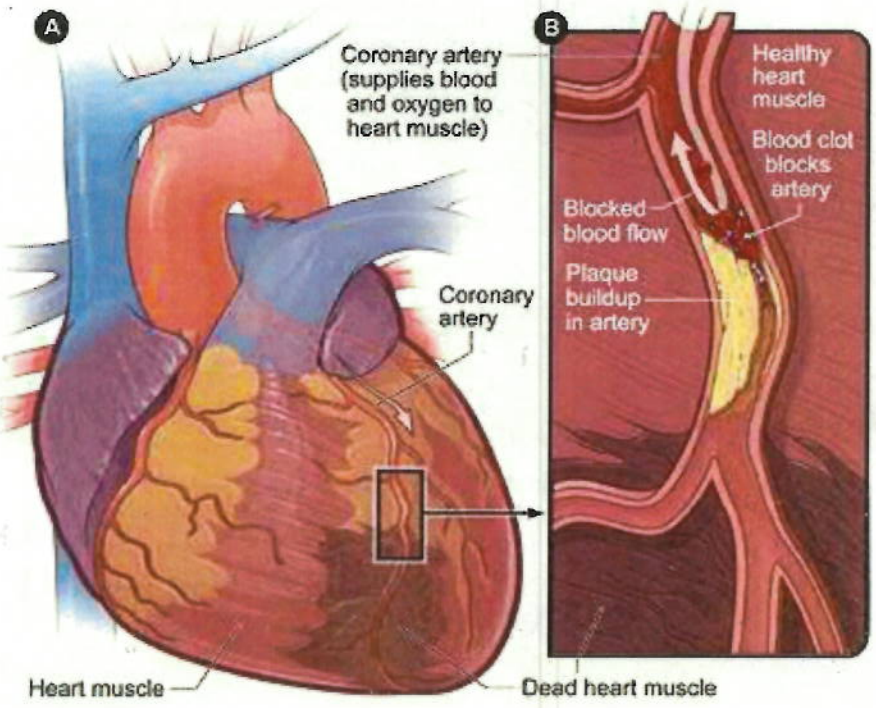


Figure 2.2: Left: heart and coronary artery showing dead heart muscle caused by a heart attack. Right: longitudinal section of a coronary artery with plaque buildup and a blood clot. (Source: www.nhlbi.nih.gov.)

- One surgical method of increasing blood flow through partially blocked arteries is an *angioplasty*.
- In a coronary angioplasty, a cardiologist inserts a balloon-tipped catheter under local anaesthetic, typically through the groin or arm.
- When the catheter is correctly positioned within the coronary artery, the doctor inflates the balloon to expand the blood vessel (and sometimes inserts a metallic stent to maintain the expansion).
- Angioplasties are much simpler and less invasive than coronary artery bypass surgery, but have a higher rate of recurrence of the original occlusion.

(*) Play video of operation

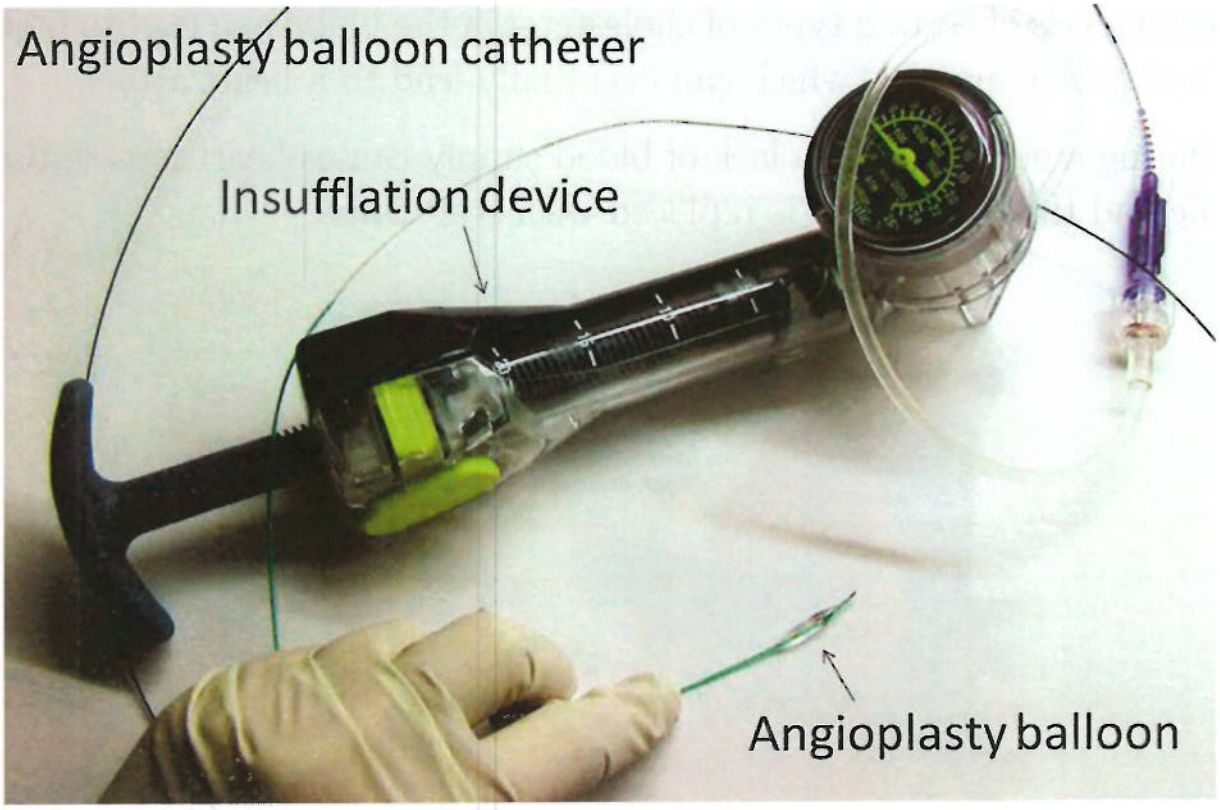


Photo 2.1: Angioplasty balloon catheter. (Source: DM.)

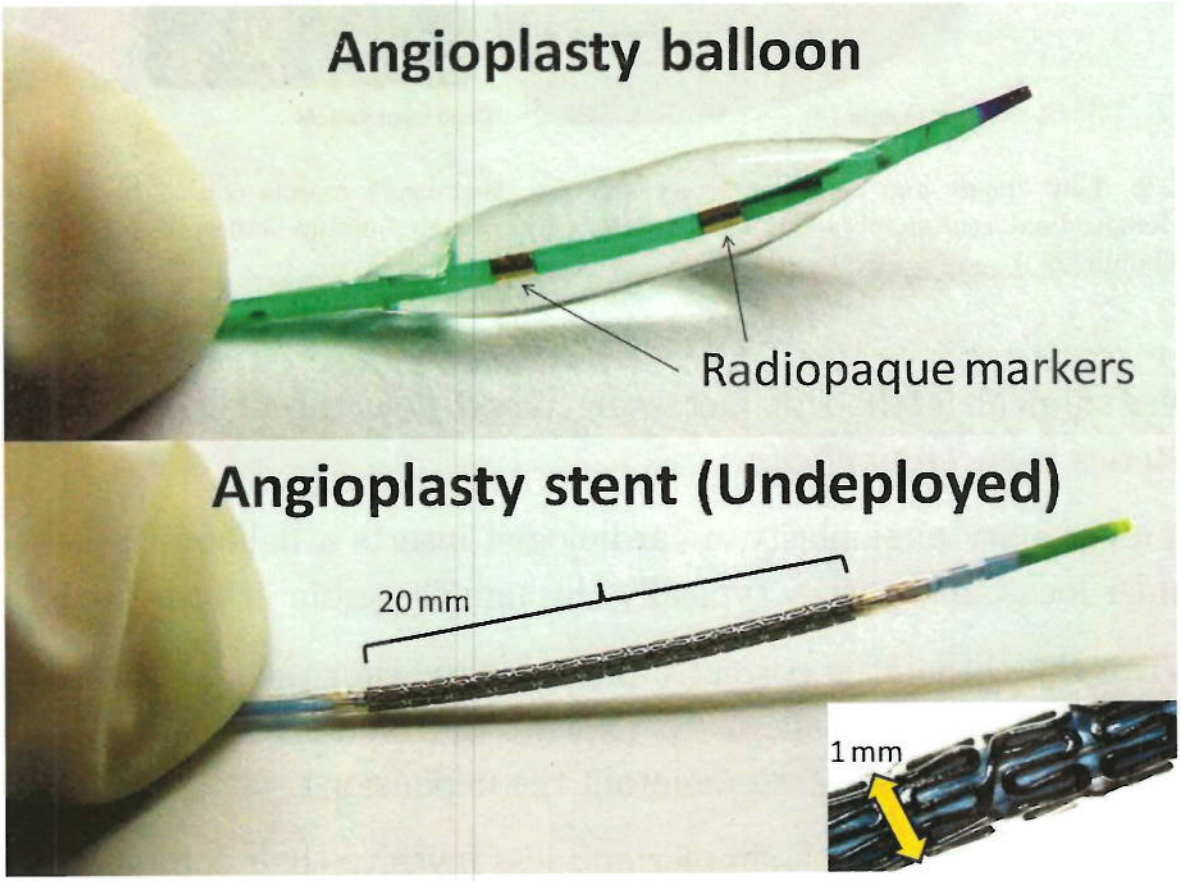


Photo 2.2: Inflated angioplasty balloon and undeployed stent. (Source: DM.)

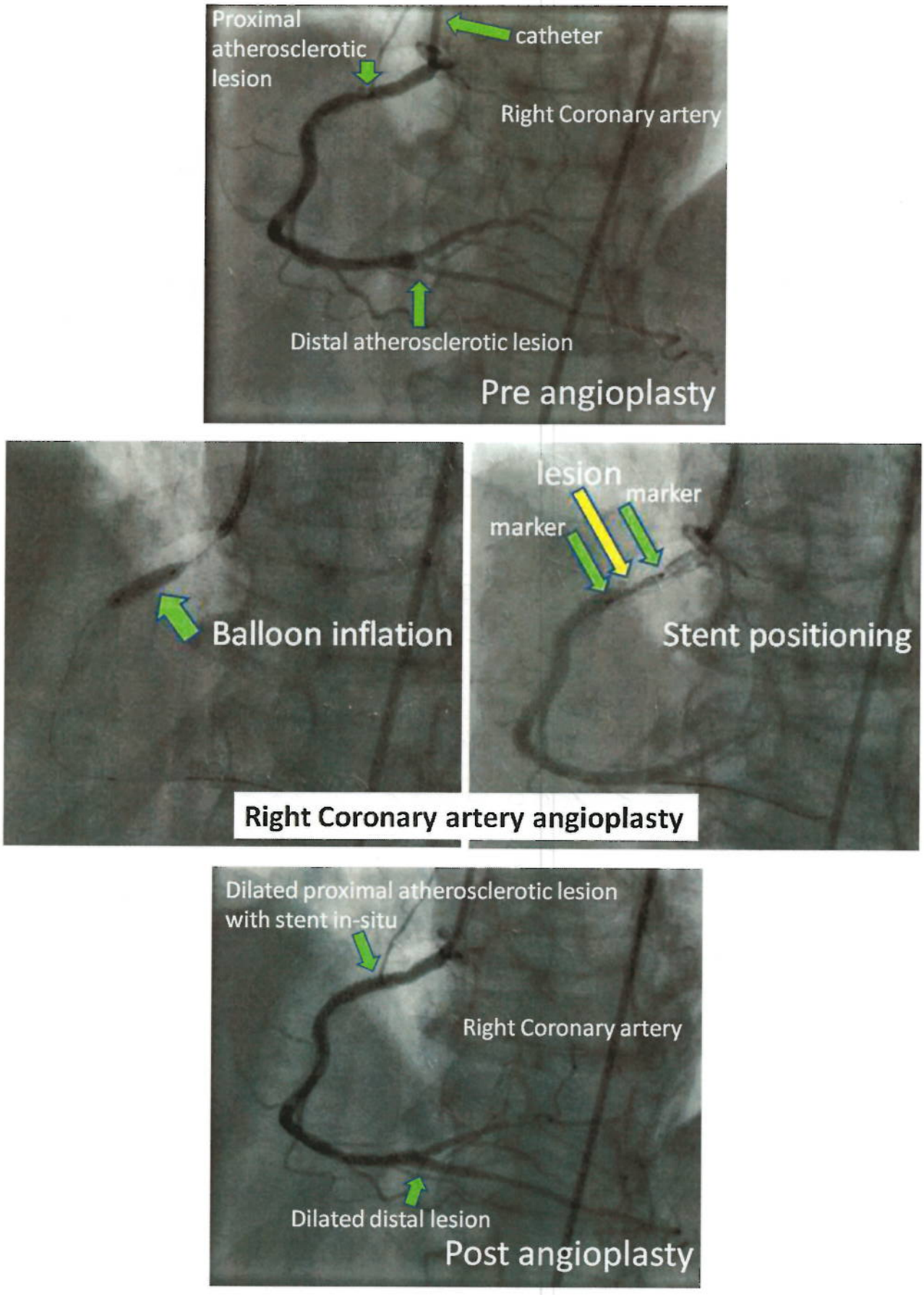


Photo 2.3: Right coronary artery angioplasty. (Source: DM.)

Question 2.3.2

Assume that a patient undergoing an angioplasty procedure shows a 30% increase in the diameter of a partially occluded artery. Use the Hagen-Poiseuille equation to calculate the resulting percentage increase in blood flow rate through that artery, and interpret your answer.

$Q = \dfrac{\Delta P \pi r^4}{8\mu L}$ 30% ↑ in diameter

radius = ? Discuss...

$r \Rightarrow 1.3r$

$$Q = \frac{\Delta P \pi (1.3r)^4}{8\mu L}$$

$$= \frac{\Delta P \pi (1.3)^4 r^4}{8\mu L}$$

So Q ↑ by $(1.3)^4 = 2.8561$

$\Rightarrow$ 186% improvement

OR

$$\frac{Q_{new}}{Q_{old}} = \frac{\dfrac{\Delta P \pi (1.3r)^4}{8\mu L}}{\dfrac{\Delta P \pi r^4}{8\mu L}} = (1.3)^4$$

End of Case Study 1: Let it flow.

Case Study 2: To the heart of the matter

- Diseases of the circulatory system (including heart disease and stroke) are the leading cause of death in many western societies.

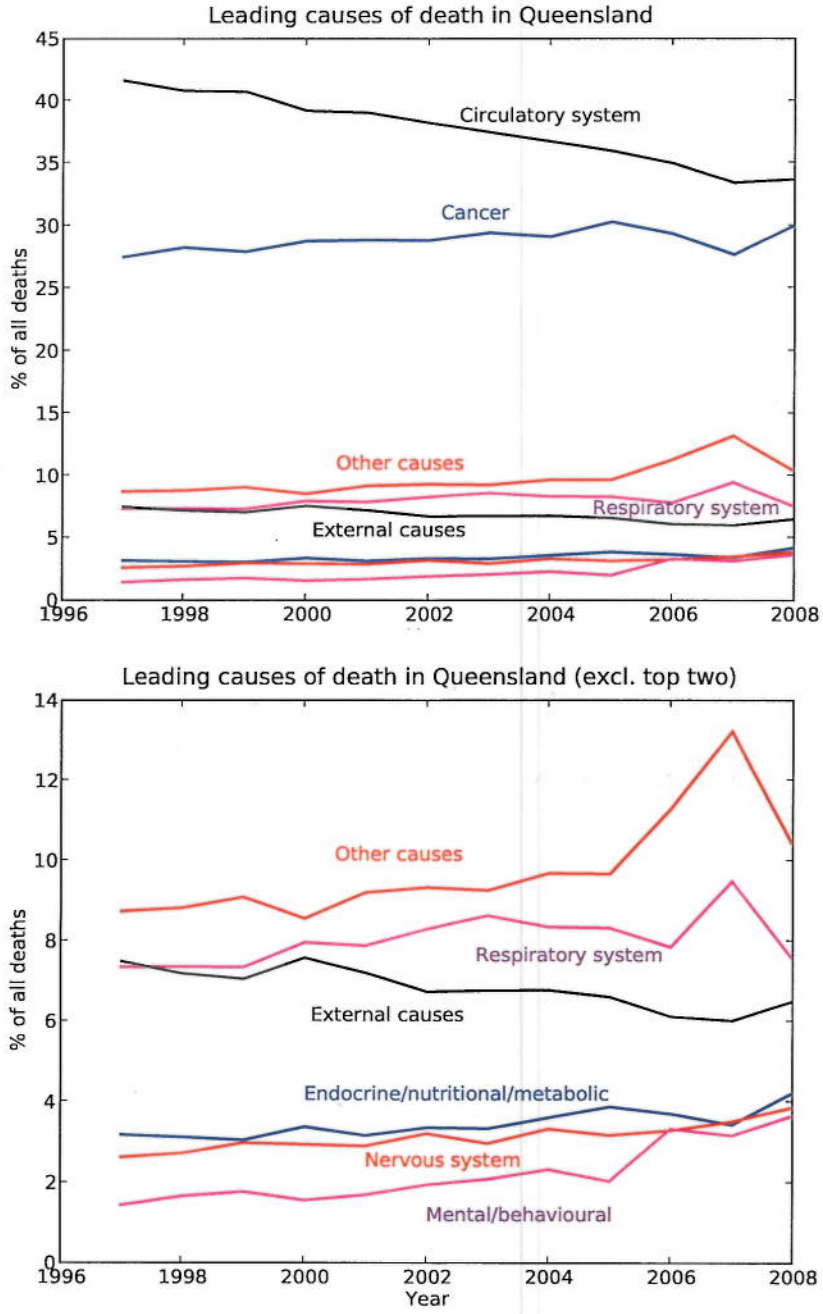


Figure 2.3: Leading causes of death in Queensland. The lower graph reproduces data from the top graph, for all except the two largest causes of death. (Data source: Qld government.)

- Individuals, doctors and public health bodies all have an obvious interest in predicting the risk of suffering cardiovascular disease.

- In medicine and population health, risks are often specified as a probability of an identified event occurring in a given time period.
- Shortly we will encounter a famous, long-running study into cardiovascular health, called the *Framingham study*[1]. The study defines Coronary Heart Disease (CHD) as including:
  - *angina pectoris*, which is severe chest pain caused by a lack of blood to heart muscle;
  - *myocardial infarction*, commonly called a heart attack, arising from complete loss of blood supply to heart muscle; and
  - death due to cardiac arrest.
- CHD is most often caused by *atherosclerosis*, which is a blockage of a coronary artery supplying blood to heart muscle tissue.
- Photograph 2.4 shows a calf heart, with coronary arteries clearly visible.

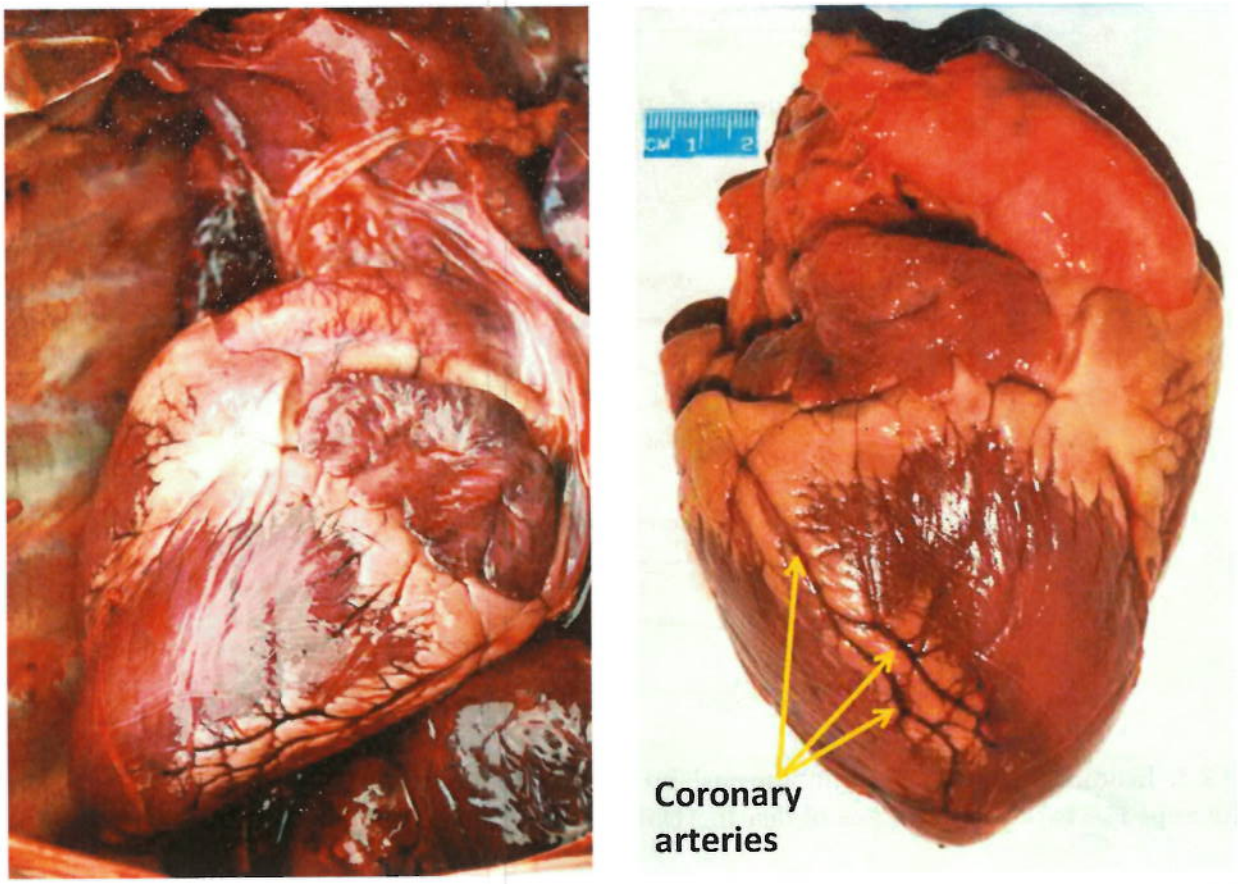


Photo 2.4: Left: calf heart in-situ. Right: calf heart showing coronary arteries. (Source: PA.)

[1]All information from the Framingham study has been reproduced with permission from the National Heart, Lung, and Blood Institute as a part of the National Institutes of Health and the U.S. Department of Health and Human Services.

Question 2.3.3

Which factors or data are crucial when developing a model for estimating the likelihood that a person will suffer from CHD in the next 10 years? Does each factor increase or decrease the risk?

Factor	Likelihood CHD
Obesity ↑	Risk ↑
Stress ↑	Risk ↑
Smoking ↑	Risk ↑
Drinking ↑	Risk ↑
Diabetes ↑	Risk ↑
Cholesterol ↑	Risk ↑
Genetics (family history) ↑	Risk ↑
Gender ♀	Risk ↓

What is your "gut feeling" of the likelihood that your lecturer will suffer from CHD in the next 10 years?

Students usually give high estimates

- Until comparatively recently, little was known about the general causes of heart disease and stroke, although the rates of cardiovascular disease (CVD) in many societies had been rising for some time.
- In 1948, a study into heart disease commenced in Framingham, Massachusetts, which has become one of the best-known longitudinal health studies.
- The Framingham study (which continues today) has monitored the cardiovascular health of participants, identified a range of risk factors for CHD and included these factors in a mathematical risk model.

- One of the resources produced from the Framingham Study is a CHD Risk Prediction score sheet, used to predict the likelihood that a person will suffer CHD in the next ten years.

Estimate this for each lecturer Guess LDL/HDL if you don't know it

Step 1: Age

Age (Years)	Points Female	Points Male
30-34	-9	-1
35-39	-4	0
40-44	0	1
45-49	3	2
50-54	6	3
55-59	7	4
60-64	8	5
65-69	8	6
70-74	8	7

Step 2: LDL cholesterol

LDL (mmol/L)	Points Female	Points Male
≤ 2.59	-2	-3
2.60-3.36	0	0
3.37-4.14	0	0
4.19-4.91	2	1
≥ 4.92	2	2

Key	
Colour	Risk
green	very low
white	low
yellow	moderate
rose	high
red	very high

Step 3: HDL cholesterol

HDL (mmol/L)	Points Female	Points Male
≤ 0.9	5	2
0.91-1.16	2	1
1.17-1.29	1	0
1.30-1.55	0	0
≥ 1.56	-2	-1

Step 4: Blood pressure (F: Female, M: Male)

Systolic (mm Hg)	Diastolic (mm Hg) < 80 Points	80-84 Points	85-89 Points	90-99 Points	≥ 100 Points
≤ 120	F: -3 M: 0				
120-129		0			
130-139			F: 0 M: 1		
140-159				2	
≥ 160					3

Note: When systolic and diastolic pressures provide different estimates for point scores, use the higher number.

Step 5: Diabetes

Diabetes	Points Female	Points Male
No	0	0
Yes	4	2

Step 6: Smoker

Smoker	Points Female	Points Male
No	0	0
Yes	2	2

Step 7: Sum points from Steps 1-6

Category	Points
Age	
LDL	
HDL	
Blood pressure	
Diabetes	
Smoker	
Point total	

Step 8: Determine risk from point total

Point total	10 Year CHD risk Female	Male
≤ -3	1%	1%
-2	1%	2%
-1	2%	2%
0	2%	3%
1	2%	4%
2	3%	4%
3	3%	6%
4	4%	7%
5	5%	9%
6	6%	11%
7	7%	14%
8	8%	18%
9	9%	22%
10	11%	27%
11	13%	33%
12	15%	40%
13	17%	47%
14	20%	≥ 56%
15	24%	≥ 56%
16	27%	≥ 56%
≥ 17	≥ 32%	≥ 56%

Step 9: Compare to others of the same age

Age (Years)	Average 10 Yr risk	Low 10 Yr risk
30-34	F: <1% M: 3%	F: <1% M: 2%
35-39	F: 1% M: 5%	F: <1% M: 3%
40-44	F: 2% M: 7%	F: 2% M: 4%
45-49	F: 5% M: 11%	F: 3% M: 4%
50-54	F: 8% M: 14%	F: 5% M: 6%
55-59	F: 12% M: 16%	F: 7% M: 7%
60-64	F: 12% M: 21%	F: 8% M: 9%
65-69	F: 13% M: 25%	F: 8% M: 11%
70-74	F: 14% M: 30%	F: 8% M: 14%

Note: low risk was calculated for an individual of the same age, with normal blood pressure, LDL 2.60-3.36 mmol/L, HDL 1.45 mmol/L, non-smoker and no diabetes.

Figure 2.4: Framingham CHD risk assessment sheet for males and females

Question 2.3.4

Use the Framingham CHD risk assessment sheet in Figure 2.4 to estimate the probability that your lecturer will suffer CHD within 10 years. Compare this with your answers to Question 2.3.3.

Ans: ~10%

Question 2.3.5

Briefly discuss some key points highlighted by the risk prediction sheet. (You may wish to mention such things as the comparative impact of different risk factors, some 'risk factors' commonly mentioned in the media that are not included, and some differences between males and females.)

↑ risk in ♂

↑ risk in ♀ in menopause

age > smoking

Low blood pressure: no effect ♂ but ↓risk ♀

High HDL cholesterol protective (-ve values)

Age & gender you can't do anything about

End of Case Study 2: To the heart of the matter.

Case Study 3: **Get your BAC up**

- Blood Alcohol Concentration (BAC) is usually measured as the percentage of total blood volume which is alcohol.
- Figure 2.5 (see [51]) shows some physical and behavioural effects typically experienced after different levels of alcohol consumption.
- There are strict laws about driving and operating machinery after consuming alcohol.
- In Australia the maximum legal BAC for driving is 0.05%, or 0.5 g/L.

Stages	BAC	Likely Effects
Feeling of well-being	Up to .05%	Talkative; Relaxed More confident
At-risk	.05−.08%	Talkative Acts and feels self-confident Judgment and movement impaired Inhibitions reduced
Risky state	.08−.15%	Speech slurred Balance and coordination impaired Reflexes slowed Visual attention impaired Unstable emotions Nausea, vomiting
High-risk state	.15−.30%	Unable to walk without help Apathetic, sleepy Laboured breathing Unable to remember events Loss of bladder control Possible loss of consciousness
Death	Over .30%	Coma; Death

Figure 2.5: Typical physical and behavioural effects of alcohol at various BAC levels.

Extension 2.3.6 (from [51])

"**Intoxication risks**: Intoxication is the most common cause of alcohol-related problems, leading to injuries and premature deaths. As a result, intoxication accounts for two-thirds of the years of life lost from drinking. Alcohol is responsible for:

- 30% of road accidents
- 44% of fire injuries
- 34% of falls and drownings
- 16% of child abuse cases
- 12% of suicides
- 10% of industrial accidents

As well as deaths, short-term effects of alcohol result in illness and loss of work productivity (e.g. hangovers, drink driving offences). In addition, alcohol contributes to criminal behaviour – in Australia over 70% of prisoners convicted of violent assaults have drunk alcohol before committing the offence and more than 40% of domestic violence incidents involve alcohol."

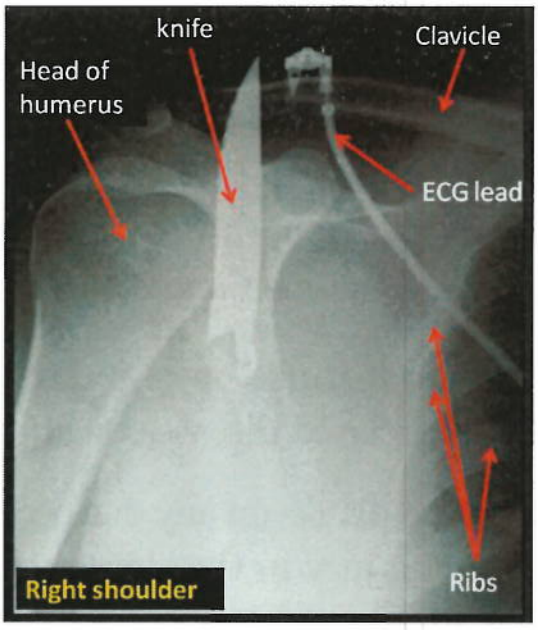


Photo 2.5: A sharp, stabbing pain in the right shoulder: excessive alcohol consumption is a major contributor to violence and crime. (Source: Qld Health and DM.)

(continued over)

***Extension* 2.3.6** *(from [51]) (continued)*

"**Long-term effects**: [In Australia,] each year approximately 3000 people die as a result of excessive alcohol consumption and around 101 000 people are hospitalised. Long-term excessive alcohol consumption is associated with:

- heart damage
- high blood pressure and stroke
- liver damage
- cancers of the digestive system
- other digestive system disorders (e.g. stomach ulcers)
- sexual impotence and reduced fertility
- increasing risk of breast cancer
- sleeping difficulties
- brain damage with mood and personality changes
- concentration and memory problems

Drinking guidelines: National guidelines for alcohol consumption have been developed by the National Health and Medical Research Council to help reduce the risk of harm from alcohol consumption. The guidelines can help in making informed choices and help keep the risk of alcohol-related accidents, injuries, diseases and death, low – both in the short and long term. The guidelines recommend that healthy men and women should drink:

- no more than two standard drinks on any day in order to reduce the risk of harm from alcohol-related diseases or injury in the long term;
- no more than four standard drinks on any one occasion in order to reduce the risk of alcohol-related injury in the short term."

- It is often useful to be able to estimate the time for BAC to return to 0 after

consuming alcohol. The time will vary somewhat between individuals, but governments and health bodies publish general guidelines.

Question 2.3.7

Figure 2.6 shows approximate times required for the BACs of males of different masses to return to 0. (The term *weight* is more common, but *mass* is technically more correct.) Derive a mathematical model of these data.

num. drinks	Mass (pounds)							
	120	140	160	180	200	220	240	260
1	2	2	2	1.5	1	1	1	1
2	4	3.5	3	3	2.5	2	2	2
3	6	5	4.5	4	3.5	3.5	3	3
4	8	7	6	5.5	5	4.5	4	3.5
5	10	8.5	7.5	6.5	6	5.5	5	4.5

Figure 2.6: Approximate times (in hours) for BACs to return to 0 for males.

n, M, t

$n \uparrow$, $t \uparrow$

$M \uparrow$, $t \downarrow$

so $t = n - M$ OR $t = \frac{n}{M} * const$

Looking @ the table, which is better?

What is the value of the constant?

$t = \frac{n}{M} * const \rightarrow const = \frac{Mt}{n}$

substitute in any point or see column for $M = 240$.

$const \approx 240$

Note the constant isn't always 240? Rounding for people!

(To convert approximately from pounds to kg, divide by 2.2.)

Units: $const = \frac{16 \times hr}{\#} = 16.hr$

Question 2.3.8

Figure 2.7 shows actual and modelled times for BACs to return to zero; modelled times were calculated using the above mathematical model.

Actual times:

num. drinks	Mass (pounds)							
	120	140	160	180	200	220	240	260
1	2	2	2	1.5	1	1	1	1
2	4	3.5	3	3	2.5	2	2	2
3	6	5	4.5	4	3.5	3.5	3	3
4	8	7	6	5.5	5	4.5	4	3.5
5	10	8.5	7.5	6.5	6	5.5	5	4.5

Modelled times:

num. drinks	Mass (pounds)							
	120	140	160	180	200	220	240	260
1	2	1.7	1.5	1.3	1.2	1.1	1.0	0.9
2	4	3.4	3.0	2.7	2.4	2.2	2.0	1.8
3	6	5.1	4.5	4.0	3.6	3.3	3.0	2.8
4	8	6.9	6.0	5.3	4.8	4.4	4.0	3.7
5	10	8.6	7.5	6.7	6.0	5.5	5.0	4.6

Figure 2.7: Actual and modelled times (in hours) for BACs to return to 0 for males.

(a) Briefly discuss the effectiveness and accuracy of the mathematical model.

Actual times are more accurate
Modelled times might be more effective
because 1.3 hrs might confuse people?

(continued over)

***Question 2.3.8** (continued)*

(b) Figure 2.8 shows the approximate times for the BACs of males and females of different masses to return to 0.

num. drinks	**Males**: Mass (pounds) 120	140	160	180	200	220	240	260
1	2	2	2	1.5	1	1	1	1
2	4	3.5	3	3	2.5	2	2	2
3	6	5	4.5	4	3.5	3.5	3	3
4	8	7	6	5.5	5	4.5	4	3.5
5	10	8.5	7.5	6.5	6	5.5	5	4.5

num. drinks	**Females**: Mass (pounds) 120	140	160	180	200	220	240	260
1	3	2.5	2	2	2	1.5	1.5	1
2	6	5	4	4	3.5	3	3	2.5
3	9	7.5	6.5	5.5	5	4.5	4	4
4	12	9.5	8.5	7.5	6.5	6	5.5	5
5	15	12	10.5	9.5	8	7.5	7	6

Figure 2.8: Times (in hours) for BACs to return to 0 for males and females.

Comment briefly on the differences in times for males and females. Why do these differences occur? What do these data mean for a "typical male" and "typical female" who go drinking together?

t longer for ♀ of same M
Why? Weight already considered?
Alcohol dissolves in H_2O in body tissue
♀ less H_2O than ♂ (muscle has more H_2O than fat)

End of Case Study 3: Get your BAC up.

Part 2: Getting Hot

Hurld headlong flaming from th' Ethereal Skie
With hideous ruine and combustion down
To bottomless perdition, there to dwell
In Adamantine Chains and penal Fire,
Who durst defie th' Omnipotent to Arms.
Nine times the Space that measures Day and Night
To mortal men, he with his horrid crew
Lay vanquisht, rowling in the fiery Gulfe
Confounded though immortal: But his doom
Reserv'd him to more wrath; for now the thought
Both of lost happiness and lasting pain
Torments him; round he throws his baleful eyes
That witness'd huge affliction and dismay
Mixt with obdurate pride and stedfast hate:
At once as far as Angels kenn he views
The dismal Situation waste and wilde,
A Dungeon horrible, on all sides round
As one great Furnace flam'd, yet from those flames
No light, but rather darkness visible
Serv'd only to discover sights of woe,
Regions of sorrow, doleful shades, where peace
And rest can never dwell, hope never comes
That comes to all; but torture without end
Still urges, and a fiery Deluge, fed
With ever-burning Sulphur unconsum'd:
Such place Eternal Justice had prepar'd
For those rebellious, here their Prison ordain'd
In utter darkness, and their portion set
As far remov'd from God and light of Heav'n
As from the Center thrice to th' utmost Pole.

Paradise Lost (c. 1677), John Milton, (1608 – 1674).

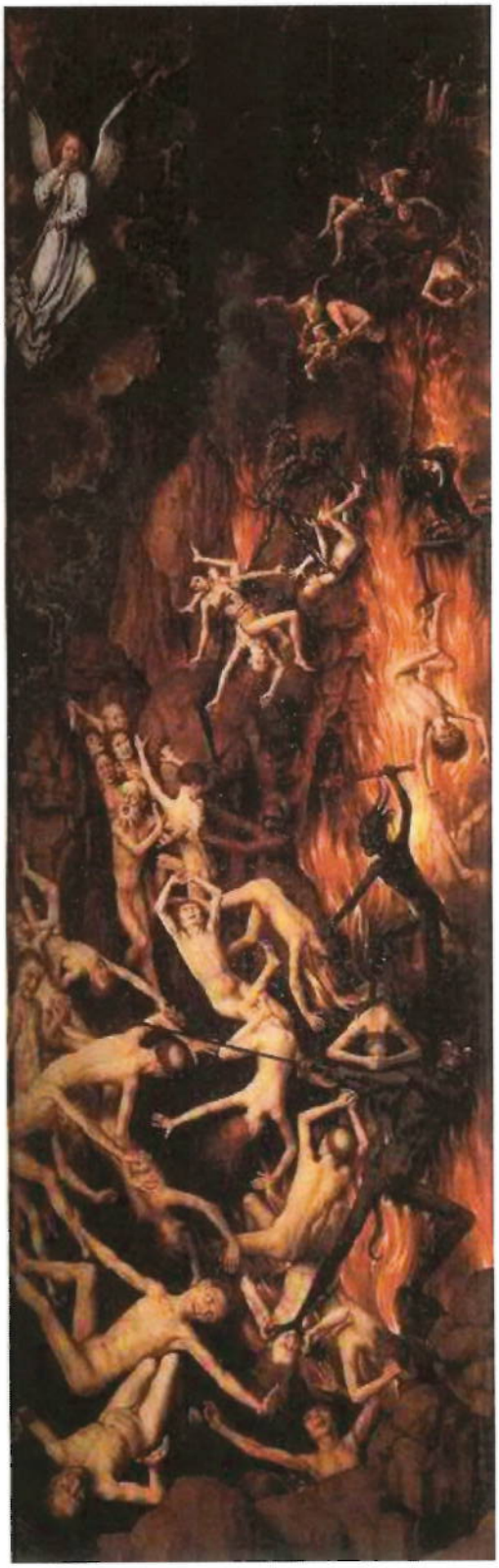


Image 2.2: *Last Judgment* (1467 – 1477), Hans Memling (c. 1430 – 1494), National Museum, Gdansk, Poland. (Source: commons.wikimedia.org).

This section of the notes (Part 2) draws examples from the broad issues associated with climate and climate change, including atmospheric conditions, temperature, species diversity, seasons and environmental degradation. Currently, the general scientific consensus is that Earth is undergoing a period of rapid *climate change*, including *global warming*, which is mostly due to human activities.

However, despite the strong scientific consensus about the existence of climate change, there is still popular and scientific debate about the nature, causes and consequences of climate change. Because there is uncertainty in all (or most) scientific "knowledge", from a scientific perspective such debate is perfectly reasonable, even essential, provided it is informed, logical and based on the best available data. Unfortunately, discussions about complex issues such as climate change are typically emotive, misinformed, parochial, adversarial, alarmist or populist. (It is also interesting to trace some concepts that are accepted now but were controversial initially: that Earth revolves around the sun; that Earth is round; that smoking is a risk factor for cancer; continental drift; that humans could deplete fish stocks.)

Over time, the values of many natural and scientific phenomena, including those associated with climate, approximately follow particular patterns. For example, a value may increase or decrease and this change may occur at a constant, increasing, or decreasing rate. Other phenomena have values that oscillate. When the values of a phenomenon follow a pattern, we can typically model them mathematically.

The next three chapters introduce some tools for mathematical modelling, specifically mathematical functions. These include linear, quadratic, power, periodic, logarithmic and exponential functions. You should have encountered all of this material in previous study.

Remember that SCIE1000 is not a course on climate or climate change, so do not attempt to memorise any climate-related details. Instead, focus on the logical process of model development, and how this relates to other models in different contexts and in different areas of science.

Chapter 3: **A place with atmosphere**

We are a rock revolving around a golden sun
We are a billion children rolled into one
So when I hear about the hole in the sky
Saltwater wells in my eyes.
We climb the highest mountain, we'll make the desert bloom
We're so ingenious we can walk on the moon
But when I hear of how the forests have died
Saltwater wells in my eyes.

Artist: *Julian Lennon.* Song: *Saltwater.*

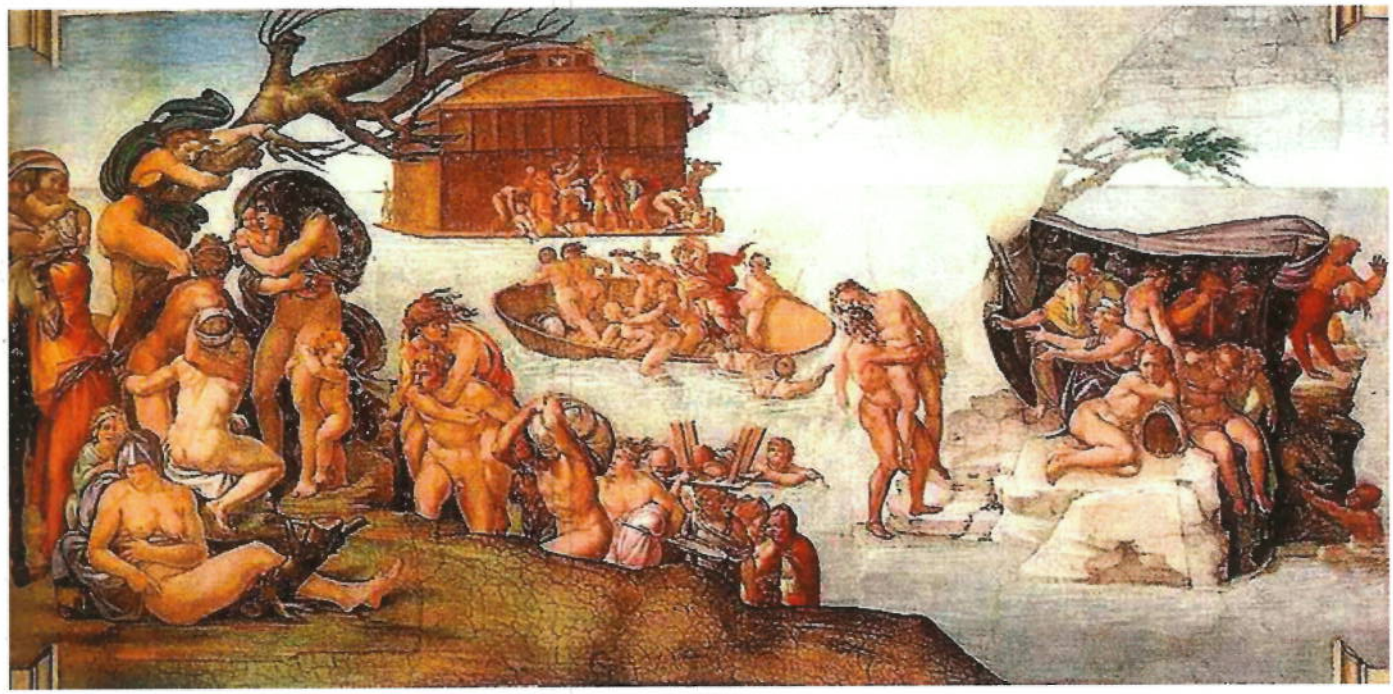


Image 3.1: *The Deluge* (1508 – 1512), Michelangelo (1475 – 1564), Sistine Chapel ceiling, Apostolic Palace, Vatican. (Source: commons.wikimedia.org)

3.1 Fully functional

Case Study 4: **Atmospheric CO_2**

- The broad scientific consensus is that:
  - Earth is undergoing a period of rapid climate change;
  - global temperatures are likely to rise rapidly over coming years;
  - the warming is related to increasing concentrations of carbon dioxide (CO_2) in the atmosphere; and
  - the increase in atmospheric CO_2 concentration is a result of human activity.
- A famous, long-running study has monitored atmospheric CO_2 concentrations at the Mauna Loa observatory in Hawaii since 1958.
- When these data are plotted, the graph is called the *Keeling curve*, named after the initiator of the study, David Keeling.
- The Scripps Institution of Oceanography (which runs the study) describes the Keeling curve as "... almost certainly the best-known icon illustrating the impact of humanity on the planet as a whole..."
- Gases in the lower atmosphere mix fairly well, so scientists consider the Keeling curve data to be representative of the atmospheric CO_2 concentration world-wide.
- By July 2016, the level of atmospheric CO_2 was consistently above 400 parts per million by volume (ppm or ppmv). (When SCIE1000 was first offered in 2008, the figure was about 380 ppm.)
- Other data from ice-core samples show that CO_2 levels remained relatively constant at 280 ppm for thousands of years, but the level started increasing in the 19th century.
- Figure 3.1 is a plot of the Keeling curve data collected between 1958 and July 2016, taken from [29, 48].

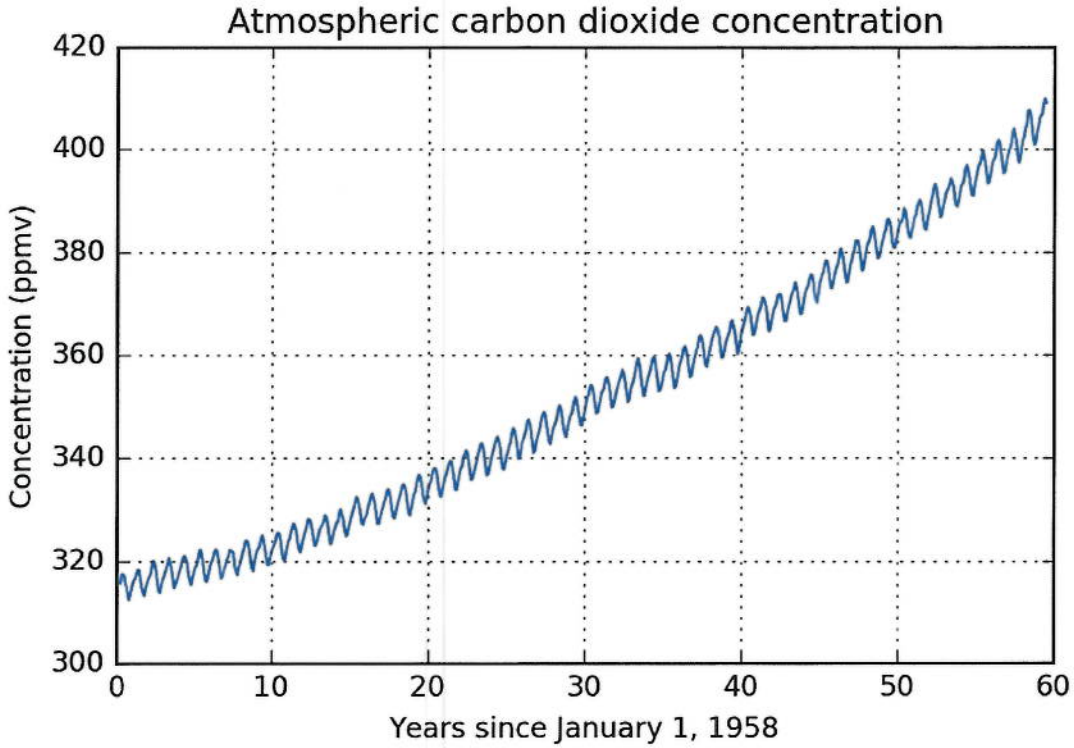


Figure 3.1: The Keeling curve.

NOAA video of $[CO_2]$ over 100s of thousands of years.

(*) Play pumphandle video

Question 3.1.1

(a) Describe the main features of the Keeling curve graph.

(1) Curves upwards

(2) Oscillations (wiggles)

(b) What physical factor(s) could cause these features?

(1) CO_2 produced by humans

(2) Annual cycle – seasons
- Why?
- deciduous trees

(c) How could you mathematically model the Keeling curve?

- linear + cycle (sine?)
- quadratic + cycle
- etc.

End of Case Study 4: Atmospheric CO_2.

- Earlier we saw the importance of modelling, why most models are quantitative, and the five common ways of presenting conceptual models (using words, values, pictures, mathematics and computer programs).
- The basic mathematical tool used to describe quantitative relationships and patterns in models is the mathematical *function*.

Mathematical functions

A *function* is a rule that converts input value(s) to output value(s).

Often, the input values are called x or t (for time). On a graph, they are represented on the *horizontal* axis. If f is the name of a function, then $f(t)$ denotes the output that arises from applying f to the input value t. On a graph, these values are represented on the *vertical* axis.

- People study a range of functions, including: linear, quadratic, power, periodic, exponential, logarithmic, and combinations of these. These functions are interesting **precisely** because they model natural phenomena.
- A key skill when modelling is to recognise which type of function will most usefully represent the observed data, balancing accuracy with complexity.
- In the next few sections, we will study some phenomena and see how mathematical functions allow us to represent and study these phenomena.
- Do not memorise details of the case studies (such as scientific names or equations for phenomena) or specific mathematical information.
- Instead, understand the concepts **behind** the examples, including which functions are most suitable for modelling certain types of phenomena, and how to interpret mathematics and science in an integrated manner.
- One point we will continually stress is the diversity of phenomena that can be modelled by the same families of functions.
- Briefly consider the following graphs of various phenomena, paying attention to the general shapes of the graphs. (Of course, the relationships will not be perfect, due to factors including natural variation and data errors.)
- The first functions we will study in detail are the *power functions*, which include linear, quadratic and other functions.

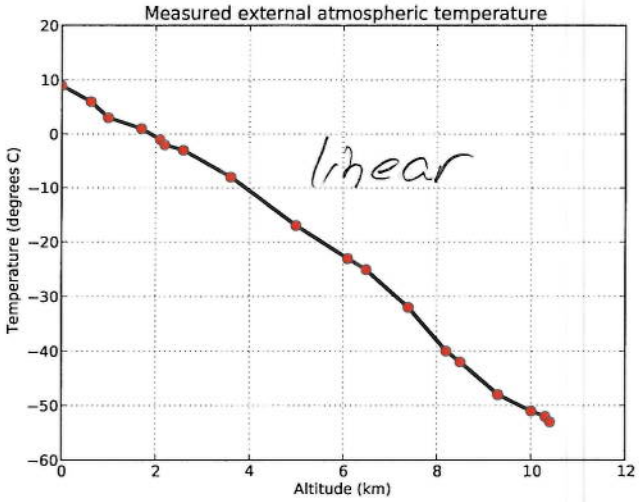


Figure 3.2: Atmospheric temperatures recorded on an international flight, at different altitudes.

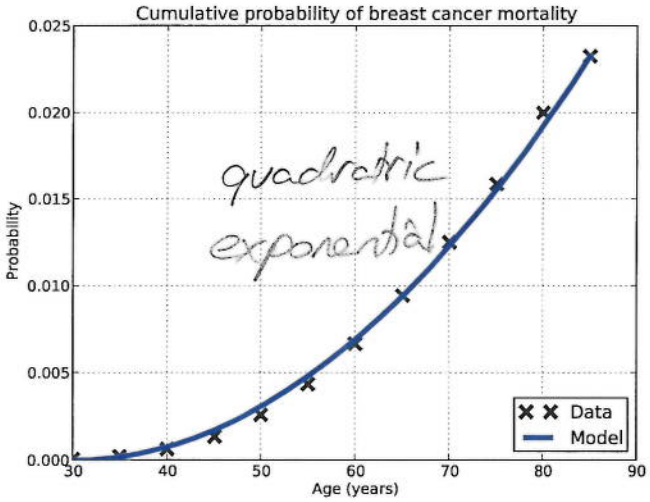


Figure 3.3: The cumulative probability of female breast cancer mortality by age.

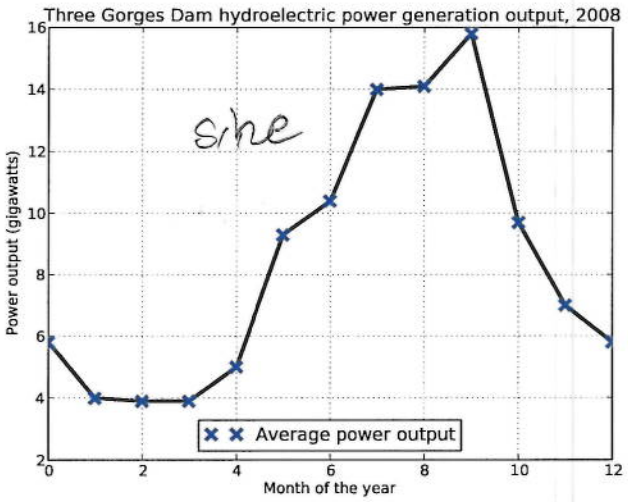


Figure 3.4: Electricity generation at the Three Gorges Dam, China, in 2008.

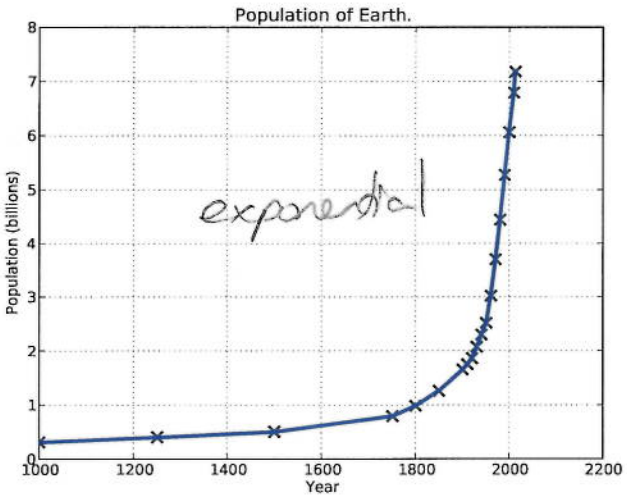


Figure 3.5: Population of Earth for the last 1000 years.

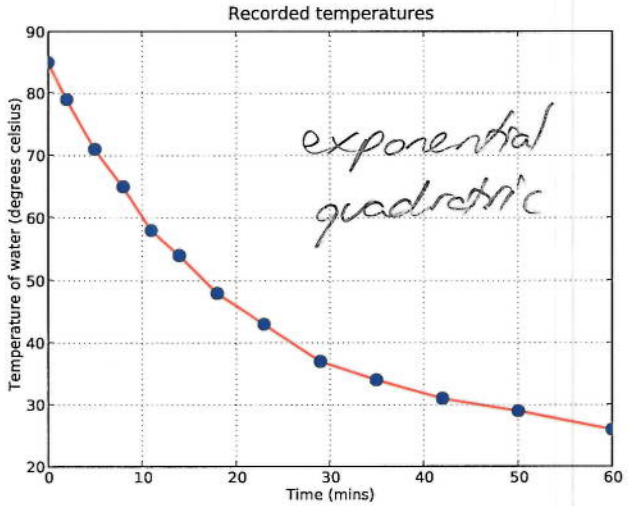


Figure 3.6: Measured water temperatures in a simple experiment.

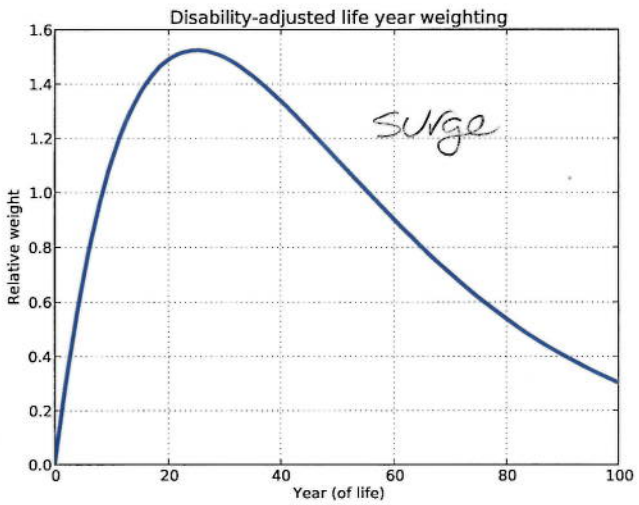


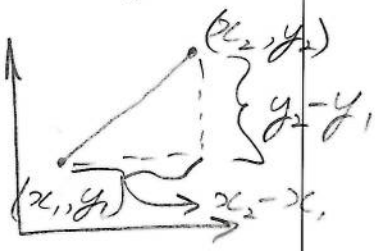
Figure 3.7: "Social weight" of the value of each year of life, used in *Global burden of disease* studies, [57].

3.2 Going straight

- Many natural phenomena are related (approximately) linearly; that is, in a *straight line*. Straight lines form the basis of many techniques we will study, including linear models, calculus, Newton's method for solving equations and Euler's method for solving differential equations.

Linear functions

Linear functions have equations $y(x) = mx + c$, where m is the *gradient* and c is the *y-intercept* of the line.



If (x_1, y_1) and (x_2, y_2) are two points on the line then

$$m = \frac{\text{rise}}{\text{run}} = \frac{\text{change in } y}{\text{change in } x} = \frac{y_2 - y_1}{x_2 - x_1}.$$

Case Study 5: **Temperature**

Photo 3.1: Hot spring, Yellowstone Park, USA. (Source: PA.)

Question 3.2.1

A temperature of c degrees Celsius can be converted to equivalent temperature F in Fahrenheit by the linear function:

$$F(c) = \frac{9c}{5} + 32.$$

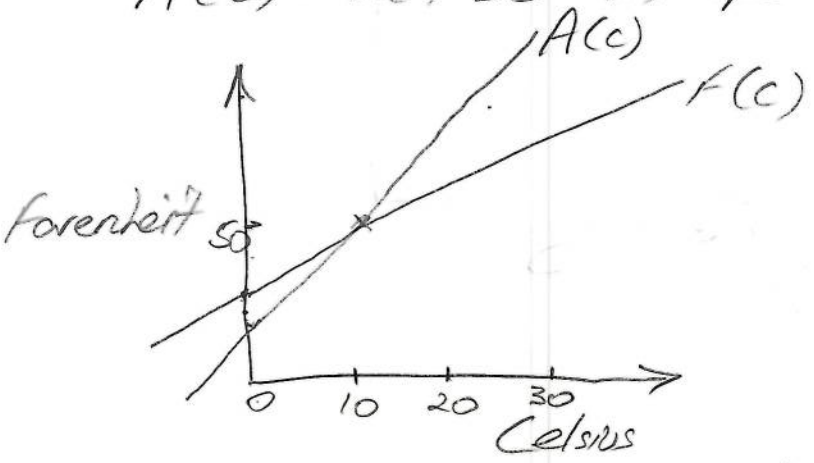
The following "rule of thumb" converts a temperature in celsius to an approximate iequivalent temperature A in Fahrenheit:

$$A(c) = 2c + 30.$$

(a) Sketch rough graphs of F and A on a set of axes.

$F(c) = \frac{9c}{5} + 32 \Rightarrow$ slope $= \frac{9}{5}$ & intercept $= 32$

$A(c) = 2c + 30 \Rightarrow$ slope $= 2$ & intercept $= 30$



Intersection point

$F(c) = A(c)$

$\Rightarrow \frac{9c}{5} + 32 = 2c + 30$

$\frac{10c}{5} - \frac{9c}{5} = 32 - 30$

$\Rightarrow c = 10$

(b) Find an expression for the error that arises when using the "rule of thumb", and comment on the suitability of the approximate value A.

Error = Estimate − Actual

$= 2c + 30 - \left(\frac{9c}{5} + 32\right)$

$= 2c - \frac{9c}{5} - 2$

$= \frac{c}{5} - 2$

So $A(c) > F(c), T > 10°C$

$A(c) < F(c), T < 10°C$

$A(c) = F(c)\ T = 10°C$

$\therefore$ $A(c)$ is a good approximation around $10°C$

End of Case Study 5: Temperature.

Case Study 6: **Higher than a kite**

Photo 3.2: Jetliner cruising at an altitude of about 10000 m. (Source: PA.)

- Scientists divide Earth's atmosphere into five primary regions: *troposphere*, *stratosphere*, *mesosphere*, *thermosphere* and *exosphere*.
- The *International Standard Atmosphere* (ISA) [28] is a model which further divides the atmosphere from the surface of Earth to the base of the thermosphere into eight layers. (Layer 0 is closest to the surface.)
- The ISA models various properties of each layer, including temperature, pressure and density.
- Layers in the ISA are defined as atmospheric regions in which *temperature is a linear function of altitude.*
- Figures 3.8 and 3.9 show various properties of the ISA temperature at different altitudes. (The *lapse rate* is the rate at which temperature changes as altitude increases.)

Layer	Name	Height at base (km)	Lapse rate (°C/km)	Temperature at base (°C)
0	Troposphere	0.0	−6.5	+15.0
1	Tropopause	11.0	+0	−56.5
2	Stratosphere	20.0	+1.0	−56.5
3	Stratosphere	32.0	+2.8	−44.5
4	Stratopause	47.0	+0	−2.5
5	Mesosphere	51.0	−2.8	−2.5
6	Mesosphere	71.0	−2.0	−58.5
7	Mesopause	84.852	NA	−86.2

Figure 3.8: Some properties of the layers within the International Standard Atmosphere.

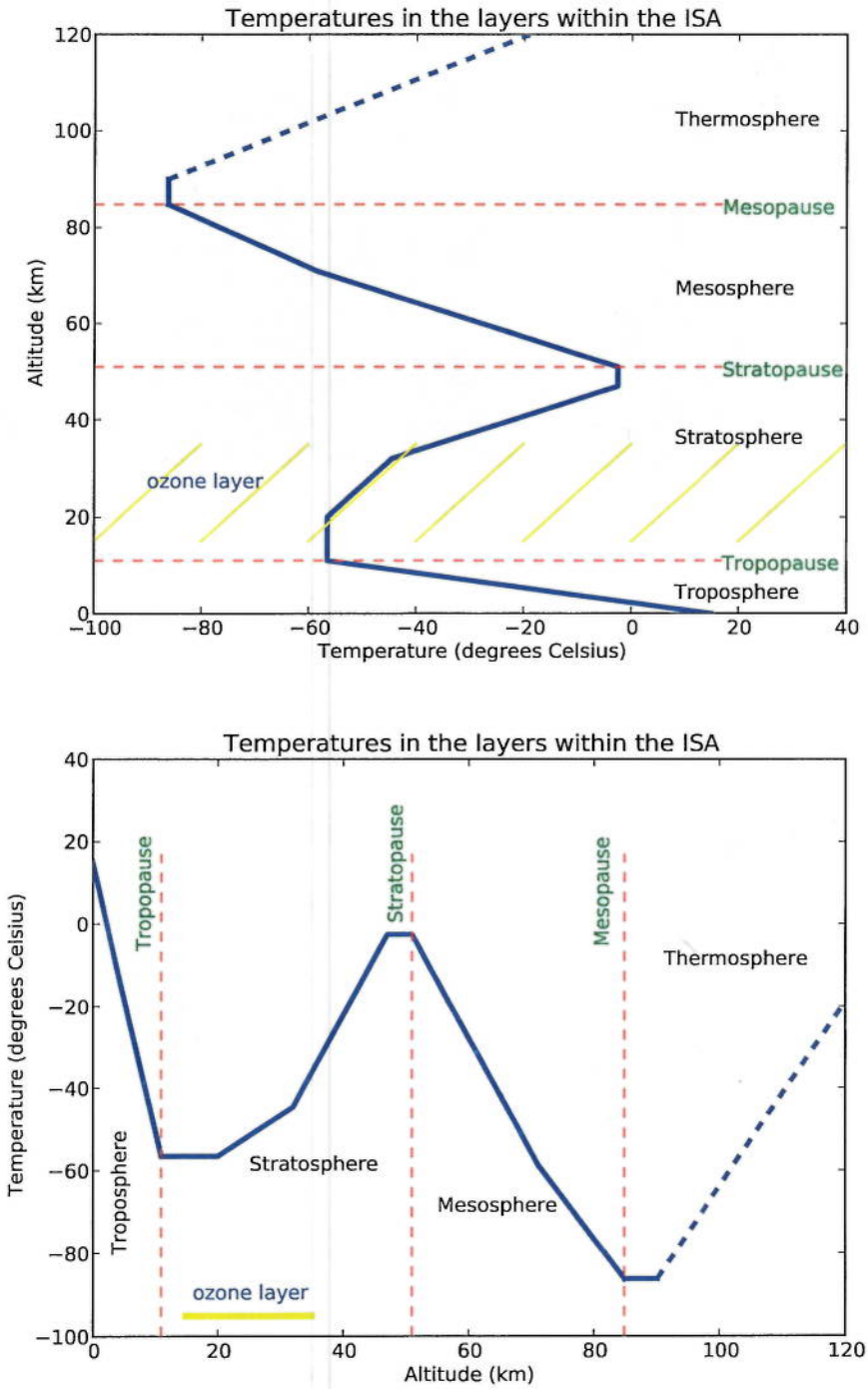


Figure 3.9: The relationships between temperature and altitude modelled by the ISA. The top graph shows altitude versus temperature, and the bottom graph shows temperature versus altitude. (The ISA does not model the thermosphere; temperature data displayed for the thermosphere are from other measurements.)

Question 3.2.2

Which of the graphs in Figure 3.9 is more correct, and which is more useful? Why?

Upper: Good visualisation. Poor mathematically, not a function

Lower: Good mathematically, piecewise equation

Poor visualisation

Question 3.2.3

Use the information about the ISA to answer the following.

(a) Write the troposphere temperature as a function of altitude. $T = mA + c$

$m = \frac{\text{rise}}{\text{run}} = \frac{-56.5 - 15}{11 - 0} = \frac{-71.5}{11} = -6.5$ (lapse rate)

& substitute in $\Rightarrow$ $15 = -6.5 \times 0 + c \Rightarrow c = 15$ OR From Table.

So $T = -6.5A + 15$

(b) The Matterhorn is a mountain in the Swiss Alps, with a height of 4478 m above sea level. The summit air temperature can range from around 0 °C to −40 °C at different times of the year. Reconcile this with the temperature predicted by the ISA.

$T = -6.5 \times 4.478 + 15 = -29.107 + 15 = -14.107°C$

This is the mean temperature, & does not consider seasonality, wind chill, cloud, etc.

- Model also assumes a constant temperature @ sea level of 15°C.
- Model must be good enough for what its used for

Photo 3.3: The Matterhorn – Italian side (Source: PA.)

(continued over)

Question 3.2.3 *(continued)*

(c) On a recent international flight, Peter recorded altitudes and external temperatures reported on the in-flight information screen. The data are graphed in Figure 3.10.

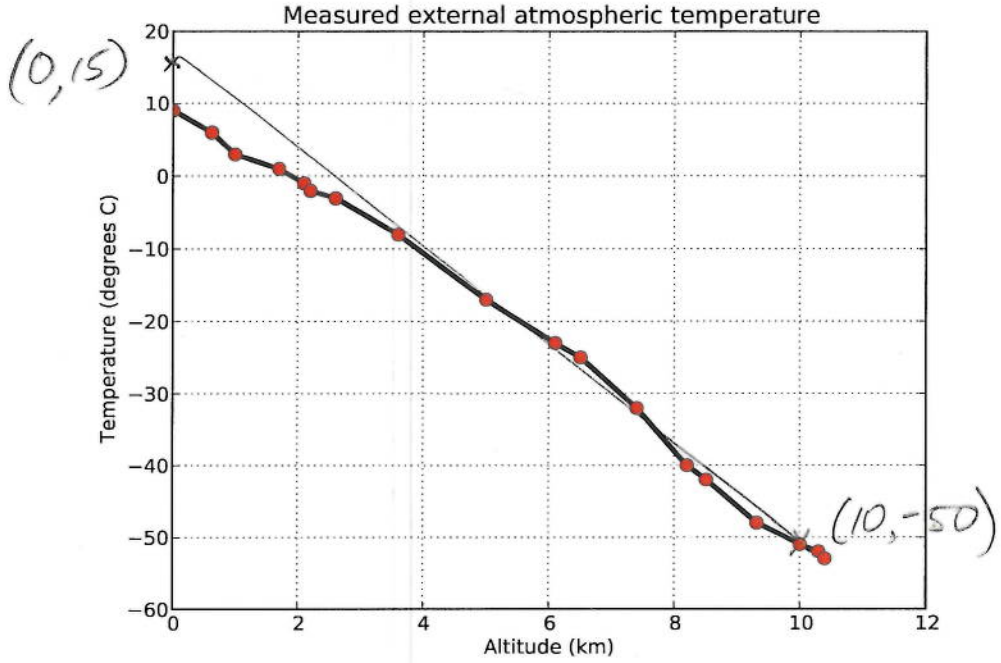


Figure 3.10: Altitudes and measured external temperatures.

Plot the function from Part (a) on the above graph and comment on the results.

Excellent!

(d) Write the temperature in ISA Layer 3 as a function of altitude.

Slope $= \frac{\text{rise}}{\text{run}} = 2.8$ (lapse rate). $T = 2.8A + C$

When $A = 32$, $T = -44.5$ $\Rightarrow$ $C = -44.5 - 2.8 \times 32 = -134.1$

$\Rightarrow$ $\boxed{T = 2.8A - 134.1}$ Why isn't $C = -44.5$?

OR $T = 2.8(A - 32) - 44.5$ $\Rightarrow$ $T = 2.8A - 2.8 \times 32 - 44.5$

$\Rightarrow$ $\boxed{T = 2.8A - 134.1}$

End of Case Study 6: Higher than a kite.

Check: when $A = 47$, $T = 2.8 \times 47 - 134.1 = -2.5$ ✓

Now we develop a rough mathematical model of the Keeling curve; as we cover more functions, we will gradually develop more accurate models.

Question 3.2.4

Keeling Model 1: Figure 3.11 shows a graph of the Keeling curve.

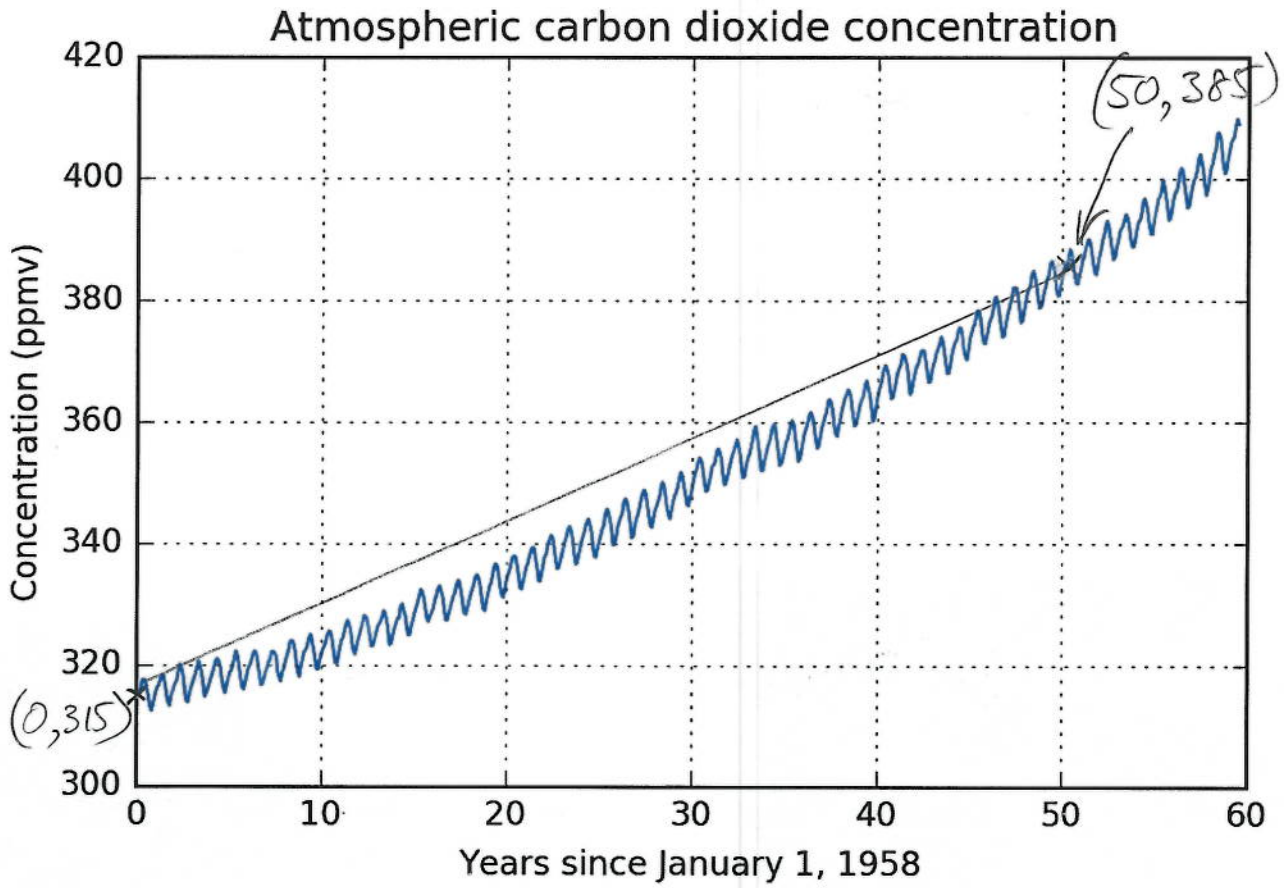


Figure 3.11: The Keeling curve.

(a) Find a rough linear model of the Keeling curve, and plot it on the graph.

$m = \text{slope} = \frac{\text{rise}}{\text{run}} = \frac{385-315}{50-0} = \frac{70}{50} = 1.4 \frac{\text{ppmv}}{\text{year}}$

So $\boxed{[CO_2] = 1.4t + 315}$

(b) Discuss the effectiveness of your model.

- Over-estimates intermediate years
- Poor for extrapolation
- Does not include cyclical variation
- Only includes trend

3.3 Bend it!

- Many scientific phenomena relate in ways that are not straight lines.

Quadratics and modelling

Quadratic functions have a power of x (or t, or ...) equal to 2, with equations of the form $y(x) = ax^2 + bx + c$, where a, b and c are constants and $a \neq 0$. The graphs of quadratics are parabolas.

Quadratics are important in practical modelling, particularly when modelling over short time periods. They are the simplest functions with *local optimal* values, that is, local *maximum* or *minimum* values.

Case Study 7: Climate change and Bicknell's thrush

Image 3.2: Bicknell's thrush, *Catharus bicknelli.* (Source: en.wikipedia.org.)

Photo 3.4: Adirondack mountains, USA. (Source: PA.)

Example 3.3.1

A paper [44] developed models for bird distributions using data from various altitudes, temperatures and locations in the north-eastern USA. The authors then used their models to predict the likely impact of rising temperatures on these distributions. Part of their study focused on Bicknell's thrush.

(continued over)

***Example 3.3.1** (continued)*

- Collecting data for the study involved: subdividing the study region into cells, each 30 m square; measuring the mean daily maximum temperature in July (summer) in each cell; and conducting fieldwork on a representative sample of these cells to identify which contained at least one resident thrush.
- Using the data, the authors created a model for thrush distribution with respect to mean July temperatures across the breadth of their habitat.
- The study found that thrush habitats with July temperatures outside the range of 9.3 °C to 15.6 °C contained insignificant numbers of thrush.

Let t be a temperature within the range 9.3 °C – 15.6 °C. The proportion $p(t)$ of cells containing thrush is closely modelled by the quadratic function:

$$p(t) = -0.0747t^2 + 1.8693t - 10.918.$$

Question 3.3.2

The graph of $p(t) = -0.0747t^2 + 1.8693t - 10.918$ is shown in Figure 3.12.

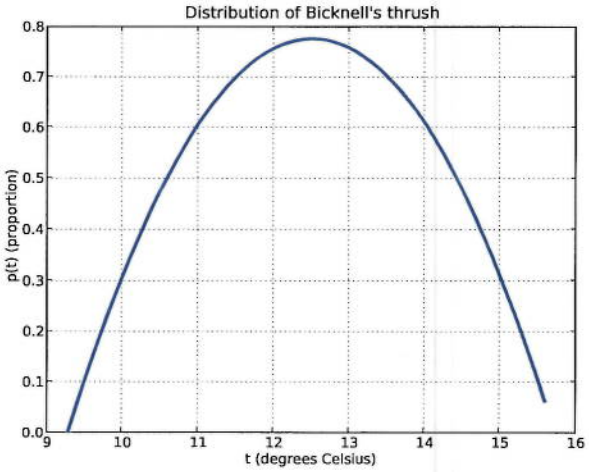


Figure 3.12: Distribution of Bicknell's thrush according to temperature.

(continued over)

***Question 3.3.2** (continued)*

(a) What is the probability that a thrush will be found in a sample area in which $t = 11$ °C?

$p(11) \approx 0.6$

(b) From the graph, at what (approximate) value of t is the thrush distribution most dense, and what is the (approximate) value of $p(t)$?

at $t \approx 12.5$°C, $p(12.5) \approx 0.78$

(c) There is no value of t for which $p(t) = 1$. Explain what this means in terms of the thrush distribution, and give reasons why it would happen.

Thrush not found everywhere it can be. Other considerations important other than temperature eg. nests?, food?, trees?, predators?

(d) Average temperature rises in the region over the next century are predicted to range from 2.8 °C under a low greenhouse gas emission scenario, to 5.9 °C under a high emission scenario.

(i) How would the graph in Figure 3.12 change if the average temperature rose by 2.8 °C? What if it rose by 5.9 °C? Explain your answers.

No change. The underlying temperature preferance of a species (its niche) doesn't change in a warmer world (assuming no adaptation/evolution)

(ii) Assuming a substantial rise in average July temperatures, which key factor of concern to resident thrush would change?

Won't it be OK? As it warms, the thrush moves out of areas that are now too warm & into areas which were too cold. A zero sum game?

No. Think about its habitat
= => less habitat => lower population

(continued over)

Question 3.3.2 (continued)

Figure 3.13 shows the total area of existing thrush habitat, and the estimated amount of viable habitat available after predicted temperature increases under the low emission scenario and the high emission scenario.

Scenario (°C)	Habitat (hectares)
(current) +0°C	140000
+1°C	32000
+2°C	10000
+3°C	1000
+4°C	200
+5°C	75
+6°C	0

Figure 3.13: Total areas of viable habitat available to Bicknell's thrush under various climate change scenarios.

(e) What is the likely impact on the thrush population if temperatures rise by 2.8 °C or 5.9 °C?

T↑ by 2.8°C ≃ 3°C => 1000 ha => $\frac{1000}{140000}$ = 0.7% left

T↑ by 5.9°C ≃ 6°C => 0 => 0% left

(f) If temperature increases occur at the higher end of predictions, what kind of survival strategies might the thrush utilise?

- Migrate to higher mountains
- Adapt (active during night/dawn/dusk when cooler)
- Evolve new temperature preferance
- Extinction?

End of Case Study 7: Climate change and Bicknell's thrush.

Now we develop a more accurate mathematical model of the Keeling curve.

Question 3.3.3

Keeling Model 2: Figure 3.14 shows two plots: a graph of the function $y(t) = 0.014t^2 + 0.7t + 315$, and the Keeling curve.

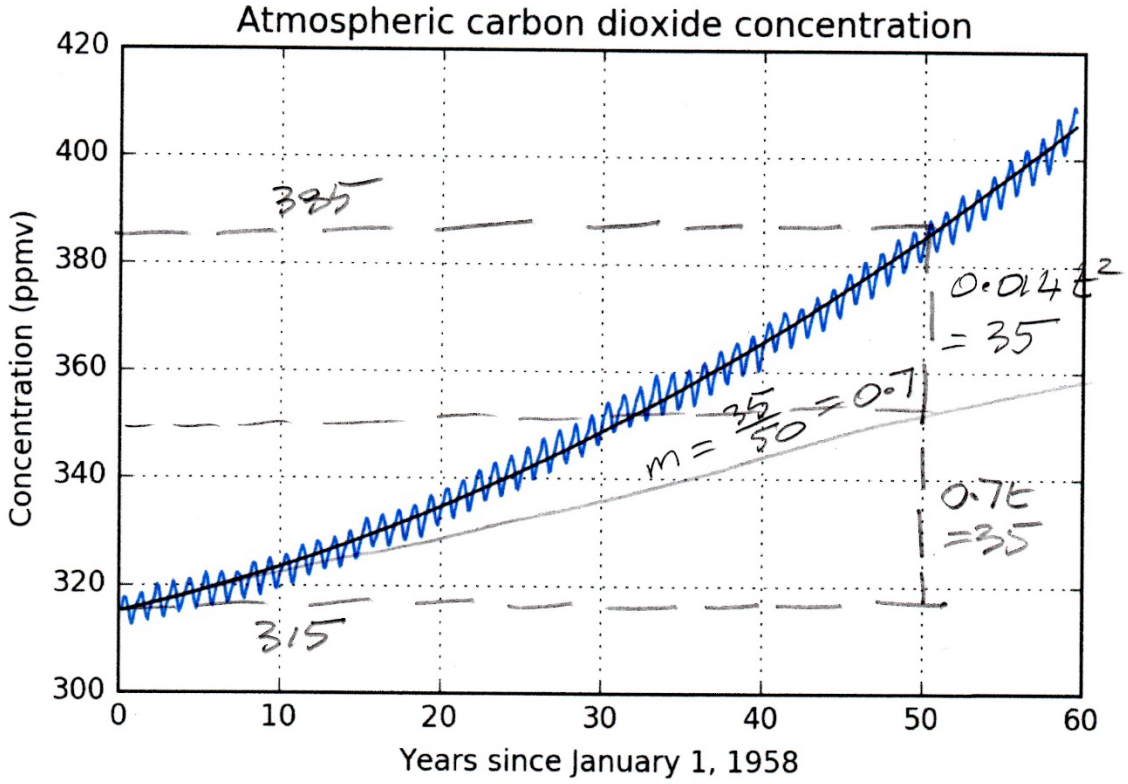


Figure 3.14: The Keeling curve and a quadratic model.

(a) Explain how each term in $y(t)$ impacts on the graph.

$Y = \underbrace{0.014t^2}_{\text{curve}} + \underbrace{0.7t}_{\text{linear}} + \underbrace{315}_{\text{y-intercept}}$

½ increase ½ increase

(b) How effectively does $y(t)$ model the Keeling curve?

good general fit

does not account for annual cycles

3.4 (Super) powers

- Recall that linear and quadratic functions are examples of the more general group of *power functions*. Functions with different powers have graphs with different shapes, and hence can model different phenomena.

Case Study 8: Species-area curves and biodiversity

Photo 3.5: Counting species in the field. (Source: DM.)

- Previously we discussed the abundance and distribution of a *single* species, *Bicknell's thrush*. Ecologists often study the *overall number of species* found in a region (sometimes called the *biodiversity* or *species richness*).

Species-area curves

In ecology, a *species-area curve* is a graph showing the number of distinct species observed, as a function of the size of the area surveyed.

Photo 3.6: Scribbly gum (*Eucalyptus racemosa*). Right: Scaly-breasted Lorikeet (*Trichoglossus chlorolepidotus*). (Source: PA.)

- Rather than performing a full count for an entire region, data from a smaller area can be extrapolated to estimate the regional species richness.

Example 3.4.1

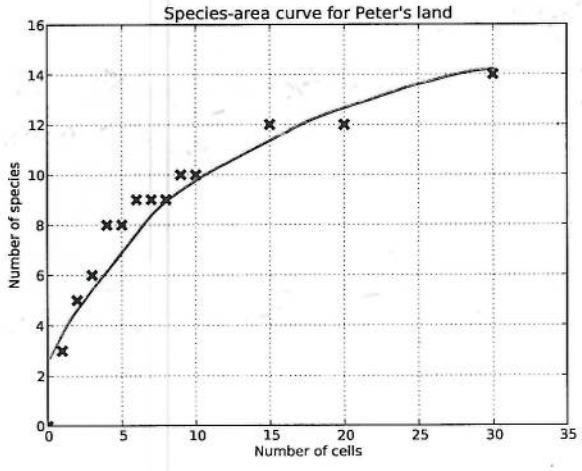
Peter lives on 4 hectares in eastern Brisbane. He wishes to estimate the number of distinct, naturally occurring, native plant species (individuals greater than 2 m in height), that occur on his land. He randomly selects 30 cells (or *quadrats*), each 10 m square, and records the occurrence of new, relevant species in each cell. Figure 3.15 shows information on the previously unseen species, including the cumulative total C of species observed so far.

Cell(s)	New species observed	C
1	*Eucalyptus racemosa, Acacia fimbriata, Banksia integrifolia*	3
2	*Eucalyptus tereticornis, Alphitonia excelsa*	5
3	*Acacia disparrima*	6
4	*Acacia leiocalyx, Lophostemon suaveolens*	8
5	–	8
6	*Glochidion sumatranum*	9
7	–	9
8	–	9
9	*Eucalyptus crebra*	10
10	–	10
11 – 15	*Banksia robur, Melaleuca quinquinerva*	12
16 – 20	–	12
21 – 30	*Allocasuarina littoralis, Angophora leiocarpa*	14

Figure 3.15: Information on additional observed species.

Example 3.4.2

Figure 3.16 is a species-area curve summarising the data in Figure 3.15:



- Not linear & not quadratic
- # spp. is cumulative & never declines

Figure 3.16: The number of distinct tree species recorded on Peter's land.

- The graph has a shape that is typical of many species-area curves: the number of distinct species initially rises rapidly as the area increases, but then rises less rapidly as the area becomes larger.

Equations for species-area curves

Species-area curves can be mathematically modelled using power functions, with power p between 0 and 1 (typically, p is between 0.2 and 0.5).

Their general form is $S(a) = Ma^p$, where S is the number of species occurring as a function of the area a, and M and p are constants depending on the geographical location, resource availability and similar factors.

Question 3.4.3

With respect to a species-area curve $S = Ma^p$ (with p between 0 and 1):

(a) Give some physical reasons for the general shape of species-area curves.

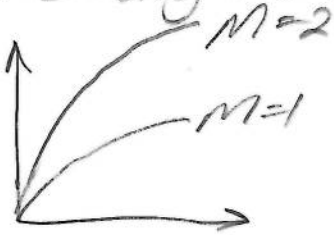
- Rapid ↑ as you initially find new spp.
- Slower ↑ later because most spp. already found

(b) How might this impact on field sampling techniques?

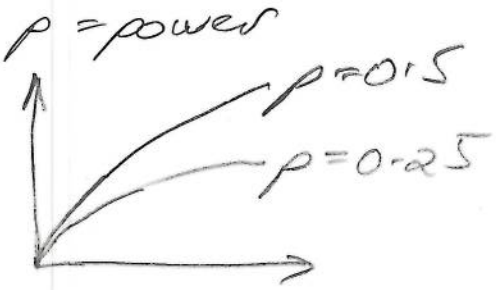
- Helps to reduce sampling effort => don't sample everything (if in similar habitats)

(c) How do the values of M and p impact on the shape of the graph?

M = scaling factor



p = power



(d) What physical factors could make M and p smaller or larger?

Resources: M↑, p↑ Competition: M↓, p↓

Habitat complexity: M↑, p↑ Predation: M↓, p↓

Example 3.4.4

Figure 3.17 shows the graph of $f(a) = 5a^{0.3}$ and the species data from Figure 3.15, where a is the number of 10 m square cells.

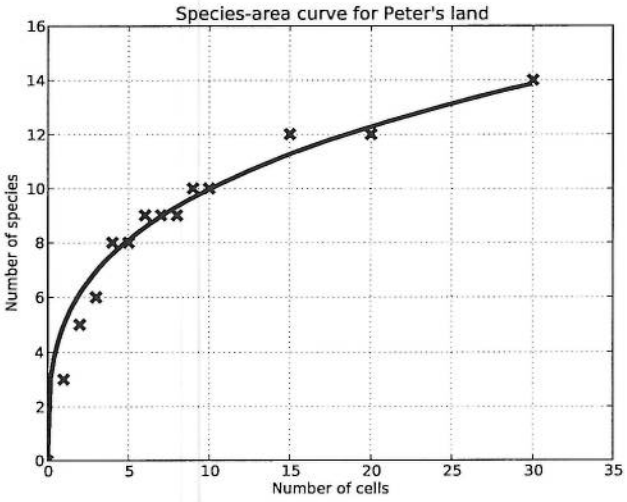


Figure 3.17: Modelling the species data from Peter's land.

Question 3.4.5

Assume that this question refers to native, naturally occurring plants more than 2 m high, growing on land ecologically similar to Peter's.

(a) Estimate the species richness on Peter's 4 hectare (40000 m^2) property.

$\#\text{ cells} = \frac{40000\ m^2}{100\ m^2/\text{cell}} = 400$ cells

$f(a) = 5a^{0.3} = 5 \times (400)^{0.3} \approx 30$ species

(b) A typical conservation goal is to establish parks that preserve 10% of the representative land area. What fraction of species richness would be represented within such a park in the area where Peter lives?

Method (1) 10% of $40000\ m^2 = 4000\ m^2 = 40$ cells

$f(a) = 5 \times (40)^{0.3} \approx 15$ species

$\therefore$ fraction $\approx \frac{15}{30} = 0.5$

Method (2) $\text{Fraction} = \frac{\text{New}}{\text{Old}} = \frac{5(0.1a)^{0.3}}{5a^{0.3}} = 0.1^{0.3}$

≈ 0.5

Is this enough?

(continued over)

Question 3.4.5 (continued)

(c) Many people believe that the figure in Part (b) is too low. If the goal is to retain 75% of species, what proportion of land should be preserved?

Let the proportion of land = k

$\therefore 0.75 = \frac{\text{New}}{\text{Old}} = \frac{S(kq)^{0.3}}{S q^{0.3}} = k^{0.3}$

$\Rightarrow k = 0.75^{\frac{10}{3}} \simeq 0.38$.

Why is it so much higher than 10%?

Question 3.4.6

The paper [41] uses species-area curves to predict the reduction in species richness of vertebrate species in Mexican cloud forests. In the table, A_0 is the current area of cloud forest in two regions, A_1 is the predicted area in 2080 after climate change, and A_2 is the predicted area after climate change and forest clearing. The corresponding numbers of endemic vertebrate species are S_0 (current), S_1 and S_2.

Region	S_0	A_0	S_1	A_1	S_2	A_2
Oaxaca	26	5160	21	2326	9	65
Chiapas	3	6037	2	797	1	45

(a) Show how to calculate the value of S_1 for Oaxaca. (Hint: $p = 0.25$.)

$S_1 = MA_1^{0.25} \Rightarrow M(2326)^{0.25}$ $M = ?$

Estimate M from $S_0 = MA_0^{0.25} \Rightarrow M = \frac{S_0}{A_0^{0.25}} = \frac{26}{(5160)^{0.25}} = 3$

$\Rightarrow S_1 = 3 \times (2326)^{0.25} = 21$ ✓

(b) Which of Oaxaca or Chiapas would you suggest as the location for a new national park? Why? What other factors might influence your advice?

- Oaxaca because it has more species now & in the future. Also proportionally more spp./unit area
- Other factors: COST, access, hunting, how pristine.

End of Case Study 8: Species-area curves and biodiversity.

Case Study 9: **Ban the tan, man**

- Earlier, we saw that temperature in the lowest atmospheric layer (Troposphere) decreases as altitude increases, but temperature in the next layer (Stratosphere) increases from altitude 20 km to 50 km.
- The rise in temperature is due to interactions between the ozone layer and ultraviolet (UV) light.
- UV light is electromagnetic radiation with wavelengths shorter than visible light, and can be divided into: UVA (wavelength 315 – 400 nm); UVB (wavelength 280 – 315 nm); and UVC (wavelength 100 – 280 nm).
- The following sequence of chemical reactions occurs in the ozone layer:

$$\begin{aligned} O_2\,(g) &\xrightarrow{\text{UVC light}} 2O\,(g) && \text{(exothermic: light energy} \longrightarrow \text{heat)} \\ O_2\,(g) + O\,(g) &\longrightarrow O_3\,(g) \\ O_3\,(g) &\xrightarrow{\text{UVB light}} O_2\,(g) + O\,(g) && \text{(exothermic: light energy} \longrightarrow \text{heat)} \end{aligned}$$

- Collectively, these reactions are called the *ozone-oxygen* cycle:
  - The net result of the first two reactions is that UVC light energy is converted to heat, and oxygen is converted into ozone, O_3.
  - In the third reaction, UVB light is absorbed; this reaction is also exothermic, again converting light energy into heat.
- This cycle is extremely important to life on Earth:
  - UVC light is very damaging to life, but is completely absorbed.
  - Most UVB light is also absorbed; only around 1 part in 350 million reaches the surface of Earth. Exposure to this light causes sunburn, eye cataracts, visible ageing, genetic mutations in cells and skin cancer.
  - Almost all UVA light reaches the surface of Earth.
- The effectiveness of sunscreens at preventing UVB light from reaching the skin is measured by their *Sun Protection Factor*, SPF. When a product with SPF n is correctly applied to the skin, it blocks a fraction of $(n-1)/n$ of the usual amount of UVB light.

Question 3.4.7

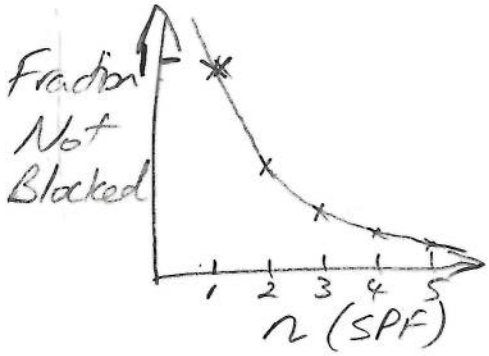
Write a function for the proportion of UVB light that is **not** blocked by sunscreen with SPF n, and draw a rough sketch of the graph of the function.

Fraction blocked $= \frac{n-1}{n}$

$\therefore$ fraction not blocked $= 1-\left(\frac{n-1}{n}\right)$

$= \frac{n-(n-1)}{n}$

$= \frac{1}{n}$



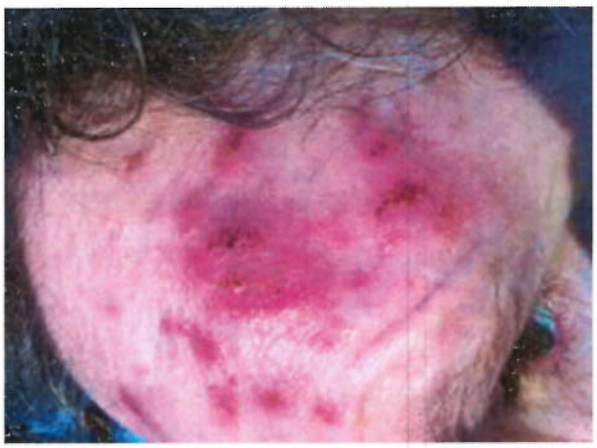


Photo 3.7: Medical treatment of skin damage caused by exposure to UV light. (Source: PA.)

- There has been a substantial depletion of total atmospheric ozone in recent decades, including formation of the *ozone hole* over Antarctica.
- It is well-accepted that this has anthropogenic causes, particularly the release of ozone depleting substances such as *chlorofluorocarbons* into the atmosphere. These had been used as refrigerants and aerosol propellants.
- The *Montreal Protocol*, adopted in 1989 and ratified by around 200 states, is an agreement on phasing out the use of CFCs. It represents one of the most significant international climate agreements ever.
- There is evidence that the ozone layer has started repairing, and scientists believe that it will recover by the year 2050.

End of Case Study 9: Ban the tan, man.

Summary of power functions

Power functions have equations

$$y(x) = Mx^p + c$$

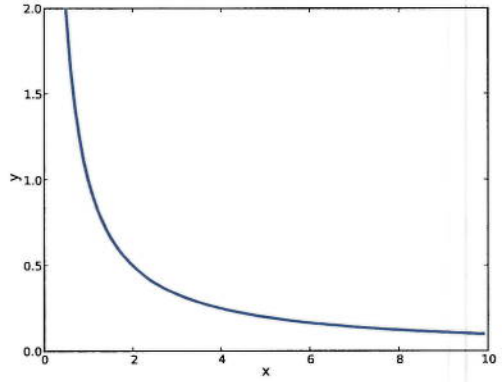
where M, p and c are constants. Changing the value of these constants generates graphs with different shapes, which makes power functions useful for modelling a range of phenomena. For example, changing the value of:

- the power p creates graphs that increase or decrease, at different rates;
- the constant M *scales* the vertical height of the graph at each point; and
- the constant c *shifts* the graph up or down.

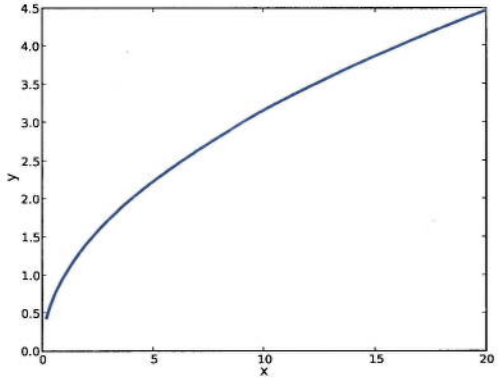
Figure 3.18 illustrates how the value of the power p affects the general shape of the corresponding graph, **for positive values of C and x.** Figures 3.19 and 3.20 illustrate this, showing some equations and their graphs.

Power, p	General shape of the graph
< 0	curve, **decreasing less rapidly** as x increases
0	horizontal line
> 0 and < 1	curve, **increasing less rapidly** as x increases
1	straight line
> 1	curve, **increasing more rapidly** as x increases

Figure 3.18: Different powers and the general shapes of the corresponding graphs.



$y = x^{-1}$

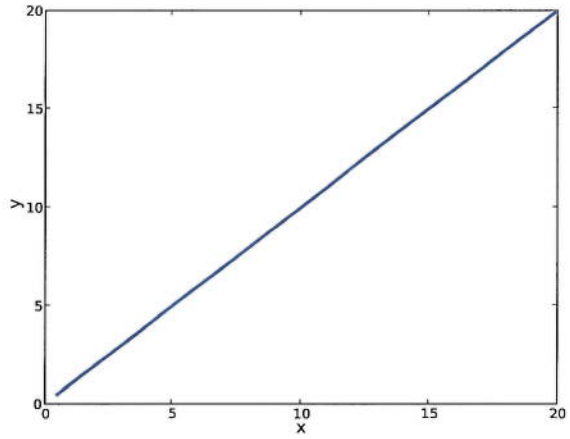


$y = x^{1/2} = \sqrt{x}$

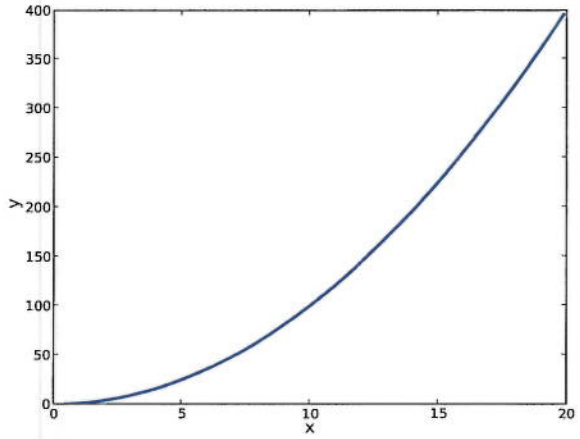
Figure 3.19: Graphs showing the shapes of some power functions.

(continued over)

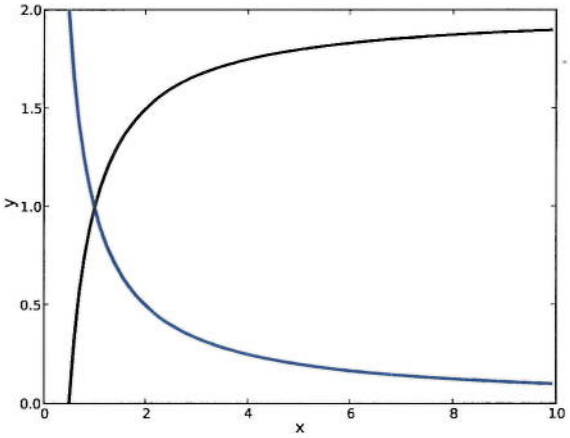
Summary of power functions *(continued)*



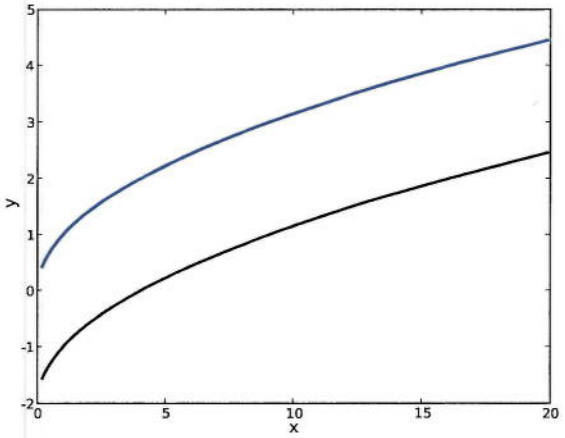
$y = x^1$



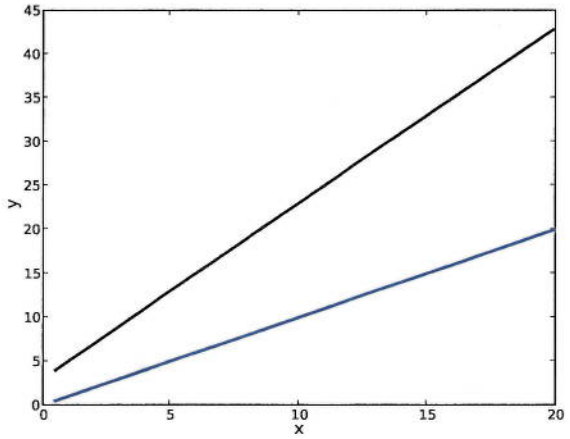
$y = x^2$



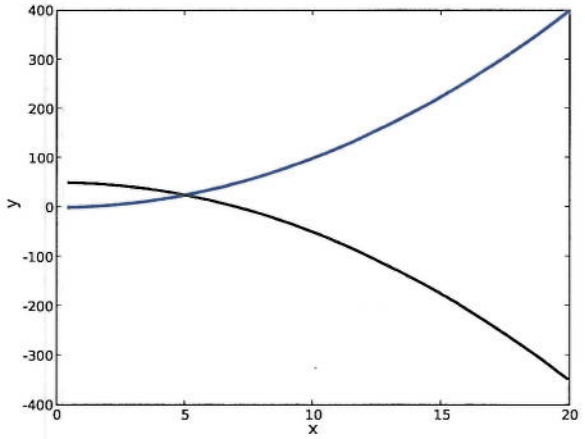
$y = x^{-1}, \quad y = 2 - x^{-1}$



$y = \sqrt{x}, \quad y = \sqrt{x} - 2$



$y = x, \quad y = 2x + 3$



$y = x^2, \quad y = -x^2 + 50$

Figure 3.20: More graphs showing the shapes of some power functions. (Where there are two functions, the graph of the second function is coloured black.)

3.5 Got wind?

- Previous examples have shown how some simple mathematical functions are used to model various phenomena, and how to interpret these models.
- Next we build on the functions we have studied, by combining multiple physical factors into models.
- Rather than a single independent variable (such as time t or area a), the next example considers how *two factors*, ambient temperature and wind speed, combine to change the apparent temperature that we perceive.

Case Study 10: **Wind chill**

- Windy days can feel much colder than calm days, even if ambient air temperatures are the same on both days. The *apparent* temperature to the human body is called the *wind chill* temperature.
- Because wind chill can cause major discomfort, and in cold climates can lead to serious injuries such as frostbite or even death, it is important to measure, model and predict the severity of wind chill.

Question 3.5.1

Derive an equation that models wind chill. (Hint: start by deciding which factors are important, whether they increase or decrease the apparent temperature, whether their effect is linear, and how they interact.)

W = Wind chill, V = Wind velocity, T = ambient temp.

V↑, W↓

T↑, W↑

So $W = \frac{T}{V}$ or $W = T - V$

How could you model the interaction? $\pm TV$

Explaining interaction: As V↑, W↓ but > effect @ low T

Photo 3.8: Blizzard, West Yellowstone, USA. (Source: PA.)

- It is possible to measure wind chill in a number of ways. In 2001, the US National Weather Service developed the most widely accepted model.
- Researchers exposed volunteers to various low temperatures and high wind speeds in a wind tunnel, recording their perceptions of temperatures, along with measurements of the physiological impact of wind chill on their faces.
- The researchers then formulated an equation that modelled the perceived wind chill temperature as a function of the ambient air temperature and the wind speed (for speeds of at least 5 km/h).

Question 3.5.2

Let t be the ambient air temperature in °C and v be the wind speed in km/h. The perceived wind chill temperature W in °C is:

$$W = 13.112 + 0.6215T - 11.37v^{0.16} + 0.3965Tv^{0.16}$$

Example 3.5.3

On a cold Brisbane bike ride, the ambient temperature is 2 °C and the effective wind speed is 30 km/h. Thus,

$$W \approx 13.112 + 1.24 - 11.37 \times 1.723 + 0.79 \times 1.723 \approx -3.88 \text{ °C}$$

We can now develop a computer model for calculating wind chill.

Program specifications: Write a program that inputs wind speed in km/h and air temperature in °C, then calculates the apparent wind chill temperature.

Program 3.1: Wind chill

```
# A program to calculate apparent wind chill temperatures.
from pylab import *

airT = eval(input("Enter air temp. in degrees Celsius: "))
windS = eval(input("Enter wind speed in km/h: "))
x = pow(windS,0.16)
windC = 13.112 + 0.6215 * airT - 11.37 * x + 0.3965 * airT * x
rt = round(windC,1)

print("An air temp. of ",airT," Celsius and wind speed of")
print( windS,"km/h gives a wind chill of",rt," Celsius.")
```

Here is the output from running the above program twice:

```
Enter air temp. in degrees Celsius: -19
Enter wind speed in km/h: 19
An air temp. of  -19  Celsius and wind speed of
19 km/h gives a wind chill of -29.0  Celsius.

Enter air temp. in degrees Celsius: -36
Enter wind speed in km/h: 135
An air temp. of  -36  Celsius and wind speed of
135 km/h gives a wind chill of -65.5  Celsius.
```

Photo 3.9: Mont Blanc. (Source: PA.)

Example 3.5.4

Figure 3.21 shows the average maximum temperatures and wind speeds on the summit of Mount Everest for the years 2002 – 2004 (see [16]), and Figure 3.22 shows the corresponding wind chill temperatures.

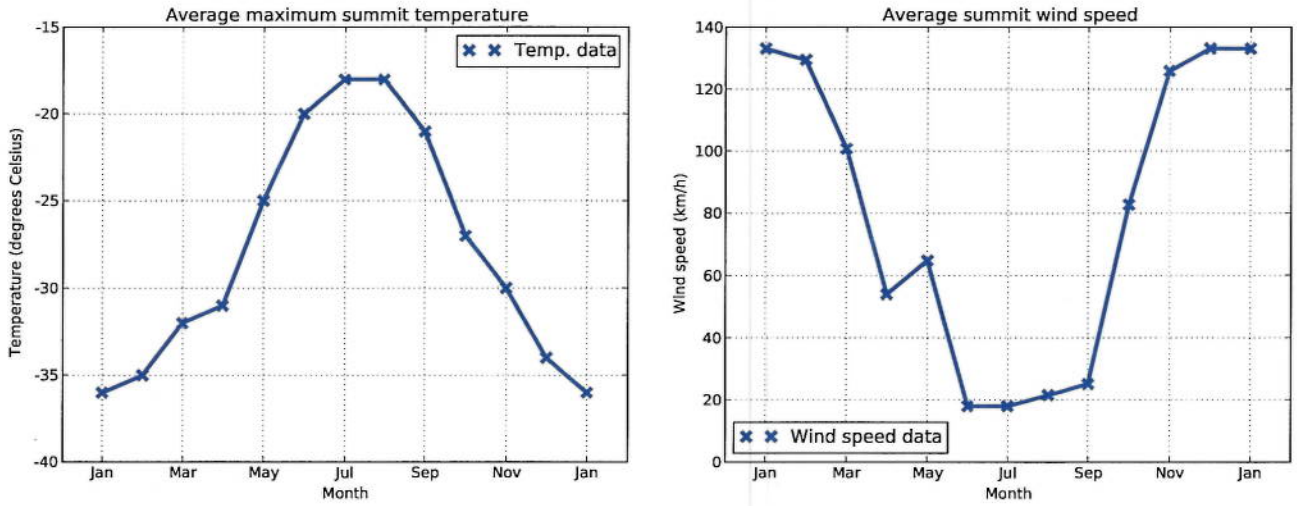


Figure 3.21: Wind speeds and temperatures on the summit of Mount Everest.

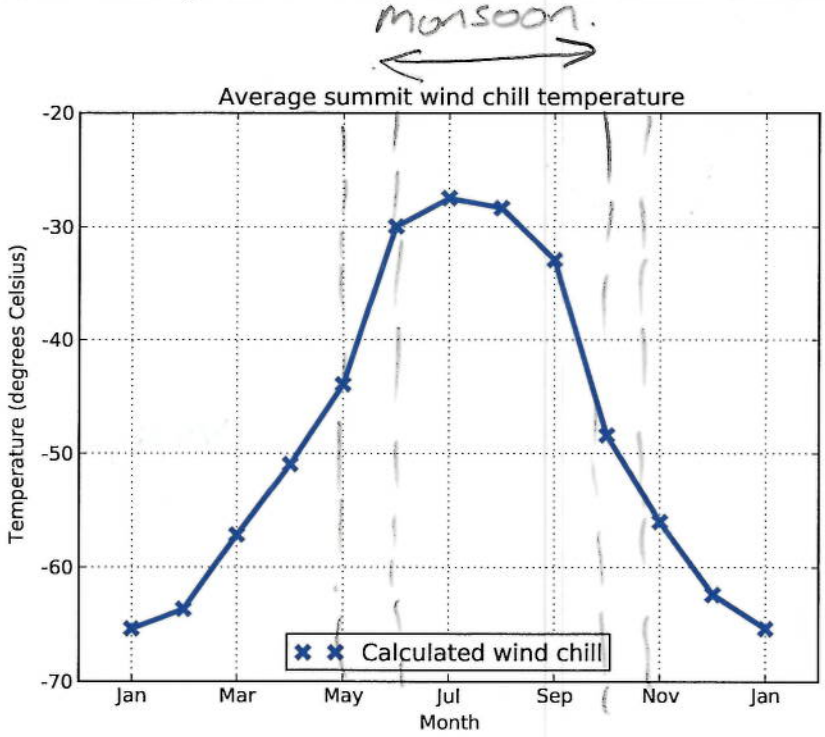


Figure 3.22: Wind chill temperatures on the summit of Mount Everest

There are two very short annual "windows" during which conditions are typically most suitable for ascending to the summit of Mount Everest: May 20 to June 6, and Oct 1 to Oct 20.

- A common way to present information from the wind chill model is via a table of values, often with colour coding to show the risk of developing *frostbite*; see Figure 3.23.

Table-based model

Wind Speed (km/h)	Air temperature (degrees Celsius)												
	10	5	0	-5	-10	-15	-20	-25	-30	-35	-40	-45	-50
5	10	4	-2	-7	-13	-19	-24	-30	-36	-41	-47	-53	-58
10	9	3	-3	-9	-15	-21	-27	-33	-39	-45	-51	-57	-63
15	8	2	-4	-11	-17	-23	-29	-35	-41	-48	-54	-60	-66
20	7	1	-5	-12	-18	-24	-30	-37	-43	-49	-56	-62	-68
25	7	1	-6	-12	-19	-25	-32	-38	-44	-51	-57	-64	-70
30	7	0	-6	-13	-20	-26	-33	-39	-46	-52	-59	-65	-72
35	6	0	-7	-14	-20	-27	-33	-40	-47	-53	-60	-66	-73
40	6	-1	-7	-14	-21	-27	-34	-41	-48	-54	-61	-68	-74
45	6	-1	-8	-15	-21	-28	-35	-42	-48	-55	-62	-69	-75
50	5	-1	-8	-15	-22	-29	-35	-42	-49	-56	-63	-69	-76
55	5	-2	-8	-15	-22	-29	-36	-43	-50	-57	-63	-70	-77
60	5	-2	-9	-16	-23	-30	-36	-43	-50	-57	-64	-71	-78
65	5	-2	-9	-16	-23	-30	-37	-44	-51	-58	-65	-72	-79
70	5	-2	-9	-16	-23	-30	-37	-44	-51	-58	-65	-73	-80

Risk of developing frostbite:	
Low:	< 5% chance of developing frostbite
Increasing:	5% - 95% chance of developing frostbite in 10 to 30 mins.
High:	> 95% chance of developing frostbite in 5 to 10 mins.
Very high:	> 95% chance of developing frostbite in 2 to 5 mins.
Extreme:	> 95% chance of developing frostbite in 2 mins.

Figure 3.23: Wind chill temperatures at various ambient temperatures and wind speeds, colour-coded with frostbite risk factors.

- Frostbite is a medical condition in which intense cold causes tissues to freeze and die, most commonly in body extremities, particularly fingers and toes.
- Severe cases can lead to gangrene and the need for amputations.

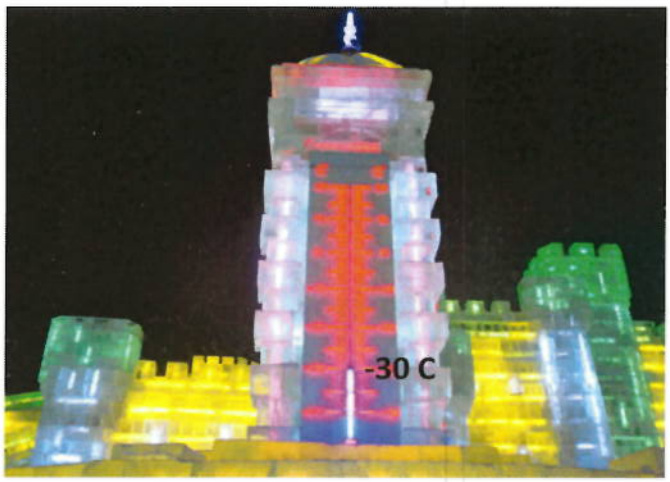


Photo 3.10: A chilly night out. (Source: PA.)

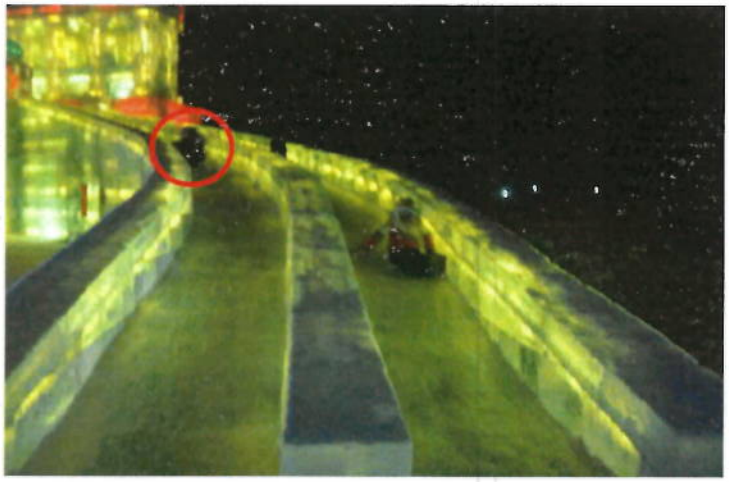


Photo 3.11: An idiot on an ice slippery-slide. (Source: PA.)

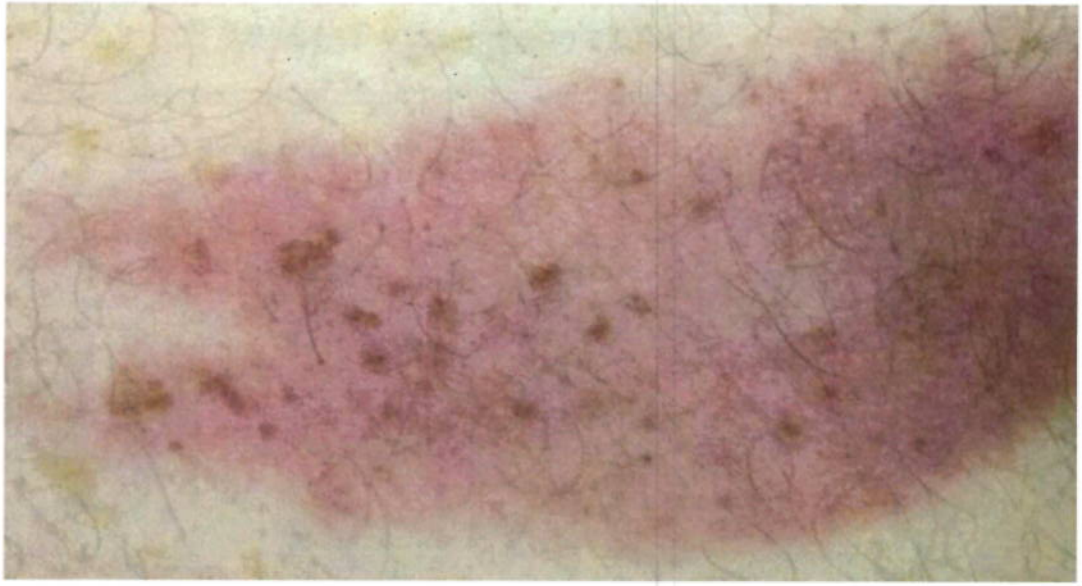


Photo 3.12: Frostbite (Source and sufferer: PA.)

End of Case Study 10: Wind chill.

Now we develop another mathematical model of the Keeling curve.

Question 3.5.5

Keeling Model 3: Figure 3.24 shows two plots: a graph of the function $y(t) = 1/3 \times t^{1.37} + 315$, and the Keeling curve.

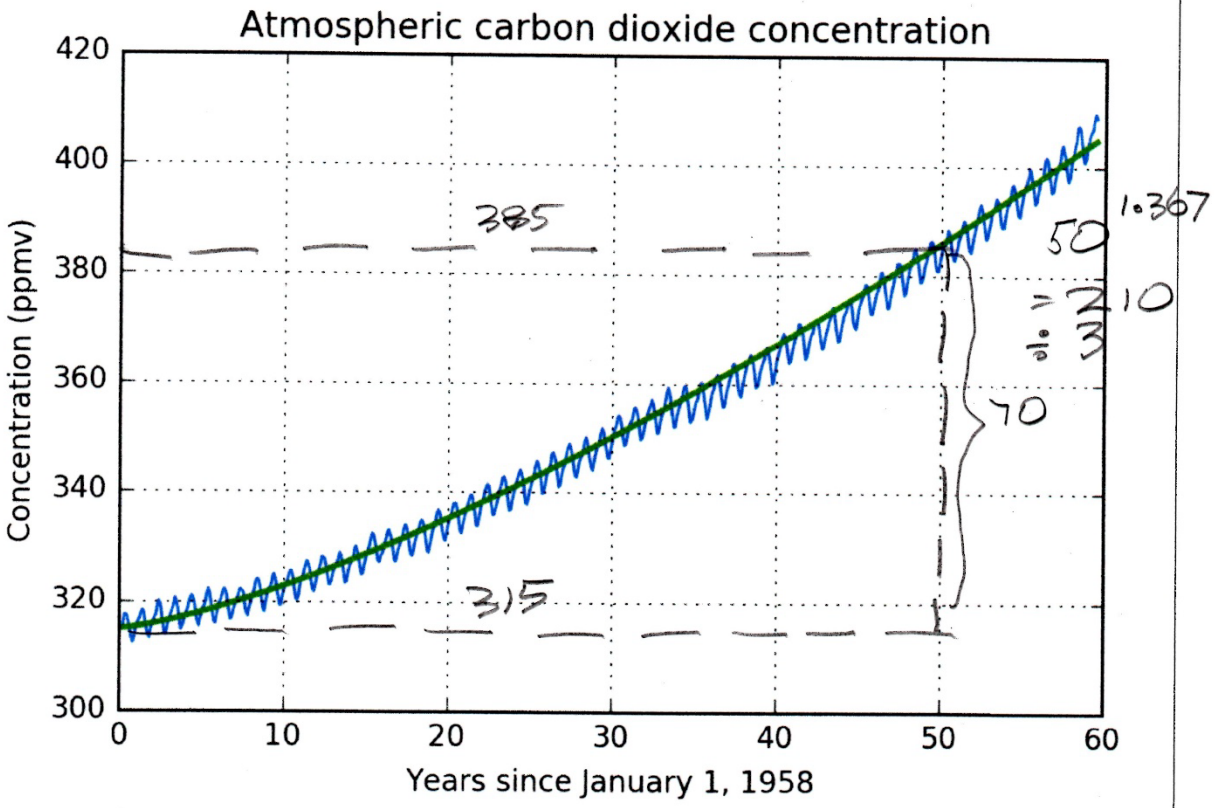


Figure 3.24: The Keeling curve and a power function model.

(a) Explain how each term in $y(t)$ impacts on the graph.

$y = \frac{1}{3} t^{1.367} + 315$

$\frac{1}{3}$: scaling

$t^{1.367}$: power fn (increases)

315: y-intercept

(b) How effectively does $y(t)$ model the Keeling curve?

good general trend
does not account for annual cycles

Chapter 4: **Give us a wave!**

Goodbye papa it's hard to die
When all the birds are singing in the sky
Now that the spring is in the air
Little children everywhere
When you see them I'll be there.
We had joy we had fun,
We had seasons in the sun.
But the wine and the songs
like the seasons have all gone.

Artist: *Terry Jacks.* Song: *Seasons in the sun.*

Image 4.1: *God creating the Heavens and Earth* (*The separation of Light and Darkness* (left), *The creation of the Sun, Moon and Earth* (centre), *The separation of Land and Water* (right)) (1508 – 1512), Michelangelo (1475 – 1564), Sistine Chapel ceiling, Apostolic Palace, Vatican. (Source: en.wikipedia.org)

4.1 Waves, cycles and periodic functions

- Many phenomena in Science and nature *repeat* or *cycle*. These include: many aspects of weather and climate; ocean waves and tides; physiological processes, such as breathing and hormone levels; sound waves; and the voltages and currents in alternating current electricity.
- Consider the four graphs in Figure 4.1, each of which shows climate-related data for Brisbane over a period of one year. If the graphs were extended over subsequent years, then an approximate cycling pattern would be observed.

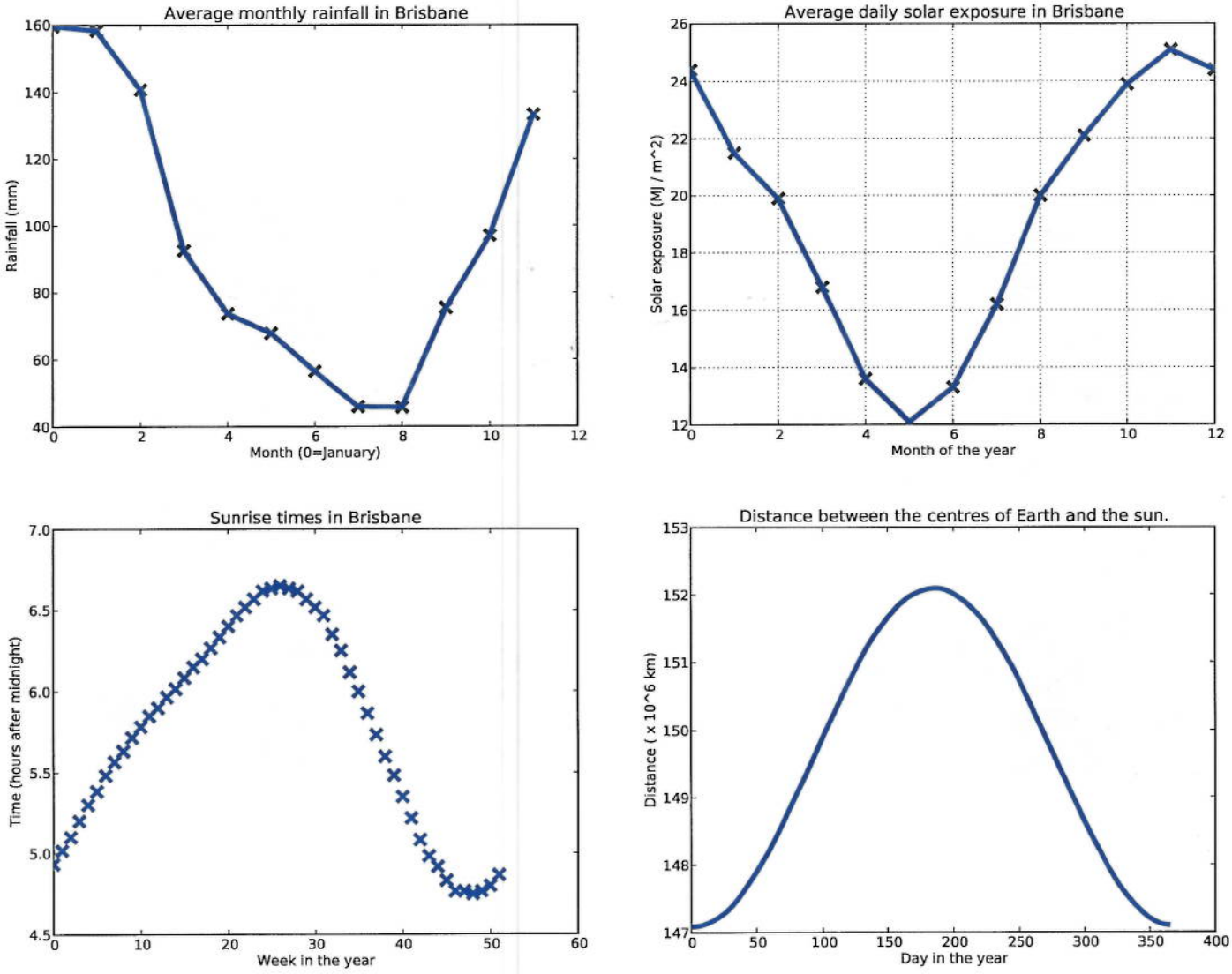


Figure 4.1: Four climate-related graphs. Top left: average monthly rainfall in Brisbane. Top right: average daily solar exposure in Brisbane. Bottom left: weekly sunrise times in Brisbane. Bottom right: daily distances between the centres of Earth and the sun.

Question 4.1.1

Graphs of cyclic phenomena are called *waves*. Using the given graph, identify the following important properties of waves.

- **peaks and troughs** – highest and lowest values on the wave;
- **centre value** – value around which the wave is centred.
- **wavelength** – distance of one **cycle**, from one peak to the next;
- **amplitude** – largest deviation from the centre during a cycle;
- **phase shift** – partial horizontal shift of the wave;
- **period** – time taken for one complete cycle; and
- **frequency** – the rate at which peaks pass a given point, equal to the reciprocal of the period, often measured in cycles per second, **hertz** or **hz**.

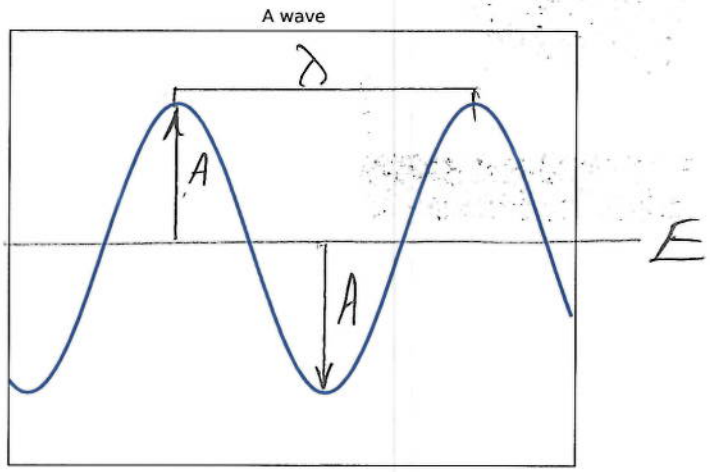


Figure 4.2: The graph of a wave.

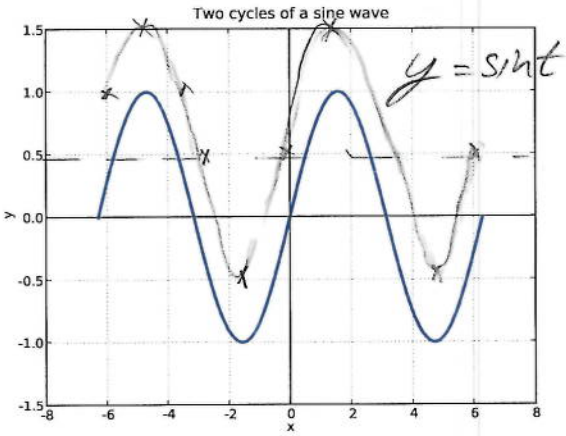
- To represent waves accurately we require a new type of function. It doesn't matter **which** functions are used, provided they give rise to waves.
- The most common choices are the *trigonometric functions* sin and cos.
- These functions are defined in the context of geometry and angles. However, **do not** think of them in a geometric context when modelling.
- They are useful precisely because they cycle, so can be used to model cyclic phenomena. This has nothing directly to do with angles!
- In SCIE1000 we will always use the function sin (we could have used cos).
- By definition, the function $\sin t$ has period 2π, amplitude 1, centre value 0, and equals 0 when $t = 0$.

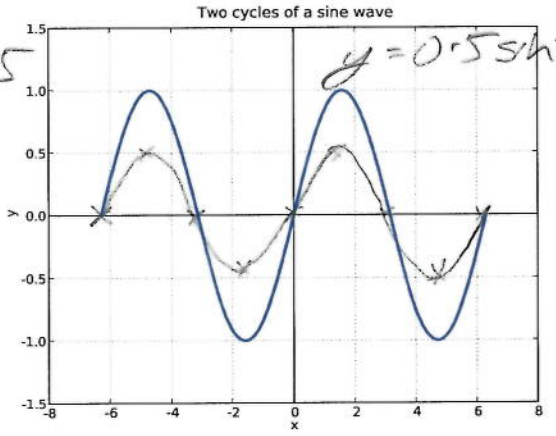
- Varying the values of the "constants" within a sin function alters the cyclic model, allowing us to model a range of cyclic phenomena.

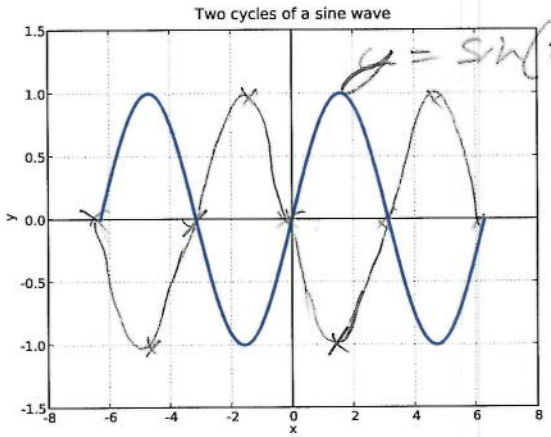
Question 4.1.2

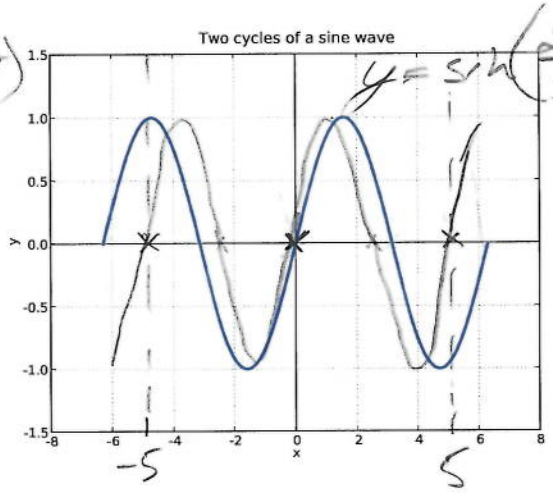
In each case, write the equation for a sin function with the given property, then sketch the new function on the corresponding graph of $y = \sin t$. Each graph shows two full cycles of a *sine wave*.

(a) centred around $y = 0.5$ $y = \sin t + 0.5$	**(b)** an amplitude of 0.5 $y = 0.5 \sin t$
(c) phase shifted by half a cycle $y = \sin(t \pm \pi)$	**(d)** a period of 5 $y = \sin\left(\frac{2\pi}{5} t\right)$









Question 4.1.3

The general equation of a sine function is

$$y(t) = A\sin\left(\frac{2\pi}{P}(t-S)\right) + E.$$

Centre value or Equilibrium.

Explain the meaning and impact of each of the constants A, P, S and E.

Amplitude

Period

Phase shift

Case Study 11: **Heavy breathing**

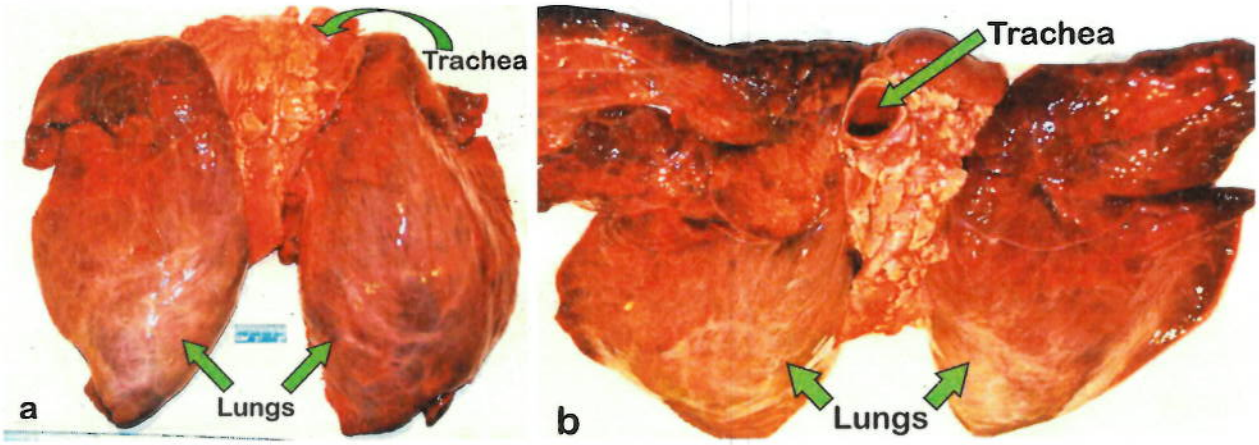


Photo 4.1: Calf lungs, (a) from the front, and (b) from the top. (Source: PA.)

- Normal breathing involves rhythmic inhalation and exhalation of air. The **tidal volume** is the total volume of air breathed in and out with normal breathing. After each exhalation the lung retains a volume of air, called the **functional residual capacity**.
- The volume and rate of air movement into and out of the lungs can be measured using a *spirometer* and graphed in a *spirogram*. (One common type of spirometer uses the Hagen-Poiseuille equation to measure air flow rates.) This information can be used to diagnose possible respiratory impairment.

(*) Coke bottle experiment

Question 4.1.4

(a) Estimate the tidal volume, period between breaths and vital capacity for a resting adult human, and sketch a rough graph of lung capacity (that is, the volume of air in the lung) over time.

(1) Rate = breaths/min => 12 breaths/min = 1 breath in 5s

$P = 5$ (1 cycle of 2π in 5s)

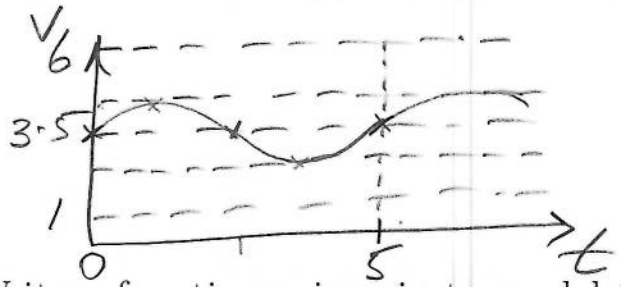
(2) Tidal volume = normal breadth in balloon = 500 ml => $A = 250$ ml $= 0.25$ L

(3) Chest volume = $L \times B \times H = 30\text{ cm} \times 20\text{ cm} \times 10\text{ cm} = 0.3\text{m} \times 0.2\text{m} \times 0.1\text{m}$

$= 0.006\text{ m}^3$

$= \pi r^2 h = 3 \times 10^2 \times 20 = 6$ L

$= 6$ L



(4) Vital capacity = big breath in bag = 5L

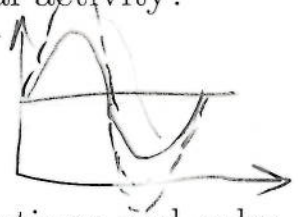
∴ functional residual = $6 - 5 = 1$L

(b) Write a function using sin to model the lung capacity in Part (a).

$y = A \sin\left(\frac{2\pi}{P}t - \underline{S}\right) + E = 0.25 \sin\left(\frac{2\pi}{5}t\right) + 3.5$

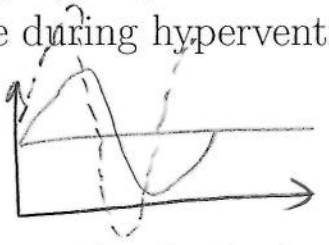
(c) How would the function change after moderate physical activity?

Exercise $A\uparrow$, curiously $P \rightarrow$



(d) Hyperventilation is characterised by rapid, deep inhalations and exhalations. How would the function change during hyperventilation?

Hyperventilation $P\downarrow A\uparrow$



(e) Smoking and air pollution causes inflammation in the lungs, gradually destroying the lung tissue and leading to *emphysema*, a type of Chronic Obstructive Pulmonary Disease (COPD). The reduction of lung surface area decreases the ability to exchange carbon dioxide and oxygen.

(continued over)

Question 4.1.4 (continued)

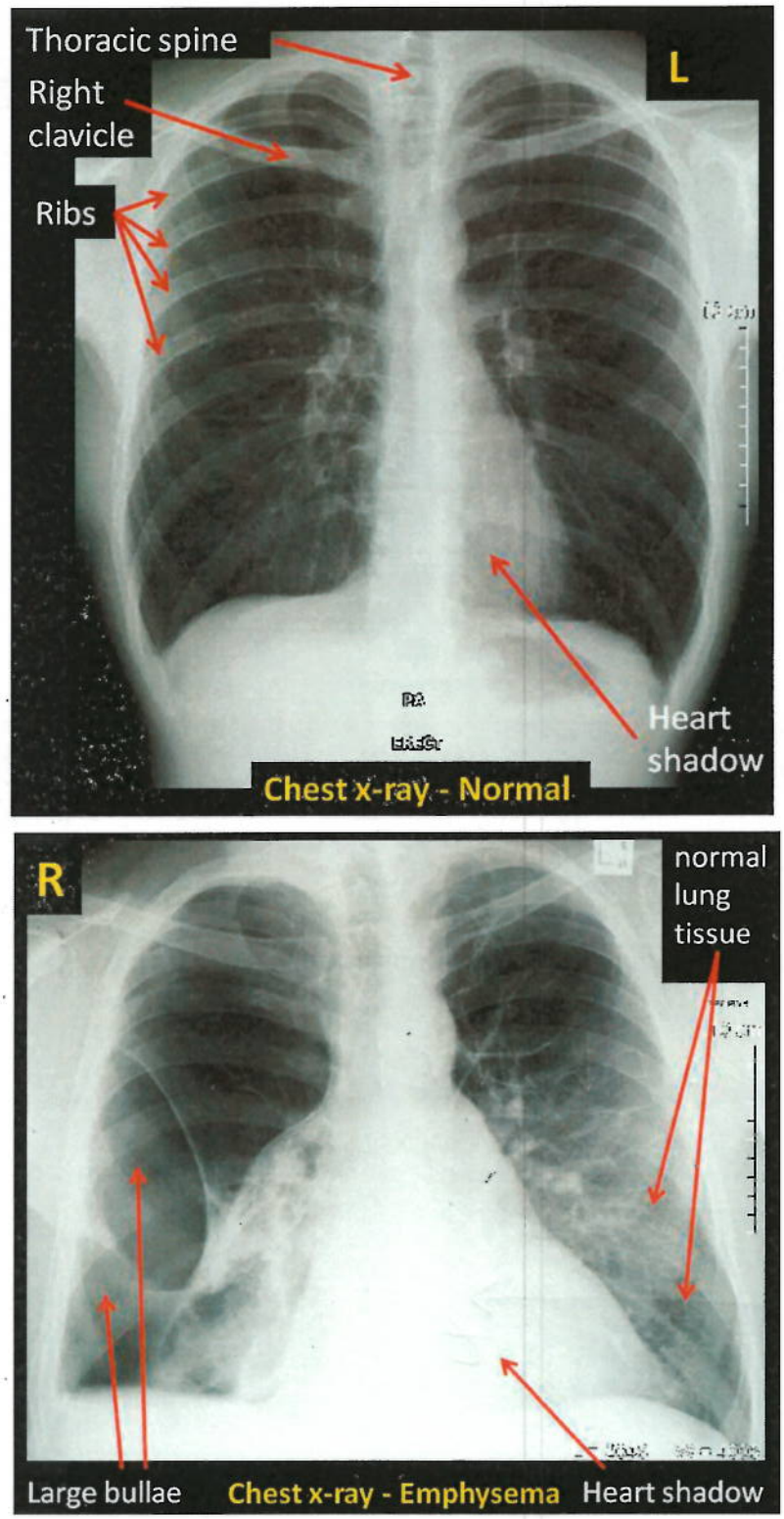


Photo 4.2: Top: x-ray of an adult male chest displaying normal lung tissue architecture and normal heart shadow. Bottom: x-ray of a chest showing large emphysematous bullae within the right lung. (Source: Qld Health and DM.)

(continued over)

***Question** 4.1.4 (continued)*

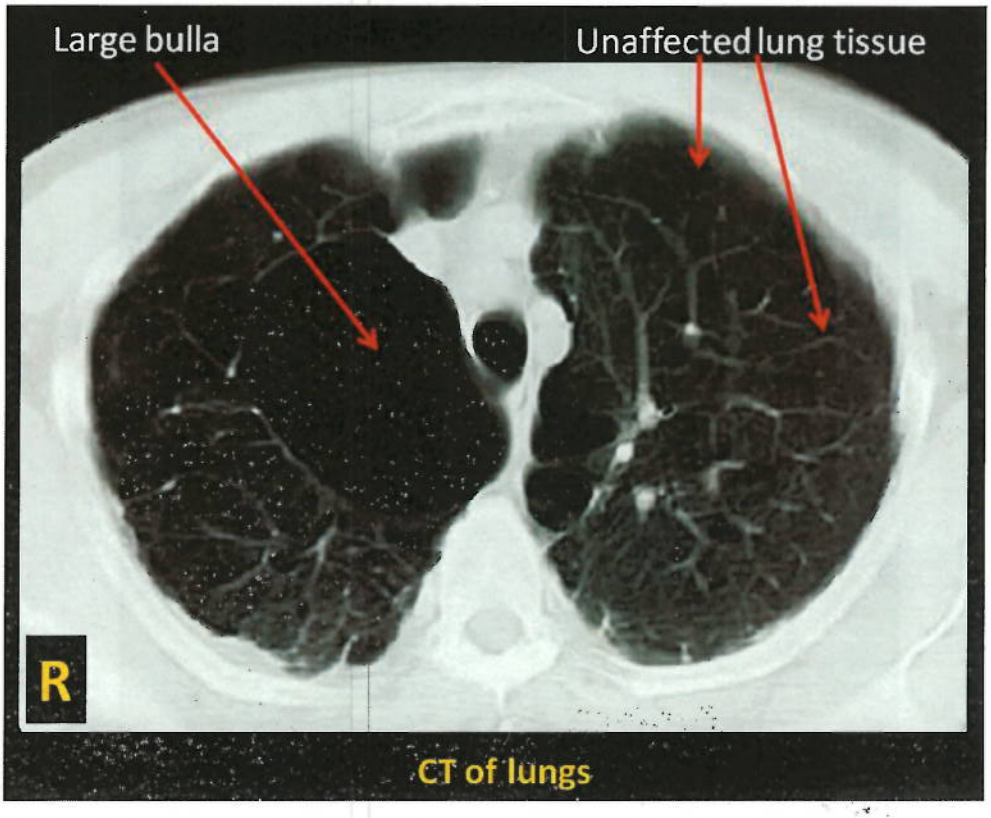
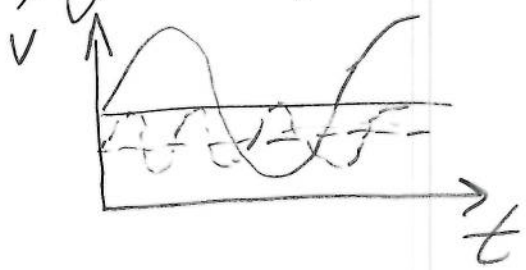


Photo 4.3: Axial CT showing one large, and multiple small, bullae of the alveolar air spaces in the right lung. (Source: Qld Health and DM.)

How would the function change for an individual with emphysema?

Emphysema: P↓ A↓ E↓



End of Case Study 11: Heavy breathing.

4.2 Days, seasons, cycles

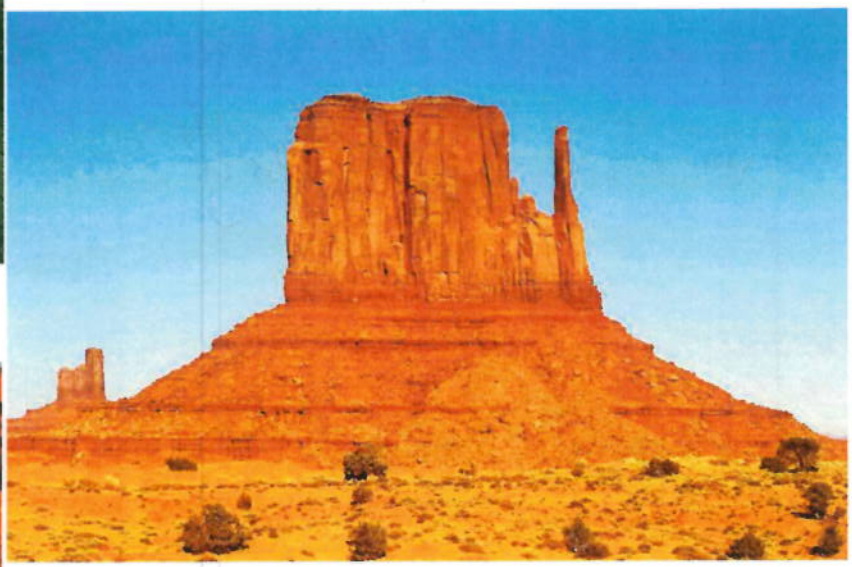


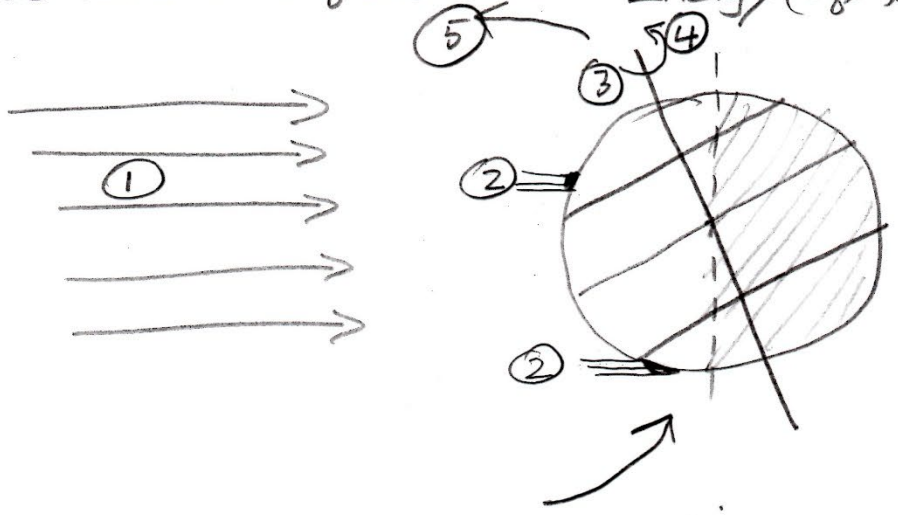
Photo 4.4: Spring – Lotus flower, *Nelumbo nucifera* (Tokyo, Japan); Summer – Monument Valley (Utah, USA); Autumn – sugar maple, *Acer saccharum* (Vermont, USA); Winter – Pine tree (Canyonlands, Utah, USA). (Source: PA.)

- The amount of sunlight available at a location on Earth on a given day can be modelled using *daytime*, defined as the time between sunrise and sunset. (This is independent of clouds or weather events.)

- Daytime lengths vary through the year. The **summer solstice** and **winter solstice** are the days with the longest and shortest daytimes (respectively). The **vernal equinox** and **autumnal equinox**, are the days in spring and autumn (respectively) with daytimes of exactly 12 hours.

Question 4.2.1

What causes seasons? Define! Quarterly. Wi-Sp-Su-Au NH/SH phase shift
Climate variation dependant on Energy (eg temp.)



① Energy source = solar radiation
② Earth round — Latitudinal variation in incidence/surface area of solar rays
③ Earth tilted (obliquity) — axis fixed 23.45° explains NH/SH variation
④ Earth rotates (24 hr cycle) (RH rule)
[if no rotation, 6 month light/hot, 6 month dark/cold]
[if annual, one face constantly exposed to sun]
⑤ Earth circles sun (annual cycle) (RH rule)
[almost circular orbit
[not elliptical - centric
" " = eccentric hot cold } whole planet

Question 4.2.2

Discuss the daytime lengths in midsummer and midwinter in each of:

(a) Brisbane;

midsummer: 13·5 hrs
midwinter: 10·5 hrs
} $12 \pm 1{\cdot}5$

(b) Singapore (which is very close to the equator); and

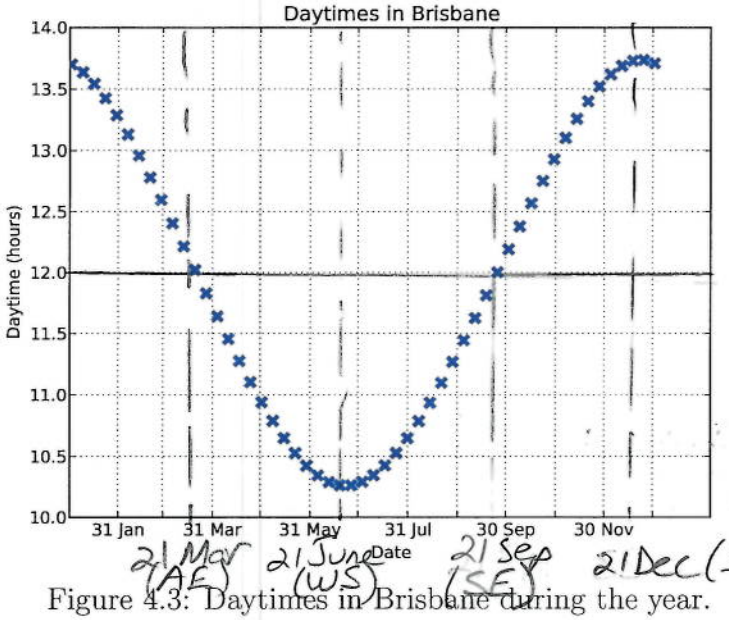
midsummer: 12·5 hrs
midwinter: 11·5 hrs
} $12 \pm 0{\cdot}5$

(c) Santa Claus village, Rovaniemi, Finland (north of the Arctic Circle).

midsummer: 24 hrs
midwinter: 0 hrs
} 12 ± 12

Photo 4.5: Top left: road sign to Santa (Rovaniemi, Finland). Right: the official home of Santa (Santa Claus Village, Finland). Bottom left: Singapore. (Source: PA.)

- At large distances from the equator, summer daytimes are very long; on some occasions there is no sunrise or sunset for a period greater than one day. For simplicity, in such cases we say that the daytime is 24 hours.
- Similarly, in midwinter we say that the daytime is 0 hours.
- Figure 4.3 shows the daytimes in Brisbane at weekly intervals from Friday 1/1/2010 to Friday 31/12/2010.[1] The graph of daytime lengths in every year will be very similar; clearly, the graph resembles a sine wave!



SS = Summer Solstice
SE = Spring Equinox
WS = Winter Solstice
AE = Autumn Equinox

Figure 4.3: Daytimes in Brisbane during the year.

Question 4.2.3

Use the graph in Figure 4.3 to answer the following questions.

(a) When are the solstices in Brisbane, and how long are the daytimes?

SS: 13·7 hrs Dec 21

WS: 10·3 hrs Jun 21

(b) When are the equinoxes in Brisbane?

AE: Mar 21 12 hrs

SE: Sep 21 12 hrs

[1] Daytimes were found by subtracting the sunrise time from the sunset time. Sunrise time is defined as the time at which any part of the sun is first visible on a clear, cloudless day. Sunset time is defined as the time at which any part of the sun is first **not** visible on a clear, cloudless day. This definition of sunset differs slightly from standard usage.

Question 4.2.4

On some international flights, in-flight maps show areas of night and day superimposed on the surface of Earth; see Figure 4.6 for two examples.

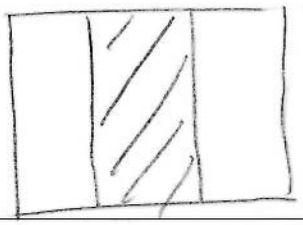
Photo 4.6: In-flight maps. (Source: PA.)

(a) Describe everything you can about the date and time of day when the first photograph was taken. Justify your answer.

Looks like SS in NH & WS in SH
ie. around June 21.

(b) Describe everything you can about the date and time of day when the second photograph was taken. Justify your answer.

Looks like an equinox - ie. equal daytime everywhere



So AE Mar 21 or SE Sep 21
Either just before SE in SH
or just after AE in SH

Case Study 12: **Modelling daytimes**

Photo 4.7: Sunrise over Kunming Lake in winter, Beijing, China. (Source: PA.)

Photo 4.8: Sunset over prison guard tower, near Krakow, Poland. (Source: PA.)

- Every location on Earth has a *latitude*, which is a measure of its distance from the equator. On any given day, **every location with the same latitude has the same daytime length.**
- At each location on Earth, the daytimes repeat in a yearly pattern. Therefore, they can be modelled using sin, as a function of the day of the year.
- (In reality, daytimes will vary slightly from these functions as days are discrete time steps whereas the Sun and Earth move continuously.)

Question 4.2.5

If t is the day number in the year (starting from $t = 0$ on January 1st) then the length of the daytime in hours at any point in the southern hemisphere is given by the function

$$D(t) = 12 + K \times \sin\left(\frac{2\pi}{365}(t - 264)\right)$$

where K is a constant determined by the latitude of the point. At the equator $K \approx 0$, and its value increases for more southerly locations. For Brisbane $K \approx 1.74$; the graph with $K = 1.74$ is plotted in Figure 4.4.

(continued over)

Question 4.2.5 *(continued)*

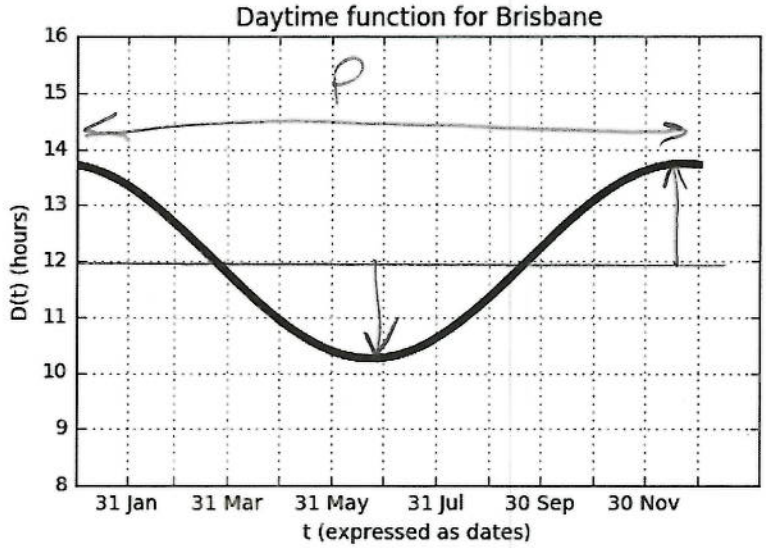


Figure 4.4: The daytime function for Brisbane.

Recall that, in Brisbane, $D(t) = 12 + 1.74 \sin\left(\frac{2\pi}{365}(t - 264)\right)$.

Discuss the physical and mathematical significance of each term in $D(t)$.

Centre point
= mean daytime

Amplitude
= dependent
on latitude

Shift
= Sep 21
= SE

Period
= 1 cycle (2π)
in 365 days

Question 4.2.6

Recall that, in Brisbane, $D(t) = 12 + 1.74\sin\left(\frac{2\pi}{365}(t-264)\right)$. = Box.

Answer each of the following from the function (see Question 4.2.3).

(a) When are the solstices in Brisbane, and how long are the daytimes?

SS: $\sin\square = 1$, when $\square = \frac{\pi}{2}$

$\Rightarrow \frac{2\pi}{365}(t-264) = \frac{\pi}{2} \Rightarrow t - 264 = \frac{365}{4}$

$\Rightarrow t = 355$ (= Dec 21) $D(355) = 12 + 1.74 \times 1$

$= 13.74$ hrs

WS: $\sin\square = -1$, when $\square = \frac{3\pi}{2}$

$\Rightarrow \frac{2\pi}{365}(t-264) = \frac{3\pi}{2} \Rightarrow t - 264 = \frac{3\pi}{2} \times \frac{365}{2\pi}$

$\Rightarrow t = 538$ WTF! $(-365 = 173)$ = Jun 21

$D(173) = 12 + 1.74 \times -1 =$ 10.26 hrs

(b) The equinoxes have daytimes of length 12 hours everywhere in the world. Explain briefly how to find the dates of the equinoxes.

SE & AE: $\sin\square = 0$ when $\square = 0, \pi$

$\frac{2\pi}{365}(t-264) = 0 \Rightarrow t = 264$ (Sep 21, SE)

OR $\frac{2\pi}{365}(t-264) = \pi \Rightarrow t - 264 = \frac{365}{2} \Rightarrow t = 446$

$(-365 = 81)$

= Mar 21 (AE)

Question 4.2.7

In $D(t)$, $K \approx 1$ for Townsville, $K \approx 1.74$ for Brisbane, and $K \approx 3.3$ for Hobart. The graph for Brisbane is shown in Figure 4.5.

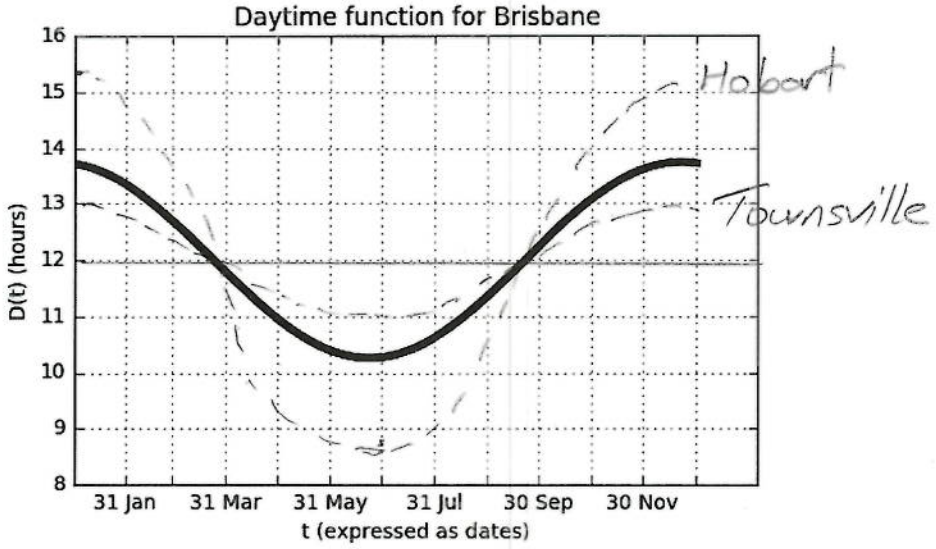


Figure 4.5: Daytimes in Brisbane over the year.

(a) Roughly sketch the graphs of $D(t)$ for Townsville and Hobart on the above graph.

(b) By how much is the daytime on the **summer** solstice in Hobart **longer** than in Townsville? By how much is the daytime on the **winter** solstice in Hobart **shorter** than in Townsville?

SS = 3.3 − 1 = 2.3 hrs

WS = 1 − 3.3 = −2.3 hrs

(c) What does your answer suggest for the total amount of daytime in a year at any location in the southern hemisphere? Is it true, and what is the ecological significance?

What latitudes have the most sunlight?

All the same. Total sunlight = 12 hrs × 365 days.

=> Migration

Question 4.2.8

Recall that the daytime equation in the southern hemisphere is

$$D(t) = 12 + K \times \sin\left(\frac{2\pi}{365}(t - 264)\right).$$

The corresponding equation in the **northern** hemisphere is

$$N(t) = 12 + K \times \sin\left(\frac{2\pi}{365}(t - 81)\right).$$

(a) With reference to these equations, explain the similarities and differences between daytimes in the northern and southern hemispheres.

Similarity: 12 hrs mean sunlight. Same period = 365 days. Same amplitude = K

Difference: Phase shift

Difference in shift = 264 − 81 = 183 days = ½ a year out of phase

(b) A graph of $D(t)$ at the Antarctic Circle is shown in Figure 4.6. Sketch a rough graph of $N(t)$ at the Arctic Circle, explain your answer and identify the solstices and equinoxes.

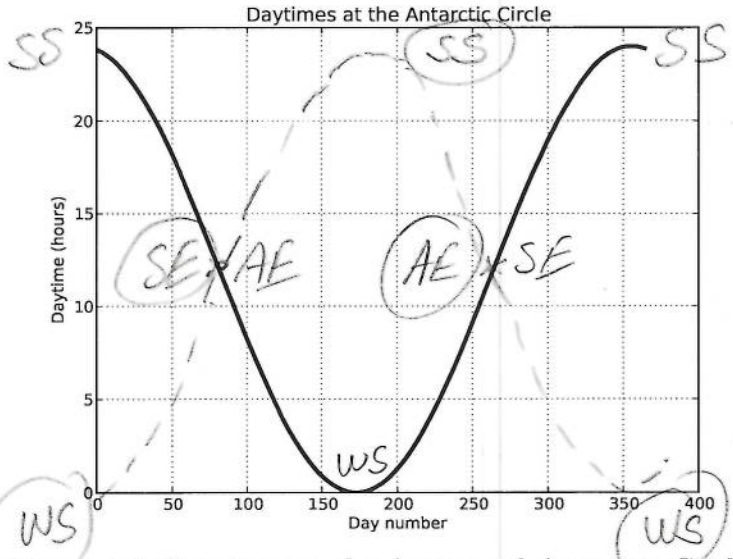


Figure 4.6: Daytimes at the Arctic and Antarctic Circles.

End of Case Study 12: Modelling daytimes.

Case Study 13: To every thing, there is a season, tern, tern, tern.

- Many animals undertake *migration*, during which they move from one area to another, and then return. This often happens on an annual basis, according to seasons or weather patterns.
- Migratory behaviour occurs in all major animal groups (birds, reptiles, mammals, amphibians, fish, insects and crustaceans); see [10].
- Examples of migration include: wildebeest and zebra on the Serengeti plains in Africa; geese "flying south for winter" in the northern hemisphere; humpback whales travelling north along the Queensland coast during winter; and sea turtles returning to beaches to lay eggs.

***Question* 4.2.9**

What are some of the reasons for, and benefits of, seasonal migration? How does this relate to daytimes?

Feeding & breeding.

You can see, so sociality. Plants photosynthesise => food. It is warm when most light.

Photo 4.9: Migrating Canada Geese, *Branta canadensis*, New York State, USA. (Source: PA.)

- The Arctic tern, *Sterna paradisaea*, is a seabird that migrates annually from its breeding grounds in the Arctic to the Antarctic and back.

Image 4.2: Arctic tern in flight. (Source: en.wikipedia.org)

- Individuals have been tracked travelling a distance of 400–700 km per day, and 80000 km in a year; this is the longest (known) migration of any animal.
- Figure 4.7 shows tracked migration routes of 11 Arctic terns (see [12]).

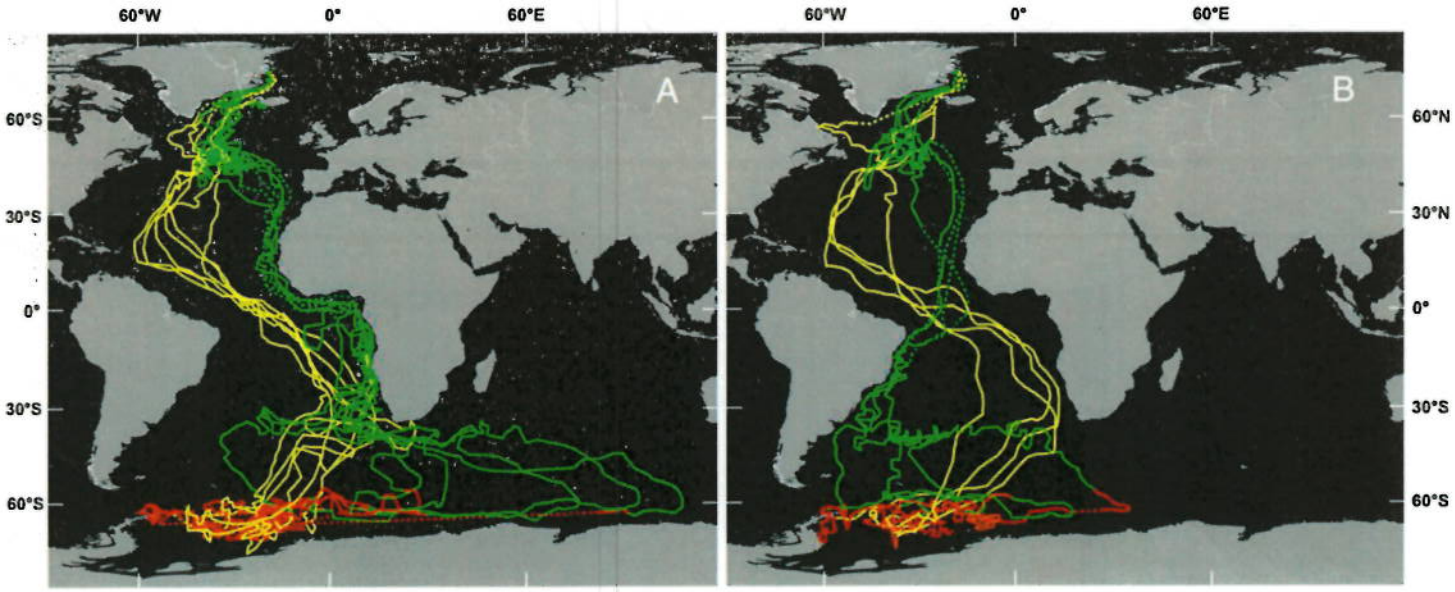


Figure 4.7: Interpolated geolocation tracks of 11 Arctic terns tracked from breeding colonies in Greenland ($n = 10$ birds) and Iceland ($n = 1$ bird). Green = autumn (postbreeding) migration (August/November), red = winter range (December/March), and yellow = spring (return) migration (April/May). Two southbound migration routes were adopted in the South Atlantic, either (A) West African coast (n = 7 birds) or (B) Brazilian coast. (This text is an extract from [12].)

- To collect the data presented in Figure 4.7, researchers in [12] tracked the migration routes of individual birds by attaching miniature archival light loggers to the legs of 20 breeding birds in Iceland in 2007, and to 50 breeding birds in Greenland in July 2007. A year later, data were retrieved from 11 birds.

$D(t) = 12 + K \sin\left(\frac{2\pi}{365}(t-264)\right)$

- The light loggers recorded and stored information about the ambient light intensity at different times and dates, which the researchers used to calculate the time of sunrise and sunset on each day, and hence the daytime.

- They then used the formulae for $N(t)$ and $D(t)$ defined in Question 4.2.8, along with the day numbers and measured daytimes, to calculate the value of K and hence the latitudes of the locations at which readings occurred.

- Next, the calculated latitude and times of sunrise/sunset were used to find the longitude of each reading, and hence pinpoint the location of the bird on each day. Finally, researchers assumed that the birds flew in direct lines between each consecutive pair of readings.

- The paper [12] states that:

> "Locations were unavailable at periods of the year when birds were at very high latitudes and experiencing 24 h daylight. In addition, only longitudes were available around equinoxes, when day length is similar throughout the world. Overall, after omitting periods with light level interference and periods around equinoxes, the filtered data sets contained between 166 and 242 days of locations for each individual."

- 24 hrs light or dark
- Equinoxes

***Question* 4.2.10**

Arctic terns migrate annually between the Arctic and the Antarctic. Figure 4.8 compares daytimes at the Arctic and Antarctic Circles, using graphs of $D(t)$ and $N(t)$ as defined in Question 4.2.8, with $K = 12$.

(continued over)

Question 4.2.10 *(continued)*

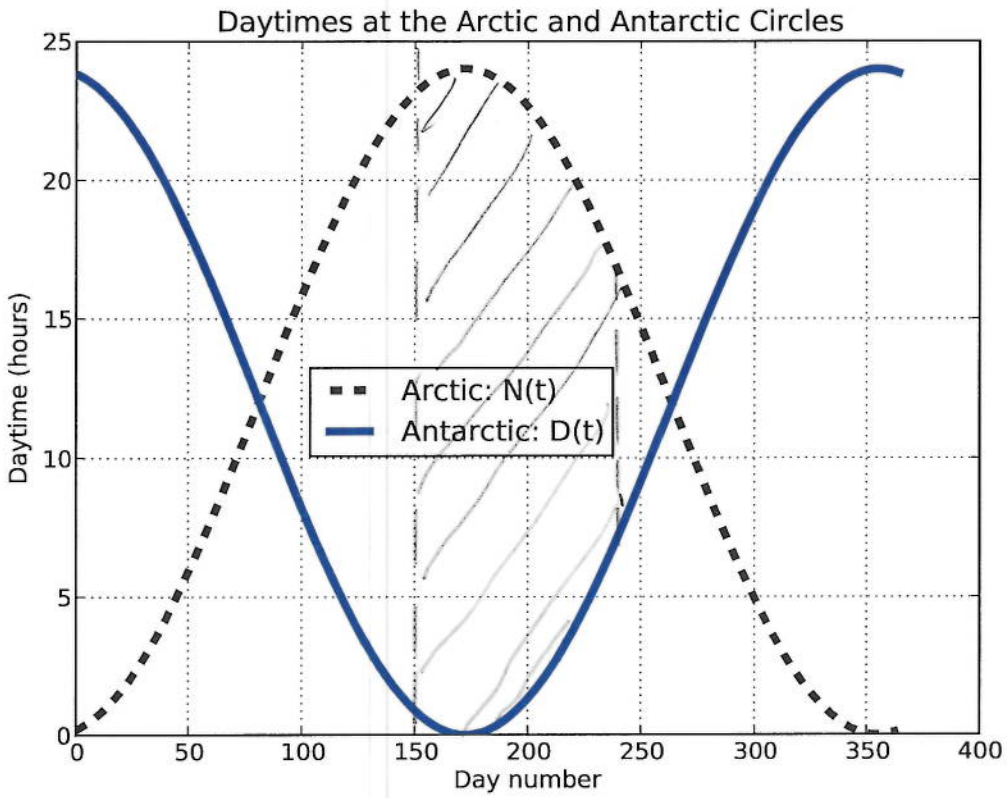


Figure 4.8: Daytimes at the Arctic and Antarctic Circles.

Assume that an individual Arctic tern arrives at its breeding grounds on June 1st (day number 150 of the year).

(a) Roughly how much more daytime occurs at the Arctic circle on June 1st than at the Antarctic Circle?

Arctic: $N(t) = 12 + 12\sin\left(\frac{2\pi}{365}(150-81)\right) = 23.13$ hrs

Antarctic: $D(t) = 12 + 12\sin\left(\frac{2\pi}{365}(150-264)\right) = 0.91$ hrs

More daytime $\sim 23 - 1 = 22$ hrs (or from graph)

(b) On the graph, identify the **total** amount of daytime at the Arctic Circle between June 1 and August 31, and the total **additional** time at the Arctic Circle compared to the Antarctic Circle. August 31 = day 241

(c) Explain how to calculate the value in Part (b) mathematically.

$AUC = \int$ => Amount of sunlight

End of Case Study 13: To every thing, there is a season, tern, tern, tern..

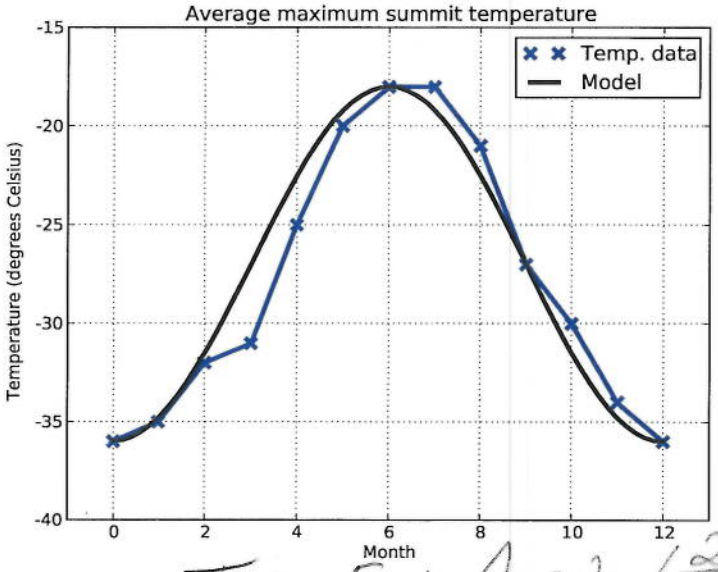
Question 4.2.11

Earlier, Figure 3.21 showed graphs of the measured summit temperatures and wind speeds on Mount Everest over the period 2002 – 2004.

(a) Figure 4.9 graphs the temperature graph, along with the function

$$T(m) = -27 + 9\sin\left(\frac{2\pi}{12}(m-3)\right)$$

where m is the month number from 0 (January) to 12 (next January).



$T = E + A\sin\left(\frac{2\pi}{P}(m-S)\right)$

Figure 4.9: Temperatures on the summit of Mount Everest and the graph of $T(m)$.

Explain the physical and mathematical meaning of each term in $T(m)$. How effectively does $T(m)$ model the temperatures?

$E = -27 =$ Centre value = Mean temperature

$A = 9 =$ Amplitude = Variation around E

$P = 12 =$ Period = 12 months

$S = -3 =$ Shift = Shift to left.

(continued over)

***Question** 4.2.11 (continued)*

(b) Figure 4.10 graphs summit wind speeds, along with the function

$$W(m) = a + b\sin\left(\frac{2\pi}{c}(m-d)\right)$$

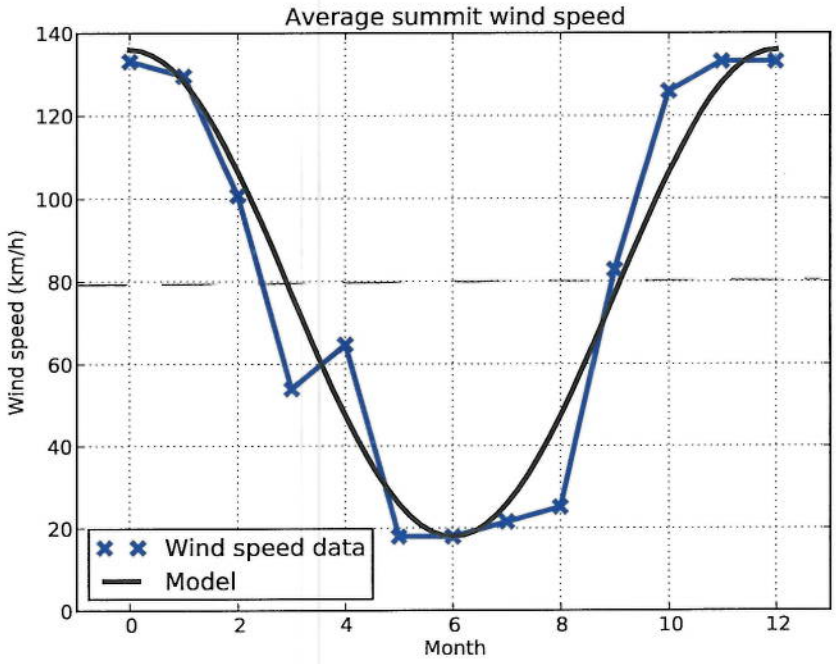


Figure 4.10: Wind speeds on the summit of Mount Everest.

Estimate the values of a, b, c and d in the equation for $W(m)$.

c = Period = 12. a = Equilibrium = 80

b = Amplitude = $\frac{1}{2}$(max − min) = $\frac{1}{2}$(138 − 18) = 60

d = Shift = 9 months

$\therefore$ $W(m) = 80 + 60\sin\left(\frac{2\pi}{12}(m-9)\right)$

Now we develop a more accurate mathematical model of the Keeling curve.

Question 4.2.12

Keeling Model 4: Figures 4.11, 4.12 and 4.13 plot the Keeling curve and the following function $y(t)$ over three different time periods.

$$y(t) = \frac{1}{3}t^{1.37} + 315 + 3.5\sin\left(\frac{2\pi}{1}(t - 0.15)\right).$$

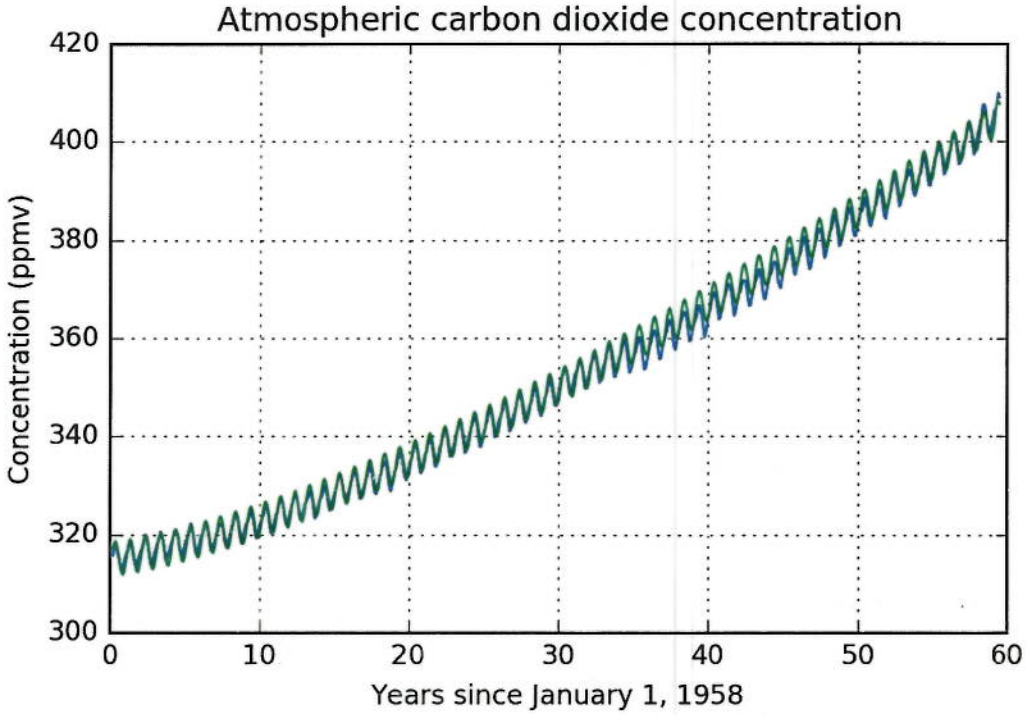


Figure 4.11: The Keeling curve and a model using sin and power functions.

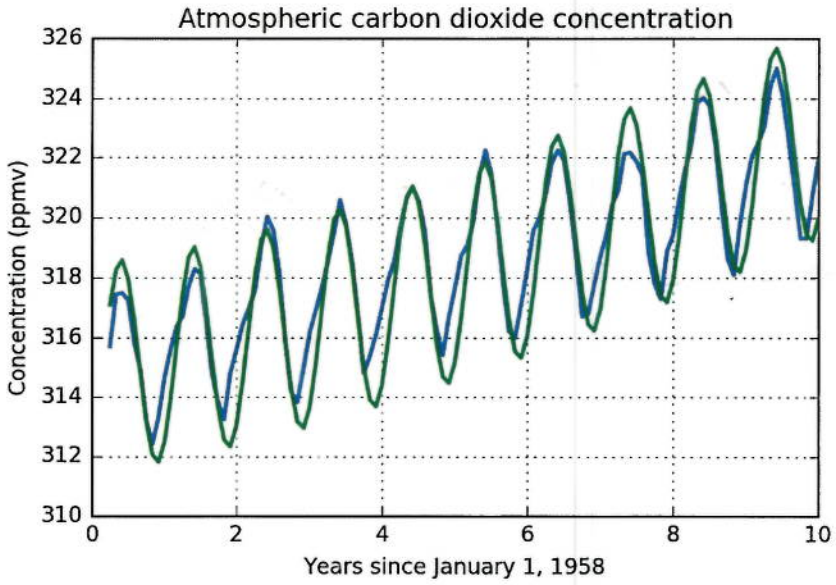


Figure 4.12: The Keeling curve and a model using sin and power functions (early years).

(continued over)

Question 4.2.12 *(continued)*

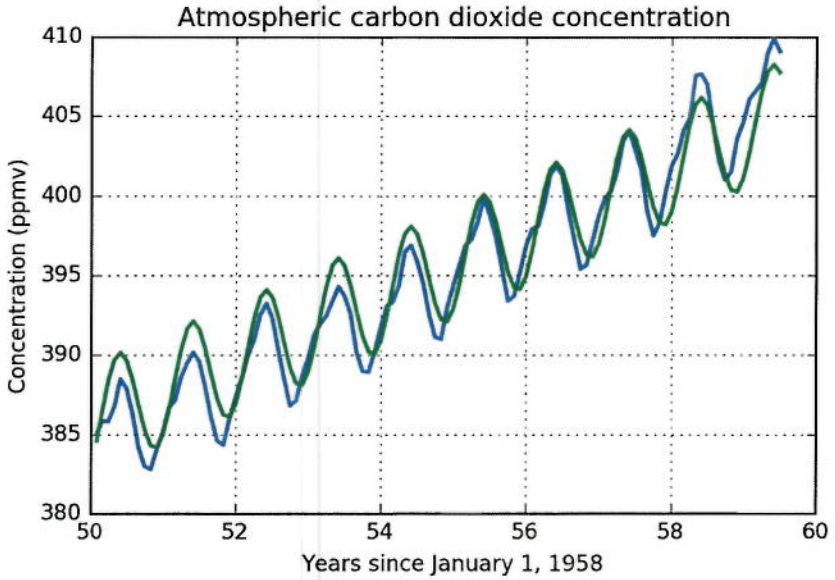


Figure 4.13: The Keeling curve and a model using sin and power functions (recent years).

(a) Explain how each term in $y(t)$ impacts on its graph. Recall that

$$y(t) = \frac{1}{3}t^{1.37} + 315 + 3.5\sin\left(\frac{2\pi}{1}(t - 0.15)\right).$$

315 = vertical shift

$\frac{1}{3}$ = scaling factor

$t^{1.37}$ = power equation curving up

3.5 = amplitude

1 = Period

0.15 = Shift.

(b) Given a physical justification for the term $(t - 0.15)$.

The shift of 0.15 is the fraction of the year that the cycle is shifted left.

(c) How effectively does $y(t)$ model the Keeling curve?

Excellently! Non-linear trend + wave!

Chapter 5: **Exponentials and logarithms**

Dum dum, diddle dum dum,
diddle dum dum, diddle dum dum.
There was a turtle by the name of Bert
And Bert the Turtle was very alert
When danger threatened him he never got hurt
He knew just what to do. (bang)
He'd duck (quack) and cover, duck (quack) and cover.
He did what we all must learn to do
You and you and you and you. (bang)
Duck (quack) and cover!

(*) Play youtube video:
Duck & Cover (1951)
Bert the Turtle
Nuclear threat.

Artist: *US Federal Government Civil Defense.* Song: *Duck and cover.*

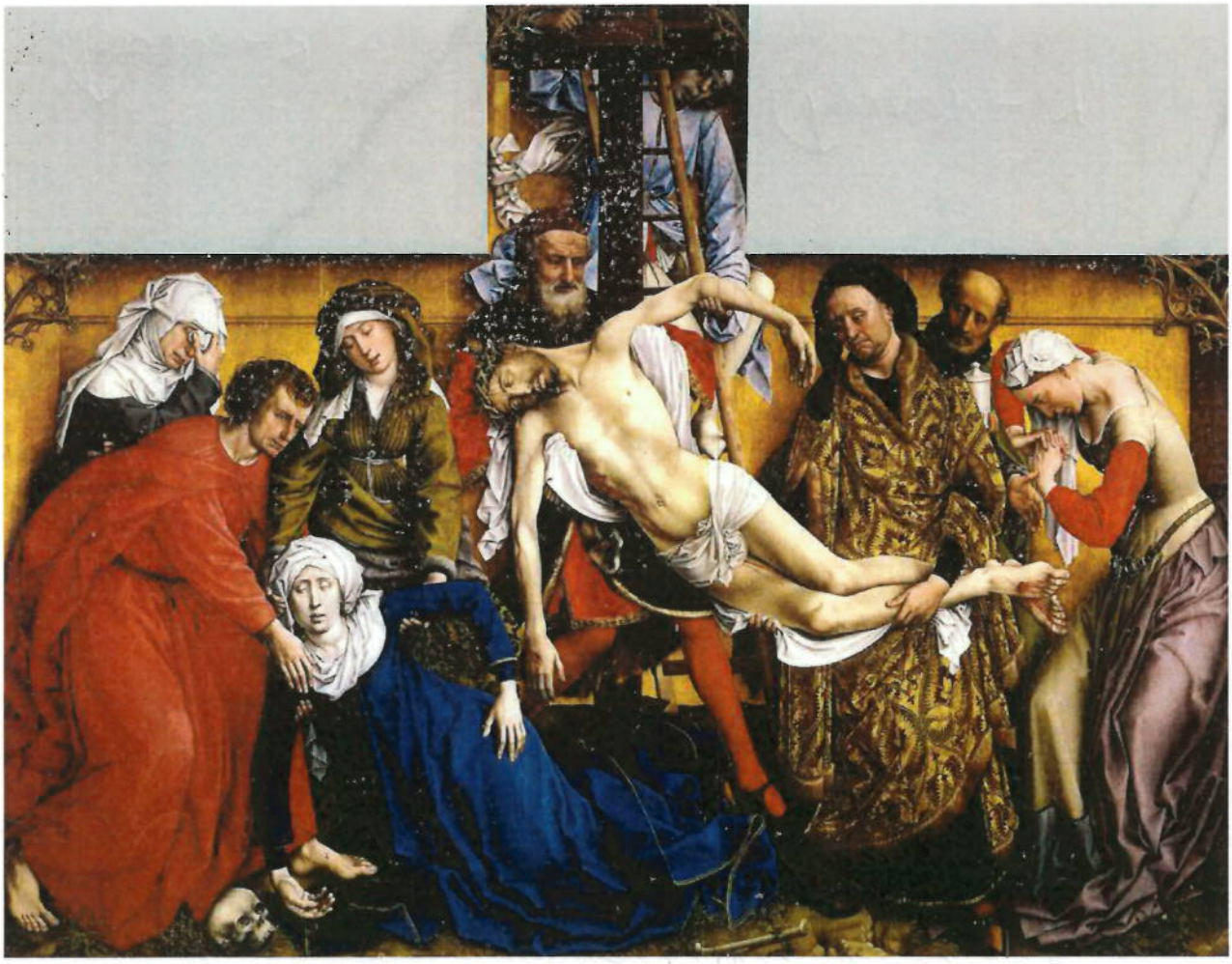


Image 5.1: *Descent from the cross*, (1435 – 38), Rogier van der Weyden (1399 – 1464), Museo del Prado, Madrid. (Source: upload.wikimedia.org)

5.1 Growth and decay

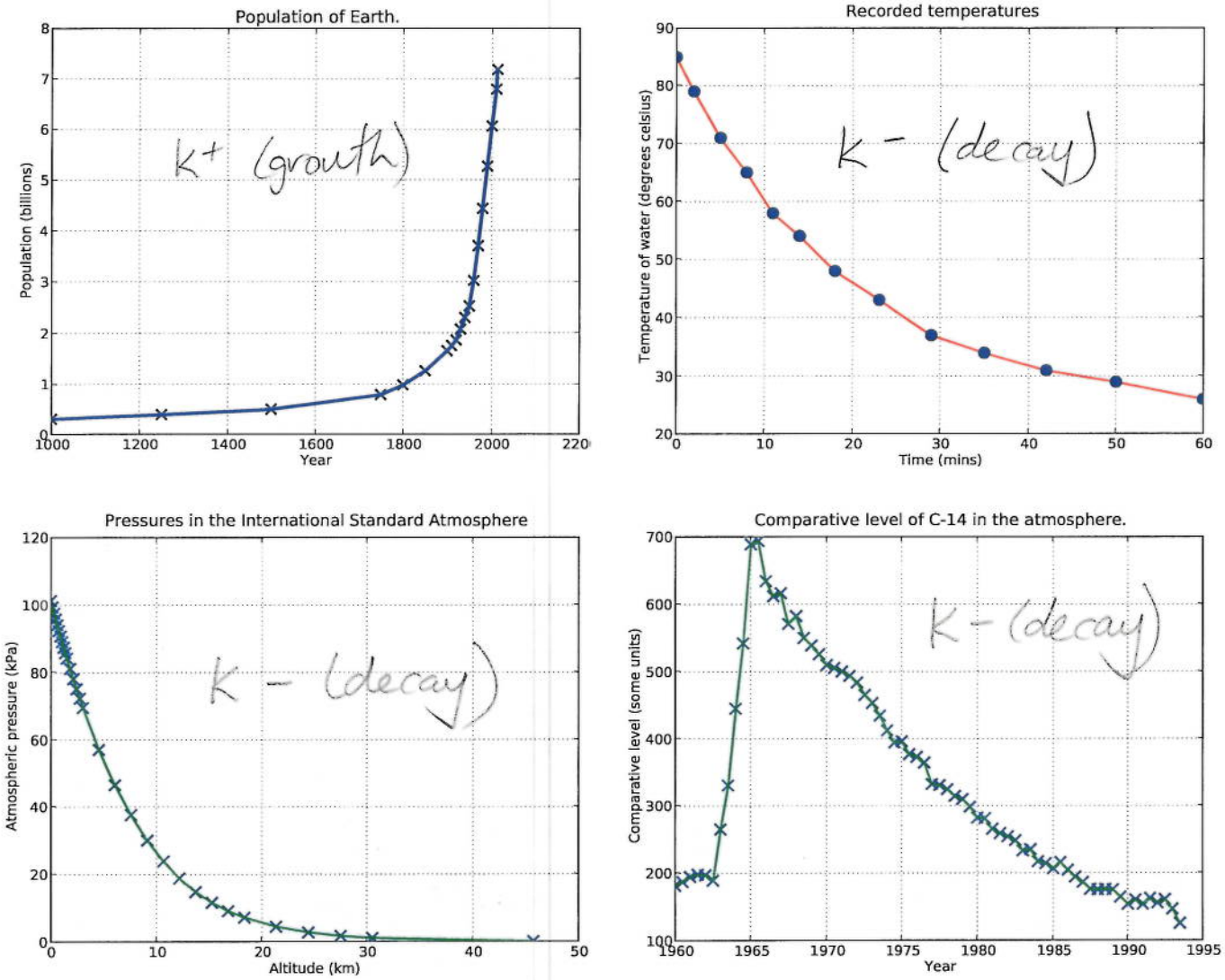


Figure 5.1: Some real, measured data. Top left: population of Earth over 1000 years. Top right: measured water temperatures in a simple experiment. Bottom left: atmospheric pressures in the international standard atmosphere. Bottom right: comparative level of atmospheric radioactive Carbon-14.

Question 5.1.1

Why did the comparative level of atmospheric radioactive Carbon-14 increase rapidly between 1960 and 1965, and why has it decreased since then?

Atomic bomb tests in 1960s.
Exponential decay after that (but C-14 ½ life = 5,700 yrs
Why? Absorbed by biological processes (plants, soils, ocean)

- Science primarily studies phenomena that change. Often, the rate of change at any time is proportional to the amount that is currently there.
- This is typical of many populations. For example, each year the size of the global human population is increasing by around 1.5% of its current size.
- Any phenomenon that has a rate of change proportional to the current amount follows an *exponential* function. (We will see why later.)

Exponential functions

Exponential functions have equations

$$f(x) = Ca^{kx},$$

where C, a and k are constants. The constant a is called the **base**. The two most common values used for the base a are

- the number 10; and
- *Euler's number*, denoted e, where $e \approx 2.71828\ldots$.

The constant k is the **growth rate** or **decay rate**. If C is positive then:

- If k is *positive*, the function displays exponential *growth*.
- If k is *negative*, the function displays exponential *decay*.

Note that when $x = 0$ the function value equals C.

Doubling time/Half-life

The **doubling time** for an exponentially growing quantity is the time it takes to increase to twice its original size.

The **halving time** or **half-life** for an exponentially decreasing quantity is the time it takes to decrease to half its original size.

Many exponential phenomena in science have relatively constant doubling times or half-lives over extended periods; knowing these values provides useful information about the phenomena.

***Example* 5.1.2**

Exponential functions occur frequently in models of nature and the social sciences. Some examples include unconstrained and constrained population growth, radioactive decay and carbon dating, modelling drug concentrations in blood, and modelling *habituation* to a stimulus.

- *Logarithms* (or *logs*) are very closely related to exponential functions.

Logarithmic functions

Logarithmic functions are of the form $f(x) = \log_a x$, verbalised as "f of x equals the **logarithm** of x to the **base** a".

In the special case that the base is Euler's number e, then the logarithmic function is often written as $f(x) = \ln x$, verbalised as "f of x equals the **natural logarithm** of x".

Logarithms and exponentials

The relationship between exponentials and logarithms is:

- If $y = 10^x$ then $x = \log_{10} y$ (and vice-versa).
- If $y = e^x$ then $x = \ln y$ (and vice-versa).

***Question* 5.1.3**

(a) Find $\log_{10} 1000$ and $\log_{10} 0.01$.

$\log_{10} 1000 = \log_{10} 10^3 = 3$

$\log_{10} 0.01 = \log_{10} 10^{-2} = -2$

(b) If $y = e^{0.02t}$, find $\ln y$.

$y = e^{0.02t}$

$\ln y = \ln e^{0.02t} = 0.02t$

- However, it *is* known that in any given time period a certain *proportion* of the total quantity in a sample will have decayed.
- Thus, radioactive material undergoes continuous decay at a rate **proportional** to the **quantity** of material, so the decay is an exponential process.

Decay constant

For a radioactive element, the *decay constant* k reflects the rate of decay of the element, and is a property of the chemical element. The half-life can be calculated from the value of k, and vice-versa.

Example 5.2.1

Decay constants and half-lives vary greatly between radioactive elements. For example:

- Polonium-212 has a half-life of about 3×10^{-7} s.
- Uranium-236 has a half-life of about 4.5×10^{9} years.
- Carbon-14 has a half-life of about 5730 years.

Example 5.2.2

Carbon-14 (C-14, also known as *radiocarbon*) is used to determine the age of organic-based artifacts (up to around 60,000 years).

Cosmic rays striking nitrogen in the upper atmosphere produce C-14. It then reacts chemically with oxygen to form radioactive carbon dioxide which permeates living creatures in a fixed proportion, either directly (by absorption from the atmosphere), or indirectly (via food chains).

When an organism dies, it ceases to accumulate C-14, and the remaining amount undergoes net decay over time. *Carbon dating* is the process of measuring the residual level of C-14 in organic artifacts, and thus deducing their age.

5.2 Exponentials in action

Case Study 14: **Radioactive decay**

(*) Demo:
Glow sticks
+
Radiation
+
Geiger counter

Photo 5.1: The B-29 Superfortress bomber "Enola Gay", National Air and Space Museum, Virginia, USA. (Source: PA.)

- Not all atoms remain the same over time; some undergo *radioactive decay*, which involves rearrangement of the nucleus of the atom, sometimes changing it into a different element.
- When an element undergoes radioactive decay but remains the same element (maintaining the original number of protons), the new atom is called an *isotope*.
- One standard way of denoting isotopes is to write the name or chemical symbol of the element, hyphenated with its atomic mass. For example, Deuterium (an isotope of Hydrogen and the main ingredient in "Heavy water") is written as Hydrogen-2 or H-2.
- Radioactive isotopes have useful applications in a range of sciences and industries, including chemistry, biology, medicine, physics and engineering. Therefore, it is important to understand how to model their decay.
- Radioactive decay is spontaneous, so there is no way of knowing *when* a *specific* individual atom is going to undergo decay.

Question 5.2.3

The half-life of C-14 is 5730 years.

(a) Find the decay constant of C-14.

$$A = A_0 e^{kt}$$

$$t_{\frac{1}{2}} = 5730 \text{ years}$$

$$\frac{A_0}{2} = A_0 e^{5730k}$$

$$\frac{1}{2} = e^{5730k}$$

$$\Rightarrow \ln\left(\tfrac{1}{2}\right) = 5730k$$

$$\Rightarrow k = \frac{\ln\left(\frac{1}{2}\right)}{5730} = \frac{-0.693}{5730} = -1.2097 \times 10^{-4}/\text{yr}$$

Note: Exponential when ROC $\propto$ current value.

$$\frac{d}{dt}\left(A_0 e^{kt}\right) = kA_0 e^{kt}$$

so ROC $\propto$ current value.

Under what situations would this occur?

(continued over)

***Question 5.2.3** (continued)*

(b) Consider the following extract from the paper [9].

> "The Shroud of Turin, which many people believe was used to wrap Christ's body, bears detailed front and back images of a man who appears to have suffered whipping and crucifixion. It was first displayed at Lirey in France in the 1350s ... Very small samples from the Shroud of Turin have been dated by accelerator mass spectrometry in laboratories at Arizona, Oxford and Zurich. As Controls, three samples whose ages had been determined independently were also dated."

Researchers discovered that 91.9% of the 'expected' C-14 was present compared to that in new organic garments. Deduce the (approximate) age of the Shroud, and comment on your answer.

$A = A_0 e^{kt}$

so in the present

$0.919 A_0 = A_0 e^{kt}$

$0.919 = e^{kt} = e^{-1.2097 \times 10^{-4} t}$

$\Rightarrow \ln(0.919) = -1.2097 \times 10^{-4} t$

$\Rightarrow t = \dfrac{\ln(0.919)}{-1.2097 \times 10^{-4}} = 698$ years

≈ 700 years old

$= 2018 - 700 = 1318$ AD

End of Case Study 14: Radioactive decay.

Case Study 15: **Hot stuff, cold stuff**

Photo 5.2: Bush fire. (Source: DM.)

Photo 5.3: Glass blowing. (Source: PA.)

- Moving an object with one temperature to a location with a different (but constant) temperature leads to a gradual change in the temperature of the object to match that of the new location.

***Question* 5.2.4**

Explain why it is reasonable that an exponential function would model the temperature of an object moved to a location with a different temperature.

ROC = Rate of change of $T \propto \Delta T = k\Delta T$

$= k(T_{Environment} - T_{Object})$ ← current value

So ROC $\propto$ to current value $\therefore$ exponential

Newton's Law of Cooling

Photo 5.4: Ice castle, Harbin Ice Festival, China. (Source: PA.)

Question 5.2.5

In an experiment, Peter recorded the temperature of hot water in a container over one hour; see Figure 5.2. The room temperature was 25 °C.

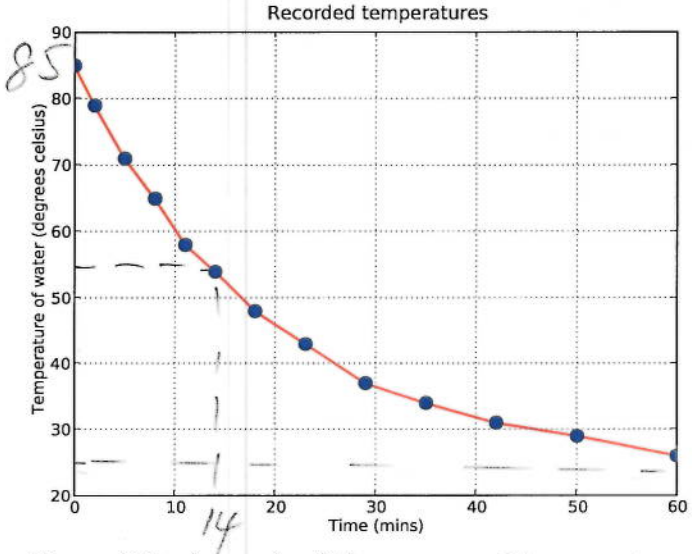


Figure 5.2: A graph of the measured temperatures.

Derive an equation for the water temperature at any time in minutes (note that the water approaches room temperature of 25 °C, not 0 °C.)

$T = T_0 e^{kt} + B$, B = vertical shift

$B = 25$

$T_0 = 85 - 25 = 60$

$\frac{1}{2}$-life when $T = \frac{T_0}{2} = 30$, so when $T = 55$

So $T = T_0 e^{kt} + B$

$55 = 60 e^{k14} + 25$ find k

$30 = 60 e^{k14}$

$\frac{1}{2} = e^{14k}$

$\Rightarrow 14k = \ln(\frac{1}{2}) \Rightarrow k = \frac{-0.693}{14}$

$\approx -0.049/\text{min}$

$\therefore T = 60e^{-0.049t} + 25$

We can develop a computer program to model the temperature.

Program specifications: Write a program that plots the measured water temperatures and the function that models these temperatures.

Program 5.1: Temperatures

```
# Program to plot measured and modelled temperatures.
from pylab import *

# Initialise variables
times = array([0,2,5,8,11,14,18,23,29,35,42,50,60])
temps = array([85,79,71,65,58,54,48,43,37,34,31,29,26])
model = 60 * exp(-0.05 * times) + 25
# Draw graphs
plot(times, temps, 'ro', markersize=10, label="Data")
plot(times, model, 'k-', linewidth=3, label="Model")
text(30,40,"model")
text(10,50,"actual")
xlabel("Time (mins)")
ylabel("Temperature of water (degrees celsius)")
title("Recorded temperatures")
grid(True)
legend()
show()
```

Output from the program is shown in Figure 5.3.

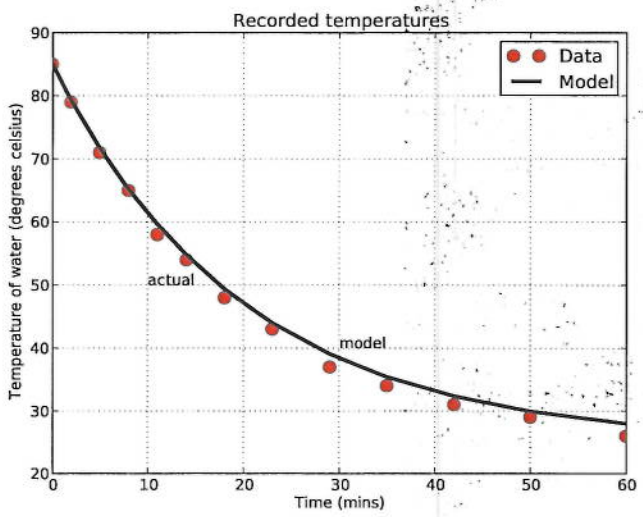


Figure 5.3: Modelled and actual water temperatures.

End of Case Study 15: Hot stuff, cold stuff.

5.3 Logarithms in action

- Logarithms provide a convenient mechanism for converting exponential data into a form that can make data analysis easier.

Question 5.3.1

Assume some data are modelled by the exponential function $D(t) = D_0 e^{kt}$. Demonstrate how a logarithmic transformation of the data values results in a linear model. Interpret the y-intercept and gradient of the linear model. (Hint: if x and y are positive then $\ln(xy) = \ln x + \ln y$.)

$D(t) = D_0 e^{kt}$

$\ln D(t) = \ln(D_0 e^{kt})$

$\ln D(t) = \ln D_0 + \ln e^{kt}$

$\ln D(t) = \ln D_0 + kt$

of form: $y = a + bx$ $\therefore$ y-intercept $= \ln D_0$

slope $= k =$ decay constant

Question 5.3.2

Earlier we saw that the *International Standard Atmosphere* (ISA) [28] models various atmospheric properties, including temperature, pressure and density. Figure 5.4 shows atmospheric pressures in kilopascals (kPA) at various altitudes in the ISA, and Figure 5.5 shows a graph of these pressure data transformed using natural logarithm, ln.

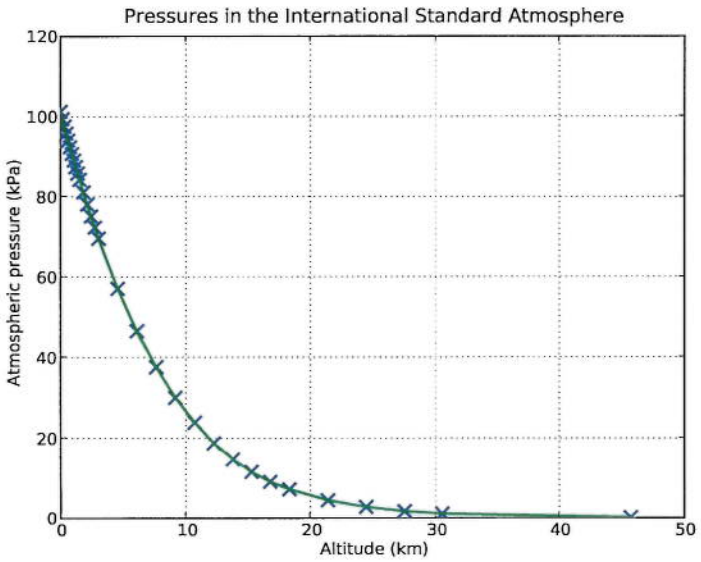


Figure 5.4: ISA pressures. *(continued over)*

Question 5.3.2 (continued)

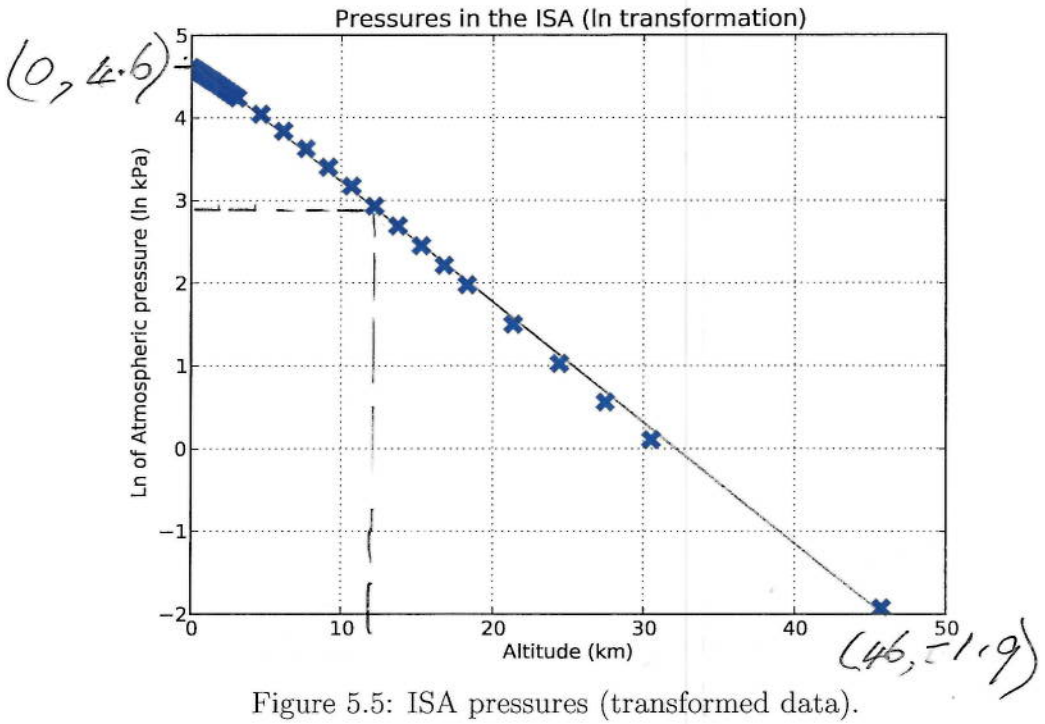


Figure 5.5: ISA pressures (transformed data).

(a) Use Figure 5.5 to estimate the pressure outside a jetliner cruising at 12000 m.

At 12 km altitude, AP = Air Pressure

$\ln AP = 2.9$

$\Rightarrow \quad AP = e^{2.9} = 18.17 \text{ kPa}$

(b) Use Figure 5.5 and Question 5.3.1 to find an exponential model of pressures in the ISA. alt = altitude

$AP(alt) = AP_0 e^{k \cdot alt}$

$\ln AP = \ln AP_0 + k \cdot alt$

Now $\ln AP_0 = 4.6 \Rightarrow AP_0 = e^{4.6} = 100$

& $k = \text{slope} = \frac{\text{rise}}{\text{run}} = \frac{-1.9 - 4.6}{46 - 0} = \frac{-6.5}{46} = -0.14$

So $AP(alt) = 100 e^{-0.14 \cdot alt}$

(continued over)

***Question** 5.3.2 (continued)*

(c) When a jetliner is in flight, the pressure in the cabin is artifically raised to a higher level than the pressure outside. The *cabin altitude* is the altitude at which atmospheric pressure matches the pressure inside the cabin. Modern planes typically cruise with a cabin altitude of about 2000 m. Express this cabin pressure as a percentage of the pressure at ground level in the ISA.

$AP(aH) = 100e^{-0.14 \cdot aH}$

$AP(2) = 100e^{-0.14 \times 2}$
$= 100e^{-0.28}$
$= 75.6$

$AP(0) = 100e^{-0.14 \cdot 0}$
$= 100e^{0}$
$= 100$

$\therefore \dfrac{AP(2)}{AP(0)} \times 100\% = \dfrac{75.6}{100} \times 100\% = 75.6\%$

(d) Use your model from Part (b) to estimate the difference between internal and external pressure on a jetliner cruising at an altitude of 12000 m.

external $AP(12) = 100e^{-0.14 \times 12}$
$= 100e^{-1.68}$
$= 18.6 \text{ kPa}.$

Internal $AP(2) = 75.6 \text{ kPa}.$

$\therefore \Delta = AP(2) - AP(12) = 75.6 - 18.6 = 57 \text{ kPa}$

Photo 5.5: Bang? (Source: PA.)

***Example* 5.3.3**

Some very well-known scientific measurement scales measure log to base 10 of particular quantities. These include: the *Decibel scale*, which measures the 'loudness' of sounds; the *Richter scale* and *moment magnitude scale*, which measure earthquake intensity; and the *pH scale*.

Case Study 16: **The pH scale**

Photo 5.6: Erosion due to acidic rain, Eyam Church, UK. (Source: PA.)

- An important application of logarithms in Chemistry is the pH scale, which is a measure of the *acidity* or *alkalinity* of solutions.
- The pH of a solution reflects its relative concentration of positive hydrogen ions $[H^+]$, in mol/L. It is defined to equal the *negative of the logarithm to base* 10 *of the concentration* of $[H^+]$ ions, so

$$pH = -\log_{10}[H^+].$$

- A pH of 7.00 represents a neutral solution, and **decreasing** pH values correspond to an **increase** in acidity. Most substances have pH between 0 (very acidic) and 14 (very alkaline).

***Question* 5.3.4**

Find the pH of gastric digestive juice in which $[H^+] \approx 10^{-2}$ mol/L.

$pH = -\log_{10}[H^+] = -\log_{10}[10^{-2}] = -2.$

Question 5.3.5

Rising atmospheric CO_2 levels threaten coral reef survival. Atmospheric CO_2 dissolves into the ocean and produces carbonic acid (H_2CO_3), leading to ocean acidification and damage to coral skeletons. Ice core samples suggest that the long-term seawater pH was about 8.25. Studies predict this pH could drop to 7.65 by the year 2100. If this is correct, find the relative concentration of hydrogen ions in sea water in 2100 compared to the long-term average.

$$[\text{Relative}] = \frac{[H^+]_{2100}}{[H^+]_{\text{long term}}} \qquad pH = -\log_{10}[H^+] \Rightarrow [H^+] = 10^{-pH}$$

$$= \frac{10^{-7.65}}{10^{-8.25}}$$

$$= 3.98 \approx 4 \text{ time more } [H^+]$$

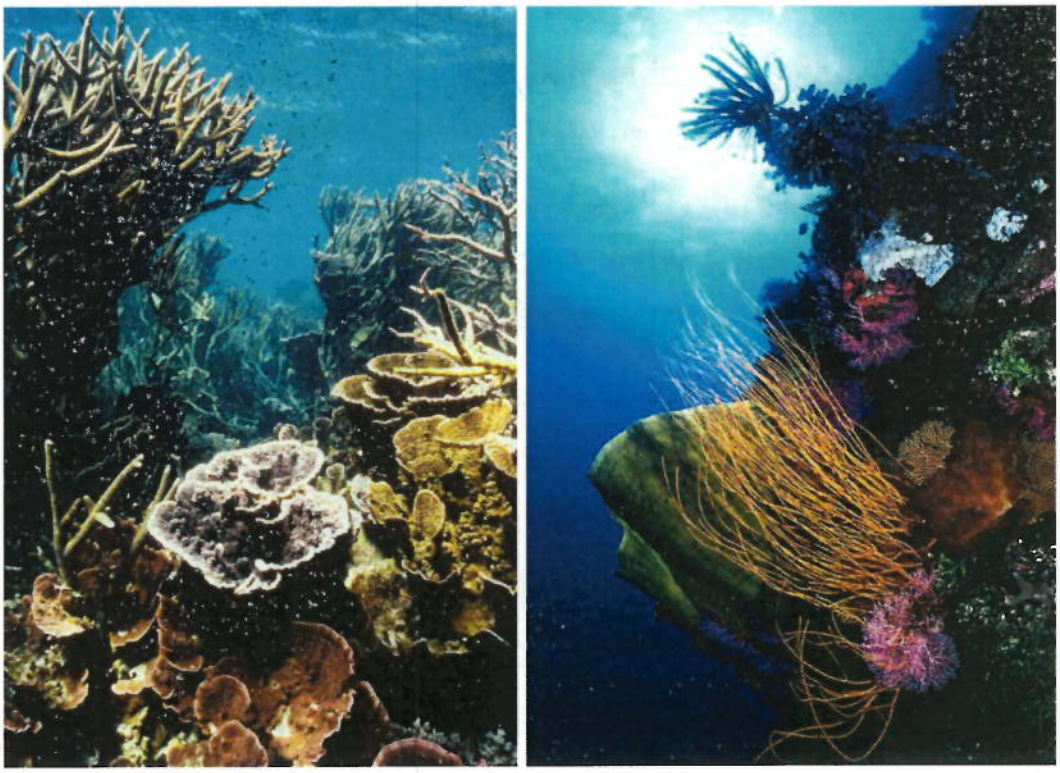


Photo 5.7: Coral reefs. (Source: DM.)

***Extension** 5.3.6 (from [27])*

"Increases in atmospheric $CO_2 > 500$ ppm will push carbonate-ion concentrations well below 200 μmol kg^{-1} ... and sea temperatures above +2 °C relative to today's values. These changes will reduce coral reef ecosystems to crumbling frameworks with few calcareous corals... Under these conditions, reefs will become rapidly eroding rubble banks such as those seen in some inshore regions of the Great Barrier Reef, where dense populations of corals have vanished over the past 50 to 100 years."

- Image 5.2 (used with permission from O. Hoegh-Guldberg, UQ) illustrates the predicted impact on coral reefs of various levels of atmospheric CO_2 and resultant ocean warming.
- The left image shows the current (comparatively) healthy condition of many reefs. The centre and right images show increasingly degraded reefs, consistent with rising levels of atmospheric CO_2 and resulting ocean warming.

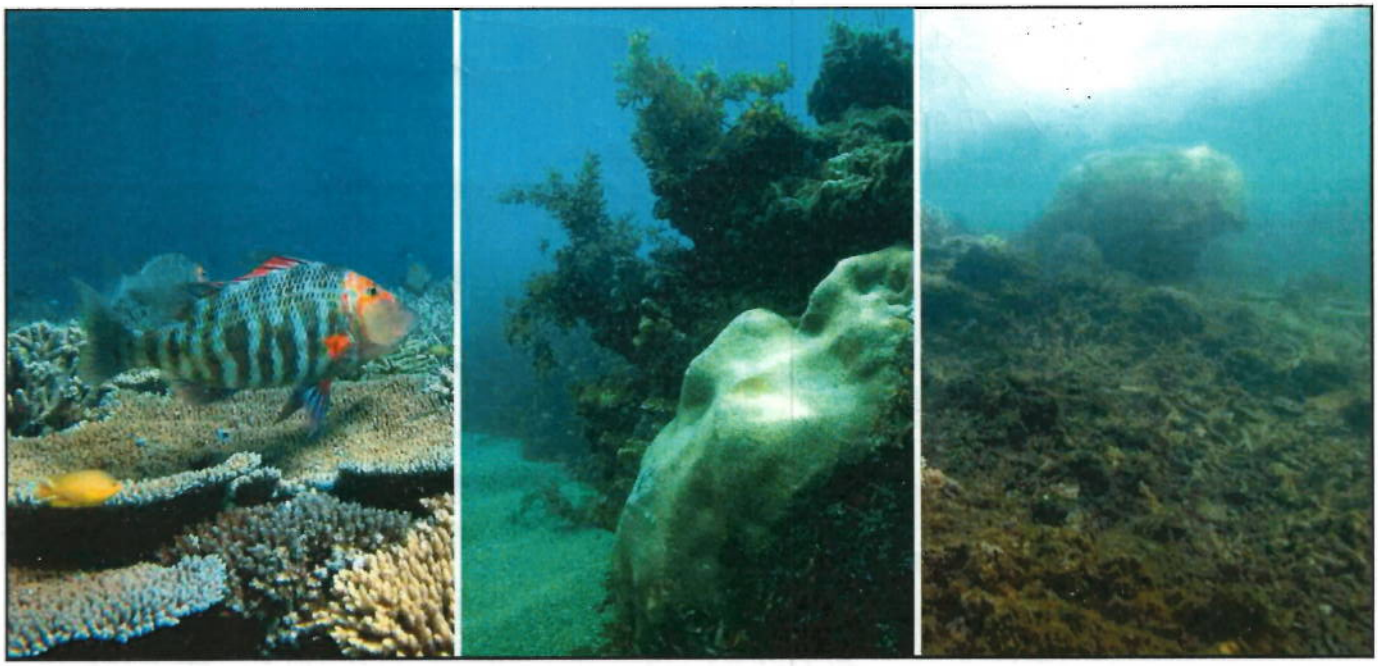


Image 5.2: Predicted impact on coral reefs of rising atmospheric CO_2 levels.

End of Case Study 16: The pH scale.

Question 5.3.7

Keeling Model 5: Figure 5.6 shows two plots: a graph of the function $y(t) = 280 + 35e^{0.022t}$, and the Keeling curve.

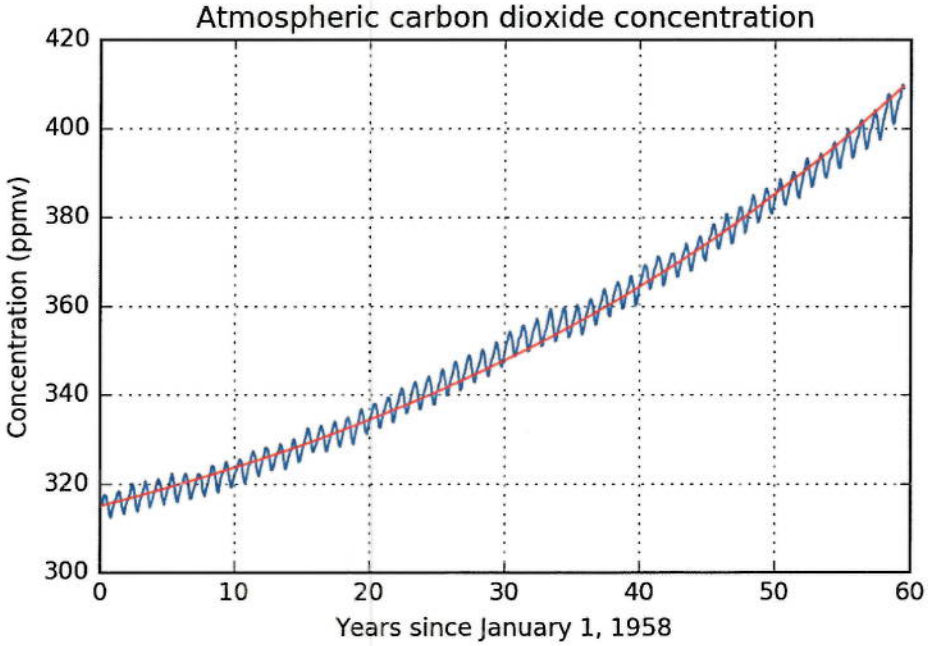


Figure 5.6: The Keeling curve and an exponential model.

(a) Explain <u>mathematically</u> how each term in $y(t)$ impacts on its graph.

$Y = 280 + 35\,e^{0.022t}$

280: vertical shift.

$35\,e^{0.022t}$: exponential growth curve.

35: scalar representing another vertical shift

(b) Data from ice-core samples show that long-term atmospheric CO_2 levels remained relatively constant at 280 ppm. Explain the <u>physical</u> significance of each term in $y(t)$.

When $t=0$ (i.e. 1958) the $Y = 280 + 35 \cdot e^{0.022 \times 0}$

$= 315$.

So by 1958 there had already been 35 ppmv of CO_2 added above long-term concentrations. There has been another 80 ppmv added since.

(c) How effectively does $y(t)$ model the underlying Keeling curve trend?

5.4 Keeling revisited

Question 5.4.1

Consider the following three models of the Keeling curve.

- Model Q+S: $y(t) = 0.014t^2 + 0.7t + 315 + 3.5\sin(2\pi(t - 0.15))$.
- Model P+S: $y(t) = 1/3t^{1.37} + 315 + 3.5\sin(2\pi(t - 0.15))$.
- Model E+S: $y(t) = 280 + 35e^{0.022t} + 3.5\sin(2\pi(t - 0.15))$.

Figure 5.7 plots graphs of the Keeling curve and all three models.

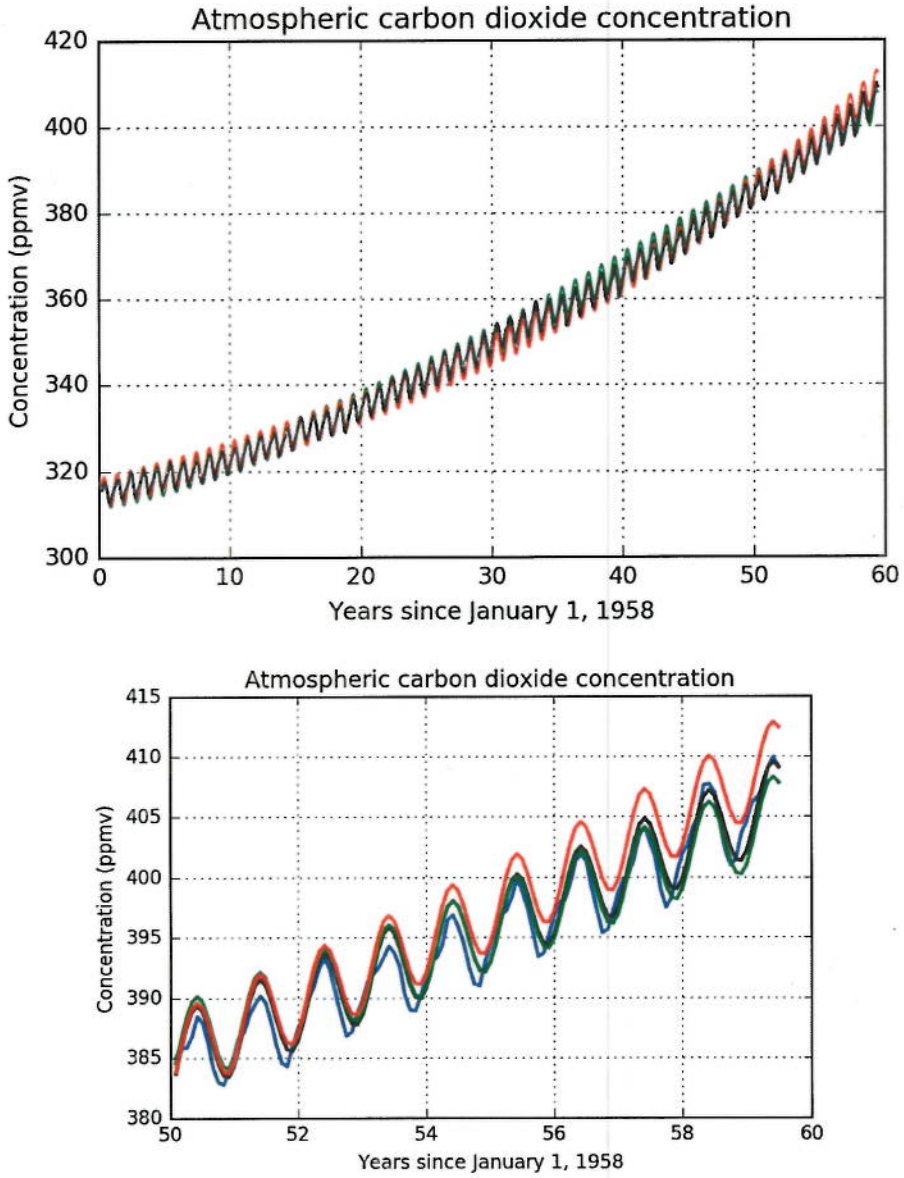


Figure 5.7: The Keeling curve and the three models for all years (top) and recent years (bottom).

(continued over)

***Question 5.4.1** (continued)*

(a) Which of the three models of the Keeling curve is correct? Why?

All ok for historical data, but who knows which is correct? None are mechanistic (ie. none have the actual mechanisms)

(b) Figure 5.8 extrapolates the models to the year 2058 (100 years after the Keeling study commenced). Which curve corresponds to each model?

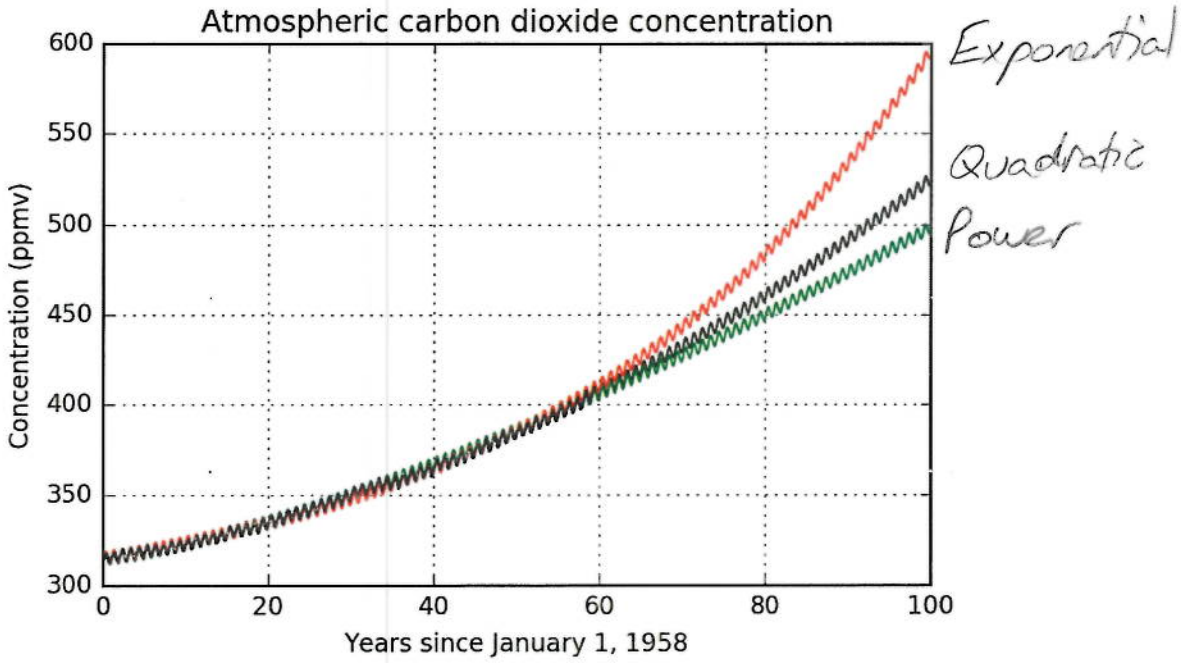


Figure 5.8: The three models of the Keeling curve, extrapolated to the year 2058.

(c) Discuss the ramifications of the different predictions.

Bleak predictions for the future

Extrapolation Exponential >> Quadratic > Power

(d) What do *you* think the real concentrations will actually look like? Why?

- Could be better or worse?
- Using oil & coal reserves
- Deforestation & ↑ animal agriculture (methane)
- Failure to act wrt mitigation

Part 3: Thinking

Tyger Tyger, burning bright,
In the forests of the night:
What immortal hand or eye,
Could frame thy fearful symmetry?

In what distant deeps or skies,
Burnt the fire of thine eyes?
On what wings dare he aspire?
What the hand dare seize the fire?

And what shoulder, and what art,
Could twist the sinews of thy heart?
And when thy heart began to beat,
What dread hand? and what dread feet?

What the hammer? what the chain,
In what furnace was thy brain?
What the anvil? what dread grasp,
Dare its deadly terrors clasp?

When the stars threw down their spears
And water'd heaven with their tears:
Did he smile his work to see?
Did he who made the Lamb make thee?

Tyger Tyger, burning bright,
In the forests of the night:
What immortal hand or eye,
Dare frame thy fearful symmetry?

The Tyger (1794), William Blake (1757 – 1827).

Image 5.3: *Truth, Time and History* (date unknown), Francisco de Goya (1746 – 1828), National Museum, Stockholm, Sweden. (Source: en.wikipedia.org).

Recall that there are five broad parts to this course. This is Part 3, "Thinking".

In Parts 1 and 2 of the notes, we considered the importance of modelling in science, and how we can use mathematics and computing to develop and apply models. In this part we take a short break from models, mathematics and computing, and instead take a "bigger picture" view.

In the first chapter of Part 3, we briefly explore the communication of scientific knowledge, particularly in the context of effectively communicating and interpreting quantitative data in a medical context.

In the second chapter, we explore the nature and logic of the so-called "scientific method". Since the birth of science, both scientists and philosophers of science have had different opinions as to exactly how and why science "works". We will discuss philosophical questions such as what distinguishes science from other human intellectual activities, and how can we rely upon science to produce 'knowledge'. We will also trace some historical attempts to answer these questions, and in the process, try to get our heads around the basic story.

You may be tempted to think that this material does not "directly fit" with the other course content we have been studying so far. However, an understanding of what constitutes science and scientific knowledge, and understanding the importance of effective scientific communication, are essential foundations for practising science.

Chapter 6: **Quantitative reasoning**

Our galaxy itself contains a hundred billion stars.
It's a hundred thousand light years side to side.
It bulges in the middle, sixteen thousand light years thick,
But out by us, it's just three thousand light years wide.
We're thirty thousand light years from galactic central point.
We go 'round every two hundred million years,
And our galaxy is only one of millions of billions
In this amazing and expanding universe.

The universe itself keeps on expanding and expanding
In all of the directions it can whizz
As fast as it can go, at the speed of light, you know,
Twelve million miles a minute, and that's the fastest speed there is.
So remember, when you're feeling very small and insecure,
How amazingly unlikely is your birth,
And pray that there's intelligent life somewhere up in space,
'Cause there's bugger all down here on Earth.

Artist: *Monty Python.* Song: *Universe song.*

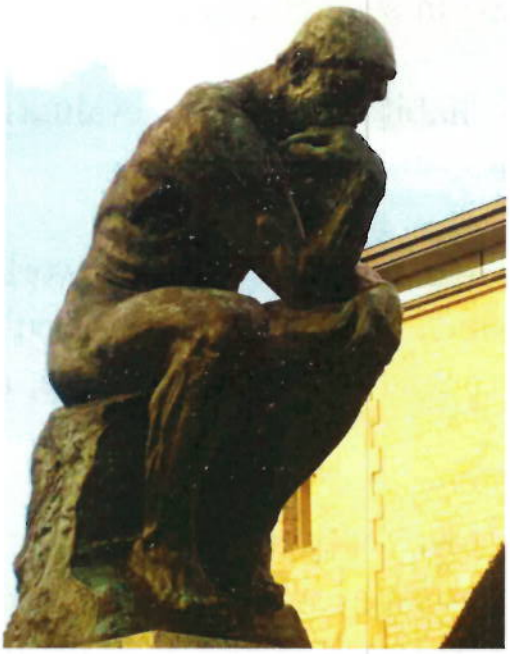


Image 6.1: *The Thinker* (1879 – 1888), Auguste Rodin (1840 – 1917), Musee Rodin, Paris. (Source: en.wikipedia.org)

6.1 Quantitative communication

- In SCIE1000, we will explore the fundamental skills and concepts that will help you with effective scientific analysis and communication.

- We are all producers and consumers of quantitative scientific information, in the form of scientific papers, assignments, media reports, the internet, and professional communications such as doctor/patient discussions.

- As a *producer* of such information, we should aspire to be accurate, honest, logical, unambiguous, concise, precise, not excessively technical, and always mindful of the intended audience.

- As a *consumer*, we should aspire to be thoughtful, reflective, sceptical, logical and analytical, while at the same time open-minded and accepting of evidence that may differ from our preconceptions or opinions.

- The media and internet provide a continual bombardment of facts, reports, summaries, interpretations and opinions, often covering sophisticated concepts but written and read by non-experts. In many cases there are errors (or deliberate falsities) in such communications.

- You should form the habit of critically evaluating information, data and (claimed) conclusions.

- A useful approach (when checking your own work, or the work of others), is *rough estimation*, which is the process of calculating approximate values. It involves building rough, conceptual models, and then evaluating them 'for sense'.

- Estimating 'gives an idea' whether a particular value is plausible. Often, we aim to find an approximate value within an *order of magnitude* of the correct value (that is, within a factor of 10 of the correct value).

6.2 Losing patients with mathematics?

- Sometimes, particularly in a medical context, critically evaluating quantitative information is a matter of life and death.
- A recent paper [21] presents the following key findings:
  - Many people (doctors, patients, journalists and politicians) do not understand health statistics.
  - Lack of understanding is due both to lack of knowledge, and intentional misrepresentation of information.
  - Sources of medical information (media, information leaflets and journals) tend to overstate benefits and understate risks.
  - Commercial and political manipulation undermines informed consent.
- The following paragraph is a quote from [21]:

  "Statistical literacy is a necessary precondition for an educated citizenship in a technological democracy. Understanding risks and asking critical questions can also shape the emotional climate in a society so that hopes and anxieties are no longer as easily manipulated from outside ..."

Question 6.2.1

In 1995, an emergency announcement in the UK warned that third-generation oral contraceptive pills doubled the risk of potentially life-threatening blood clots (thrombosis). The announcement led to widespread concern and fear, and many women ceased using the contraceptives. Reports estimate that in the following year there were an additional 13,000 abortions and 13,000 births, with 800 additional pregnancies in girls under 16 years of age. The announcement omitted the following relevant information:

- young women have an absolute risk of spontaneous thrombosis of 1 in 10,000.

(continued over)

***Question 6.2.1** (continued)*

- the absolute risk of thrombosis when taking second-generation oral contraceptive pills is about 1 in 7000.
- the relative risk of thrombosis increases by a factor of 4 to 8 during a Caesarean birth.
- the relative risk of thrombosis during and after pregnancy increases by a factor of around 4.
- the absolute risk of dying from thrombosis during or after an abortion is around 1.1 in 10,000.

(a) Discuss the difference between absolute risk and relative risk.

Absolute risk = probability of undesirable outcome

Relative risk = comparative, but relative to what?
= can be misleading when absolute risk is small

(b) The Australian Medical Association (AMA) website [2] states that:

"...in order to support and enhance the collaborative nature of the doctor-patient relationship, patients must be able to make informed choices regarding their health care. An informed choice is dependent on receiving reliable, balanced health information, free from the influence of commercial considerations, that is communicated in a manner easily understood by patients."

Discuss the contraceptive pill announcement in the context of the AMA quote.

- Irresponsible
- Much critical information ignored
- Effect of not taking pill worse than taking it.
=> consider alternative consequences
- Ill-informed choices

Question 6.2.2

Two commonly reported medical statistics are:

- the *5-year survival rate*, which is the percentage of people who are still alive five years after being diagnosed with a condition; and
- the *annual mortality rate*, which is the number of people dying from a given condition each year, often expressed as a rate per 100,000 people.

(a) The 5-year survival rate for prostate cancer in American men is 98%; for British men it is 71%.

(i) Assume that 1,000 British men and 1,000 American men receive a diagnosis of prostate cancer (at the same time). After 5 years how many men in each country are expected to have died?

US: $\frac{98}{100} \times 1000 = 980 = 20$ died

UK: $\frac{71}{100} \times 1000 = 710 = 290$ died.

(ii) Considering only the given data, which country has the 'better' health system, and why?

US >> UK

(b) The annual mortality rate for prostate cancer in American men is 26 deaths per 100,000; for British men it is 27 per 100,000. Considering only these data, which country has the 'better' health system, and why?

US ≡ UK

(continued over)

***Question* 6.2.2** *(continued)*

(c) The medical information given in Parts (a) and (b) is all correct. Explain how the (apparent) discrepancies could occur.

US - earlier diagnosis
- over diagnosis, too many false positives
=> more intervention
- cancer treatment (cut, burn, poison)

(d) Treatment for prostate cancer is invasive with many substantial side effects, including incontinence and impotence. Considering only prostate cancer, which country has the 'better' health system, and why?

UK > US?
≈ same mortality but fewer side effects.
=> better quality of life

Photo 6.1: Freedom square, Brno, Czech Republic. (Source: PA.)

Question 6.2.3

In [47], researchers asked 450 American adults (aged 35–70; 320 had attended college; 62 had a postgraduate degree) for answers to the following questions:

> "1. A person taking Drug A has a 1% chance of having an allergic reaction. If 1,000 people take Drug A, how many would you expect to have an allergic reaction?
>
> 2. A person taking Drug B has a 1 in 1,000 chance of an allergic reaction. What percent of people taking Drug B will have an allergic reaction?
>
> 3. Imagine that I flip a coin 1,000 times. What is your best guess about how many times the coin would come up heads in 1,000 flips?"

(a) What are the answers to the above three questions?

(1) $10 = 1\%$ of 1000

(2) $0.1\% = \frac{1}{1000} \times 100\%$

(3) $500 = \frac{1}{2} \times 1000$

(b) What proportion of respondents do you think gave correct answers to each of the questions?

(1) 70%

(2) 25%

(3) 76%

(c) What are the ramifications for doctors, journalists and politicians?

Scary!

poor advice

poor public information

poor decisions

Case Study 17: **Cancer**

- Cancer is the name for a large group of diseases affecting many different parts of the body. It arises from the uncontrolled, rapid growth of abnormal cells that interfere with the usual bodily functions.
- Cancerous cells can *metastasise*, spreading to other parts of the body.
- Common cancers include cancers of the lung, prostate (males), breast (mostly females), colon, skin, bladder, kidney and blood (leukaemia).
- Smoking and excessive alcohol consumption are major risk factors.
- Cancer is a leading cause of human mortality. Figure 6.1 lists all leading causes of death for Australians, by gender.

Cause of death	M	Cause of death	F
Ischaemic heart disease	10907	Ischaemic heart disease	9139
Lung/bronchus/trachea cancer	4882	Dementia/Alzheimer disease	6963
Cerebrovascular diseases	4245	Cerebrovascular diseases	6534
Chronic lower respiratory disease	3542	Lung/bronchus/trachea cancer	3255
Dementia/Alzheimer disease	3406	Chronic lower respiratory disease	3107
Prostate cancer	3079	Breast cancer	2795
Blood/lymph cancer	2327	Diseases of the urinary system	2053
Colon/rectal cancer	2240	Diabetes	2040
Diabetes	2199	Heart failure	1989
Intentional self-harm	1901	Colon/sigmoid/rectum/anus cancer	1811

Figure 6.1: Leading causes of death in Australians in 2012. (Source: Australian Bureau of Statistics.)

- Common cancer treatments include:

Poison – *Chemotherapy*, which involves the infusion of highly toxic chemicals into the body, killing rapidly dividing cells. (Recall that rapid division is a common characteristic of cancerous cells.)

Burn – *Radiation therapy*, which involves exposing cells to radiation and hence damaging their DNA, leading to cell death.

Cut – *Surgery*, which involves removing cancerous tissue from the individual.

– *Stem cell transplants* (or *bone marrow transplants*), which involves infusing healthy stem cells into an individual with cancer.

- All of these treatments can have minor to major side effects, including fatigue, nausea, mouth ulcers, hair loss, cognitive problems, infection, anaemia, infertility, graft-versus-host disease, burns, cancer (!) or death.
- Determining the precise treatment regime and dosages involves a trade-off between the beneficial impact of reducing tumour size and the (often severe or life-threatening) side-effects resulting from the treatment.

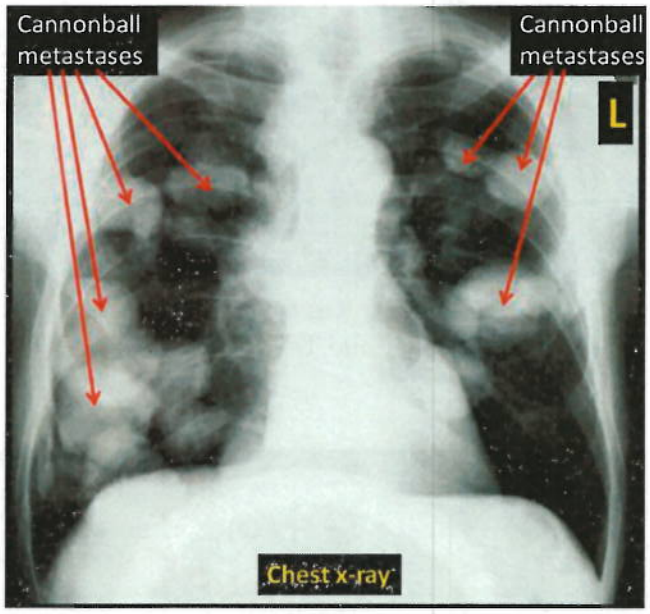


Photo 6.2: Chest X-ray displaying many classic "cannonball" metastases. (Source: Qld Health and DM.)

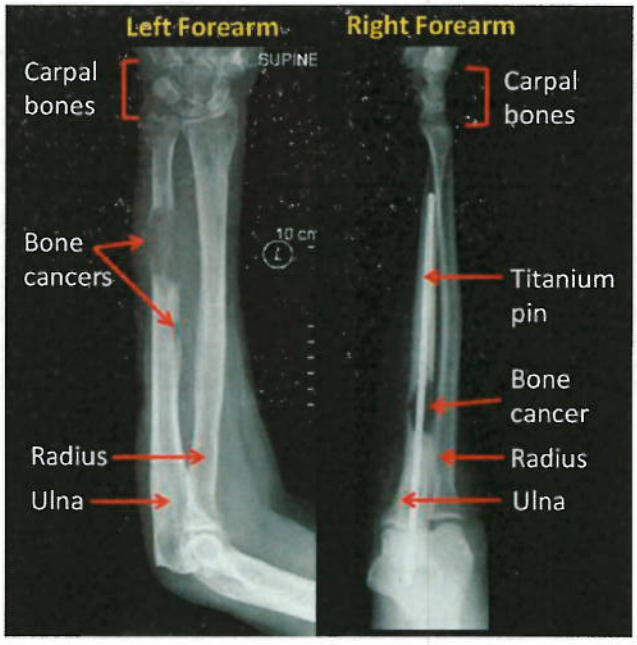


Photo 6.3: Left: x-ray of left forearm showing the destruction of bone due to cancer in the ulna. Right: x-ray of right forearm (different patient) showing a titanium pin stabilising the pathological fracture through the weakened bone. (Source: Qld Health and DM.)

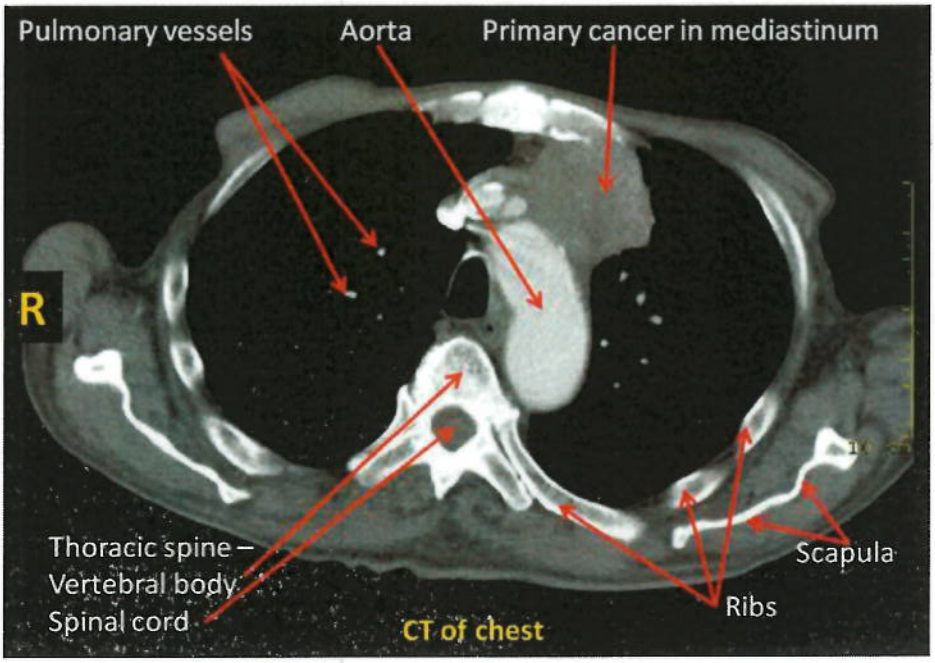


Photo 6.4: An axial CT image of the chest illustrates a primary cancer within the mediastinum extending to the anterior chest wall. (Source: Qld Health and DM.)

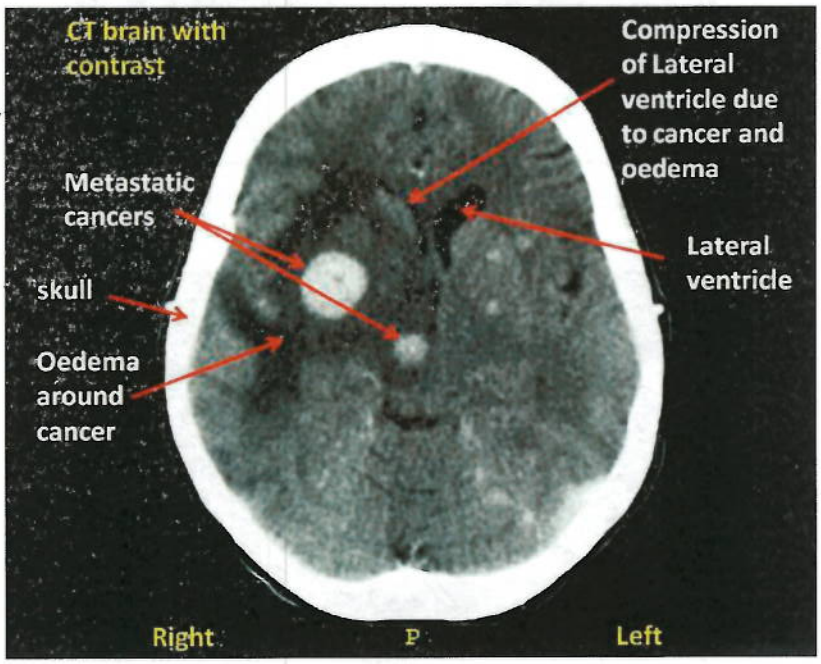


Photo 6.5: Axial CT image with contrast shows enhancing metastatic cancers with associated vasogenic oedema (swelling) in the brain. The metastases are due to breast cancer. (Source: Qld Health and DM.)

- *Breast cancer* develops due to the uncontrolled growth of cells in breast tissue, which enlarge into one or more lumps within the breast. It is a comparatively common cancer, and is a leading cause of death in women; it also affects men, but at a much lower rate.

Example 6.2.4

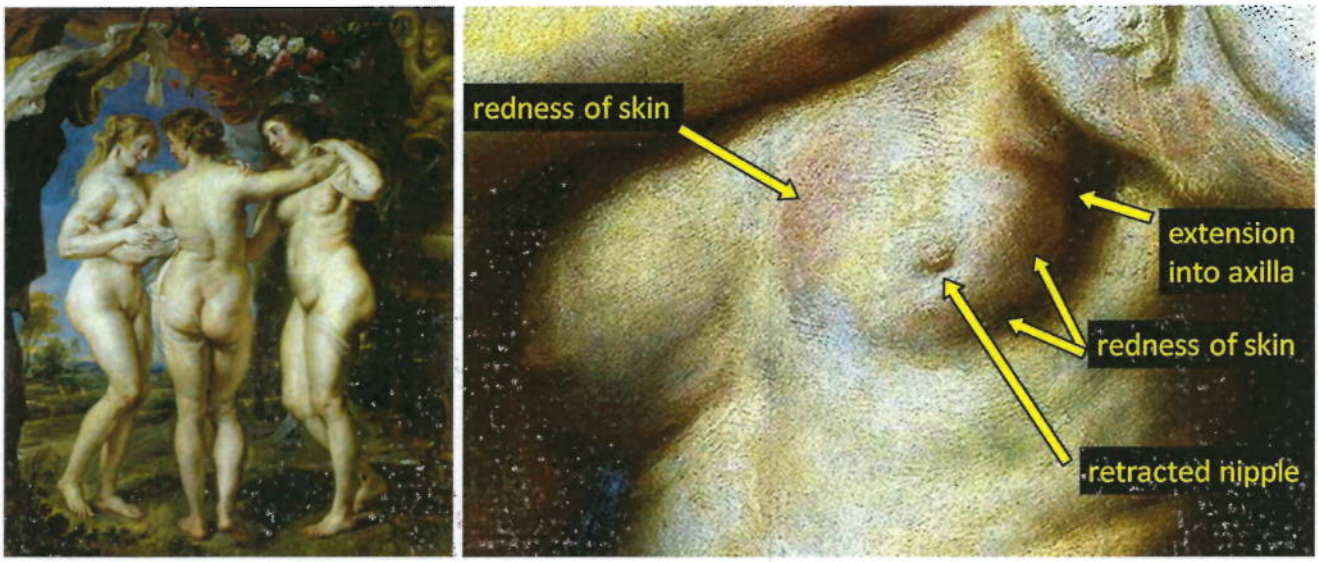


Image 6.2: *The Three Graces* (c1635), Peter Paul Rubens (1577 – 1640), Museo del Prado, Madrid. Left: full painting. Right: detail from the left breast of the right-hand Grace. (Source: upload.wikimedia.org)

A paper [24] states that the right-hand Grace has "a tumor in its external upper quadrant of the left breast which extends up to the left axila", with a "retraction in the left nipple", "the total volume of the left breast seems to be smaller than the contra-lateral one", and a "reddness of rounding skin suggesting inflamatory component" [sic]. The authors conclude that "this is a visual aspect of a locally advanced breast cancer".

- Some risk factors for breast cancer identified in the paper [35] include:
  - gender: the risk for females is around 100 times that for males;
  - age: the risk of developing breast cancer rises rapidly with age;
  - affluence: breast cancer is more common in affluent societies;
  - pre-existing breast conditions (for example, increased breast density);
  - hormonal factors (such as age at menopause or oral contraceptive use);
  - high levels of alcohol consumption.

- Some factors that reduce the risk of breast cancer include having children (more offspring at an earlier age reduces risk), breastfeeding, and increased physical activity.

Question 6.2.5

Figure 6.2 shows age-related probabilities of female breast cancer mortality (from [5]). The data are graphed in Figure 6.2, along with the function

$$d(t) = \frac{1}{43} \times \frac{1}{55^2} \times (t-30)^2 .$$

Age (yrs)	Prob.
30	1 in 19180
35	1 in 4600
40	1 in 1600
45	1 in 740
50	1 in 385
55	1 in 230
60	1 in 150
65	1 in 106
70	1 in 80
75	1 in 63
80	1 in 50
85	1 in 43

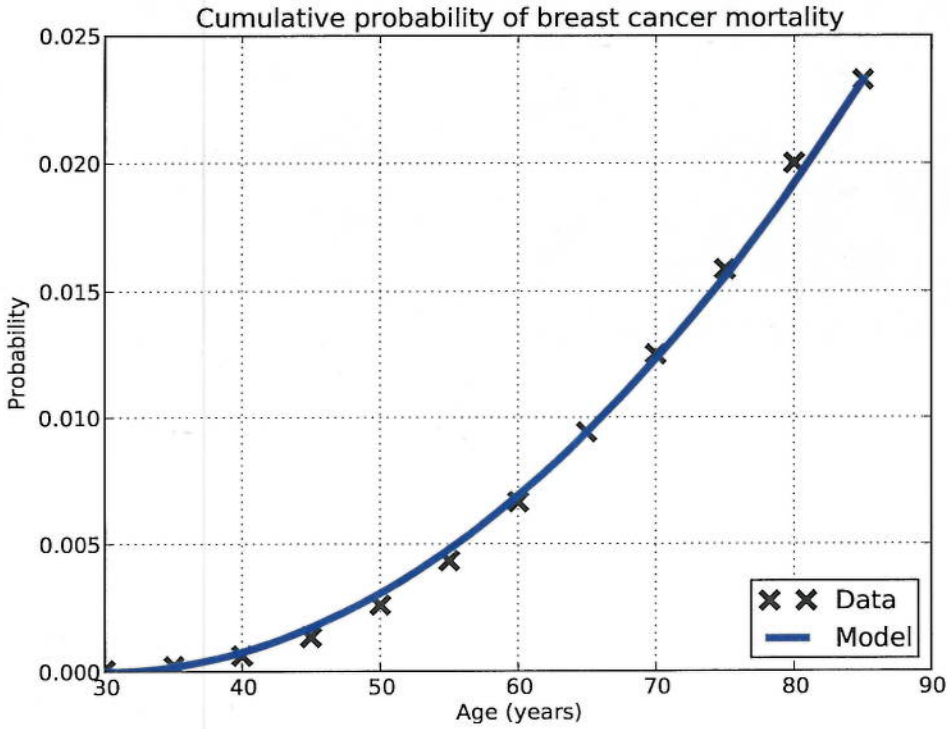


Figure 6.2: Cumulative probability of female breast cancer mortality prior to reaching various ages; see [5]. The graph also includes a plot of $d(t)$.

(a) Clearly, $d(t)$ models the probability of female breast cancer mortality prior to age t years. Explain the physical meaning of each term in $d(t)$.

$d(t) = \frac{1}{43} \times \frac{1}{55^2} \times (t-30)^2$

$= 0.023 =$ range in probability

$=$ scalar

$= (85-30)^2$

$=$ range ^{2} age.

Shift along x-axis to zero

Scale to 1 when $t = 85$

Scale to 0 when $t = 30$

Why quadratic?

Who knows why???

$d(t) = 0.000007t^2 - 0.00046t + 0.0069$

(continued over)

Question 6.2.5 *(continued)*

(b) Note that $d(t)$ is a quadratic function. What are the physical implications of this for individuals and public health organisations?

Escalation of death from breast cancer with age of t^2

=> screening for older ♀

Question 6.2.6

No medical test, for breast cancer or any other condition, is always accurate. Identify advantages and disadvantages of each of the following tests:

(a) *breast self-examination*: conducted by the individual, aiming to identify abnormal changes in her breasts

Advantages: Cheap, no side effects

Disadvantages: Relatively inaccurate.

(b) *clinical breast examination*: conducted by medical practitioners, feeling for lumps or other abnormalities

Advantages: Reasonably accurate, no side effects

Disadvantages: Expensive

(c) *mammography*: the use of X-rays to create breast tissue images, which radiologists visually inspect for abnormalities

Advantages: Accurate

Disadvantages: Expensive, radiation.

(d) *biopsy*: removal of a sample of breast tissue for laboratory analysis

Advantages: Very accurate.

Disadvantages: Very expensive, invasive

Image 6.3: Two images of mammographic procedures. (Source: www.cdc.gov, [5].)

Question 6.2.7

A *binary classification test* aims to classify objects, people or things into one of two groups. Examples include many medical tests, such as determining whether or not an individual has (or is likely to have) cancer. Most binary classification tests are imperfect: results can be *true positives*, *false positives*, *true negatives* or *false negatives*.

(a) Draw a table demonstrating true/false positives/negatives.

Test \ Disease	+	−
+	True +	False +
−	False −	True −

(b) What are some characteristics of a "good" binary classification test?

High True +ve, Low False −ve
High True −ve, Low False +ve.

(c) Identify some negative impacts of false positive or false negative cancer test results.

False +ve: Stress, expense, effects of treatment

False −ve: Premature death, effects of cancer

(d) When might a test with a higher false positive test rate be 'better' than one with a lower rate. Give an example in which it would be worse.

A higher false +ve rate might be "better" if treatment effects are minimal & cheap, & the disease is bad.
Higher false +ve would be worse if treatment invasive eg. prostate cancer

(e) Are false positive results 'better' or 'worse' than false negative results?

It all depends...

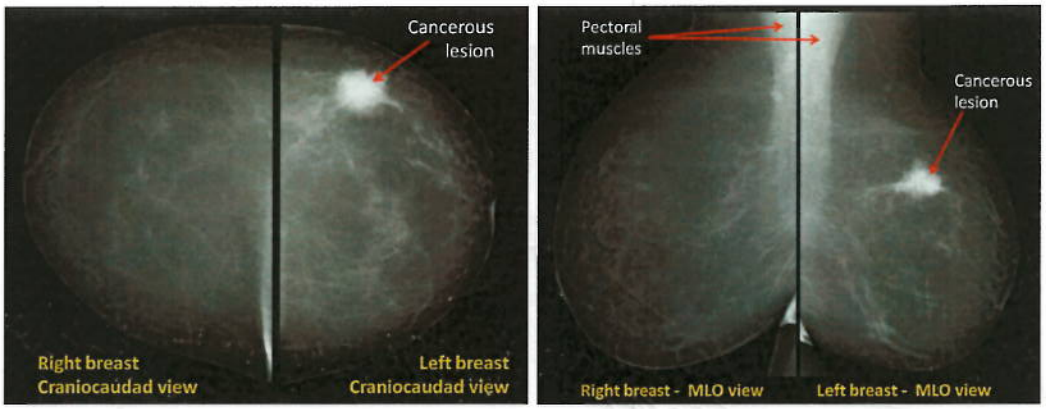


Photo 6.6: Mammographic x-ray images of both breasts in the craniocaudad (head/foot) view (left) and medio-lateral oblique (MLO) view (right). The left breast images show a dense cancerous lesion. (Source: Qld Health and DM.)

Question 6.2.8

A paper [21] quotes an example in which 160 gynaecologists were asked:

"Assume you conduct breast cancer screening using mammography... You know the following information about the women in this region:

- *The probability that a woman has breast cancer is 1% (prevalence)*
- *If a woman has breast cancer, the probability that she tests positive is 90% (sensitivity)* $= \frac{A}{A+C}$
- *If a woman does not have breast cancer, the probability that she nevertheless tests positive is 9% (false-positive rate)* = 9% specificity $= \frac{D}{C+D}$

A woman tests positive. She wants to know whether that means that she has breast cancer for sure, or what the chances are. What is the best answer?

A. The probability that she has breast cancer is about 81%.

B. Out of 10 women who test positive, about 9 have breast cancer.

C. Out of 10 women who test positive, about 1 has breast cancer.

D. The probability that she has breast cancer is about 1%."

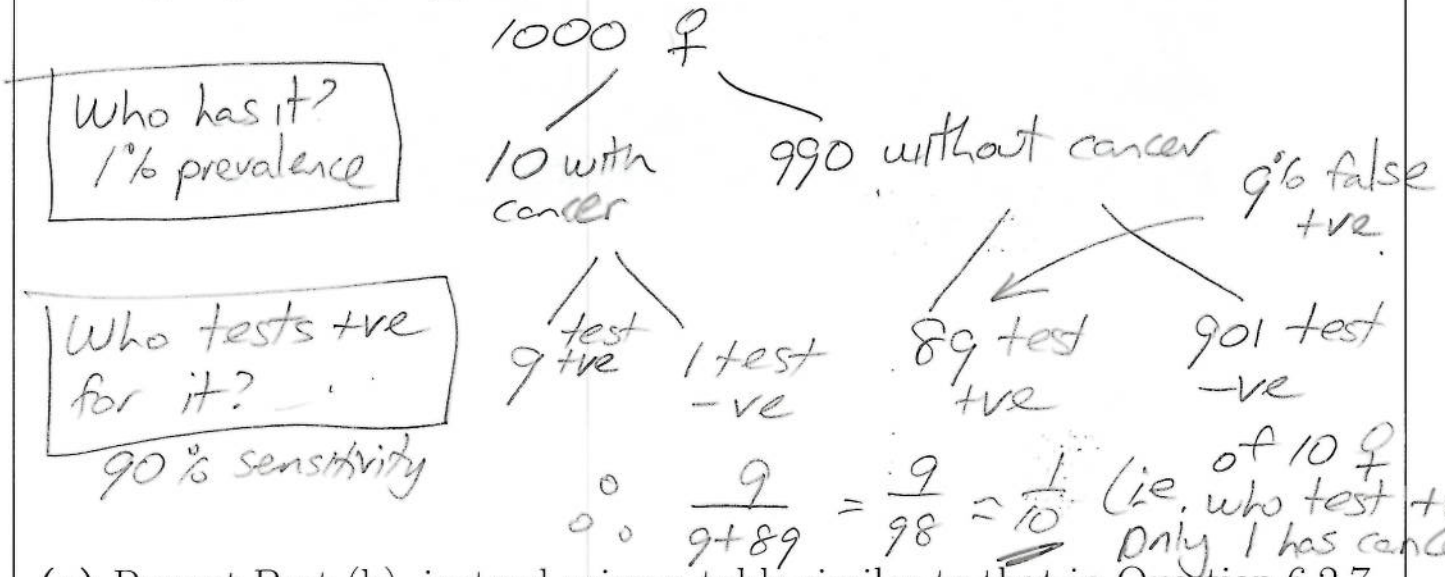
(a) Without doing detailed calculations, what is **your** answer?

???

(continued over)

***Question 6.2.8** (continued)*

(b) Investigate the answer to the question using a *probability flowchart* and a group of 1000 'typical' women.



$\therefore \frac{9}{9+89} = \frac{9}{98} \simeq \frac{1}{10}$ (i.e. of 10 ♀ who test +ve only 1 has cancer)

(c) Repeat Part (b), instead using a table similar to that in Question 6.2.7.

Prevalence = 1% = A+C.
False +ve rate = 9% = $\frac{B}{B+D}$
Sensitivity = 90% = $\frac{A}{A+C}$

Test	Disease +	Disease −	
+	A: 9 (True +)	B: 89 (False +)	
−	C: 1 (False −)	D: 901 (True −)	
	10	990	1000

So testing positive $= \frac{9}{9+89} = \frac{9}{98} \simeq 10\%$

(d) Estimate the proportion of gynaecologists who answered Part (a) correctly. What are the implications for you and/or your female relatives?

Answer: 2% (less than random guess)

Poor medical advice,

False +ves – stress, surgery, cost, side effects.

(*) Present 6.2.9 immediately before p.139

Question 6.2.9

Using a table similar to that in Question 6.2.7, define *accuracy*, *sensitivity* and *specificity*, and explain why each is important.

Test \ Disease	+	−
+	True + : A	False + : B
−	False − : C	True − : D

Accuracy $= \frac{A+D}{N}$ = probability of correct results

Sensitivity $= \frac{A}{A+C}$ = probability of getting it right if you have disease

$= 1 - p(\text{false negative})$

Specificity $= \frac{D}{B+D}$ = probability of getting it right if you don't have disease

$= 1 - p(\text{false positive})$

Question 6.2.10

A paper [15] studied the effectiveness of combined mammography and ultrasound imaging to screen for breast cancer. A total of 203 women returned "suspicious or malignant" test results, of whom 138 were later found to have cancer (via biopsy testing). A total of 3014 women returned "normal or probably benign" test results, of whom 150 were later found to have cancer. Find the accuracy, sensitivity and specificity of the combined procedures.

(continued over)

Question 6.2.10 *(continued)*

Test \ Disease	+	−	
+	A: 138	B: 65	203
−	C: 150	D: 2864	3014

$$\text{Accuracy} = \frac{A+D}{N} = \frac{138+2864}{203+3014} = 0.933$$

$$\text{Specificity} = \frac{D}{B+D} = \frac{2864}{65+2864} = 0.978$$

$$\text{Sensitivity} = \frac{A}{A+C} = \frac{138}{138+150} = 0.479$$

- Treatment options for breast cancer include chemotherapy, radiation therapy, hormonal methods and surgery, including total removal of the breast (mastectomy) and breast-conserving surgery (lumpectomy).
- Photo 6.7 shows an x-ray of a breast specimen removed during a lumpectomy.
- Wires inserted preoperatively during a mammographic biopsy procedure guide the surgeon to the cancerous lesion during the operation.
- Intra-operative x-rays of the excised breast specimen help to determine whether removal of the cancerous lesion was successful.

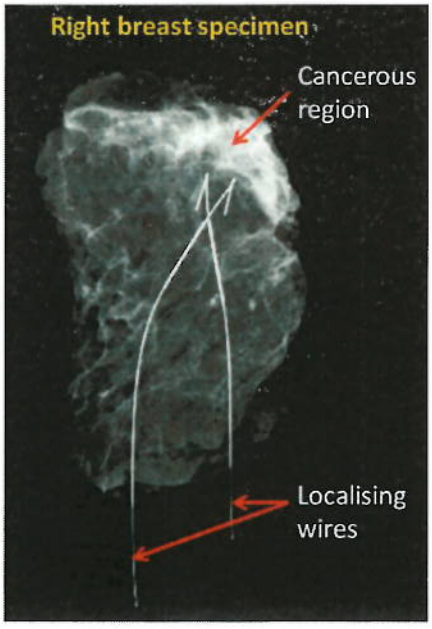


Photo 6.7: X-ray of a breast specimen containing localising hook wires. (Source: Qld Health and DM.)

End of Case Study 17: Cancer.

(*) Play Grim Reaper add from YouTube

Question 6.2.11

In the 1980s, blood screening in Florida found that 22 people who had donated blood tested positive for AIDS. Once notified of the test results, seven of these donors committed suicide. (At that time, AIDs was largely unheard of, and people were not regularly tested. Screening donors for the disease commenced after the discovery that transmission of AIDS occurred through contact with infected blood.)

The AIDS test has a very high *sensitivity* [percentage of infected individuals who correctly test positive] of about 99.9% and *specificity* [percentage of non-infected individuals who correctly test negative] of about 99.99%.

The *prevalence*, or rate of infection, for heterosexual men with low-risk behaviour, is around 1 in 10,000.

(a) What is the (approximate) probability that someone who tests positive for AIDS is infected?

10,000 people.

Who has it?

Prevalence 1 in 10,000 → 1 with AIDS, 9999 without AIDS

Who tests for it?

1 with AIDS → Sensitivity = 99.9% → 1 test +, 0 test −

9999 without AIDS → Specificity = 99.99% → 1 test +, 9998 test −

∴ 50:50 chance of having AIDS if you test positive!

(b) Assume that seven low risk, heterosexual men all test positive for AIDS. Calculate the probability that at least one of them does not have AIDS.

P(1 or more people don't have AIDS) = 1 − P(all have AIDS)

$= 1 - (0.5)^7 = 1 - 0.008 = 0.992$

∴ 99.2% chance that one of the ♂ that committed suicide don't have AIDS!

Question 6.2.12

(From [21].) To investigate the quality of AIDS counselling for heterosexual men with low-risk behaviour, an undercover client visited 20 public health centres in Germany, undergoing 20 HIV tests.

The client was explicit about belonging to a low risk group, as do the majority of people who take HIV tests. In the mandatory pre-test counselling session, the client asked: 'Could I possibly test positive if I do not have the virus? And if so, how often does this happen?'

The answers from the medical practitioners were:

No, certainly not	False positives never happen
Absolutely impossible	With absolute certainty, no
With absolute certainty, no	With absolute certainty, no
No, absolutely not	Definitely not ... extremely rare
Never	Absolutely not ... 99.7% specificity
Absolutely impossible	Absolutely not ... 99.9% specificity
Absolutely impossible	More than 99% specificity
With absolute certainty, no	More than 99.9% specificity
The test is absolutely certain	99.9% specificity
No, only in France, not here	Don't worry, trust me

(a) How would **you** answer the question?

None were exactly right (ie specificity = 99·99%.
Most were very wrong!

(b) Comment on the responses from the German doctors, relating your answer to your answers to Question 6.2.11 and the AMA statement in Question 6.2.1.

Poor statistical knowledge of doctors ⇒ poor information to patients ⇒ poor choices by patients

6.3 Huh?

Question 6.3.1

Critically evaluate and comment on the following quotes and examples.

(a) (*www.mentalhealth.org.uk*)

"The equation [by Cliff Arnall] calculates that Monday 19 January 2009 is the worst day of the year... Blue Monday was devised using the following mathematical equation:

$$\frac{(W + (D - d)) \times T^Q}{M \times N_a}.$$

The model was broken down using six immediately identifiable factors; weather (W), debt (d), time since Christmas (T), time since failing our new year's resolutions (Q), low motivational levels (M) and the feeling of a need to take action (N_a)."

(b) Equations for "Blue Monday":

www.ft.com, en.wikipedia.org: $\dfrac{(W + (D - d)) \times T^Q}{M \times N_a}$

www.msnbc.msn.com: $\dfrac{[W + (D - d)] \times TQ}{M \times NA}$

www.scotsman.com: $([W + (D - d)] \times TQ)(M \times NA)$

www.peterboroughtoday.co.uk: $W + (D) \times T^Q / M \times Na$

news.bbc.co.uk, www.cbc.ca: $1/8W + (D - d)3/8 \times TQM \times NA$

(continued over)

***Question** 6.3.1 (continued)*

(c) (*www.naturalnews.com, April* 16, 2008) "**Odds of intensive care medication errors are over one hundred percent**

A report produced by PubMed Central states that 1.7 errors per day are experienced by patients in intensive care units (ICU). At least one life-threatening error occurs at some point during virtually every ICU stay. 78% of the serious medical errors are in medications. 1.7 errors per day times 78% equals the likelihood of experiencing a medication error while in an ICU of well over 100% per day. That means the odds are that you will receive the wrong medication or the wrong amount of a medication at least once every single day of an ICU stay."

$1.7 \times 0.78 > 100\%$

Makes no sense => a likelihood $>100\%$

There are some excellent online articles that cover doubtful claims and errors in media reporting of quantitative science. The following links are very interesting, amusing, frightening and informative:

- "Behind the Headlines" [36]– provides a factual analysis of health-related claims, including the scientific background. **This is an excellent resource,** not for SCIE1000 but for your life. If you want to interpret reported health claims, this site will assist you to do so.
- "Bad Science" [22]– identifies and discusses mathematical and scientific errors in reported and published science.
- "Dodgy Boffins" [14]– discusses misuse of equations in the British media.
- "Helping Doctors and Patients Make Sense of Health Statistics" [21]– discusses the causes and impacts of errors in presenting and interpreting health statistics.

Chapter 7: **How scientific reasoning works**

Immanuel Kant was a real pissant
Who was very rarely stable.
Heidegger, Heidegger was a boozy beggar
Who could think you under the table.
David Hume could out-consume
Wilhelm Freidrich Hegel,
And Wittgenstein was a beery swine
Who was just as schloshed as Schlegel.

Artist: *Monty Python.* Song: *Philosopher's song.*

Image 7.1: *The Philosopher in Meditation* (1632), Rembrandt van Rijn (1606 – 1669), Musee du Louvre, Paris. (Source: en.wikipedia.org.)

7.1 Introduction: science and the assumption of rationality

News Headlines:

"No Rational Person Can Deny Human-Induced Global Warming"

What does this mean? What, if anything, is scientific "rationality"?

Many people believe that science is rational and that this is because there is such a thing as *the scientific method* by means of which we are able to make reliable claims about the natural world. Results arrived at by means of the scientific method have a special status — *scientific knowledge is reliable knowledge*, unlike claims made on the basis of common sense. This is not to say that common sense never leads to truth or science always does, but rather that our grounds for accepting claims made in the name of science are stronger than those for accepting claims made on the basis of common sense — this by virtue of the method used for arriving at the claims made.

To illustrate the point, consider the conflict between the common sense view that the sun revolves around the earth (the earth feels stationary, the sun seems to move across the sky, etc.) and the scientific view that, in fact, the earth revolves around the sun. This conflict is typically resolved in favour of the scientific view. In general, where science and common sense conflict, common sense gives way.

So consider two public speakers A and B. A stands and proclaims that human-induced global warming is not occurring, B claims it is. When asked, A admits that he claims no global warming due to human activities is occurring because common sense suggests that any fluctuation is more likely to be part of a natural cycle of climate change. B, on the other hand, admits that she claims human-induced global warming is occurring because scientific evidence all points to the fact. Now, *regardless of who is in fact right, who do we have more reason to believe?* B would commonly be said to be the more credible of the two by virtue of the means employed for arriving at the claim — B's method is more reliable, more rational, than A's.

The general view underlying this resolution of the conflict seems to be, again, that scientific claims to knowledge have some kind of merit not shared by common sense claims. But if merit attaches to scientific knowledge then why? What makes science rational? The common answer, again, is that science employs a method which is rational and is the means by which scientific knowledge is arrived at. This is what I want to talk about for the next four lectures.

7.2 Getting Philosophical

Of course, at this point, in discussing the general nature of science, we are engaged in an activity other than science itself. No amount of scientific experimentation will tell us whether or how science is rational. Just think about it for a minute. Asking questions about what science is and how it works is not something that we can do in a lab. White coats, bunsen burners and experiments won't help. When we question our beliefs *about science* we step out of science itself to a more abstract level of discussion. We are engaged in *philosophical* debate and argument about the nature of science.

The following lectures are directed at introducing you to some of the philosophical issues that arise in attempting to explain the apparent rationality of science and scientific method.

7.3 Some Preliminaries

7.3.1 What is Science?

What is science? As we shall see, it is not a body of *facts*. Despite common views to the contrary, science is not in the business to putting forward proven facts. Scientific laws and theories are continually being overthrown in the face of problems, or anomalies, that the laws or theories fail to adequately account for – they are always *provisional* to some degree, as are the specific scientific claims that depend on them. Newton's Laws, for example, were never scientific facts. They were conjectures that were eventually overthrown by a "better" laws, those of relativity theory.

Science is a way or method of thinking: thinking **critically** about **the empirical world** using **evidence** to try to justify **hypotheses**, **laws** and **theories** (collections of laws about some set of phenomena), put forward as **conjectures** that are subject to further critical testing against ever-increasing bodies of evidence.

Key Point: Science is not a body of facts. It is a way of thinking critically about the empirical world using evidence to make general conjectures.

Just think about the hypothesis of human-induced global warming. We appeal to evidence about humans increasing carbon dioxide levels in the atmosphere and their effect on the heat-retaining capacity of the atmosphere (the so-called "green house effect"), and go on to conjecture that *we* are thus the cause of a warmer climate. This conjecture is further tested by increasing bodies of evidence about past climate variation and its possible causes. And so it goes.

This conjecture is, of course, an hypothesis in the applied field of climate science. As applied science, it depends on a large number of even more fundamental physical and chemical laws, as well as associated mathematical principles that enable modelling of exponential growth of gas concentrations and summative effects of gas concentrations as a result of chemical reactions, etc. And these currently accepted laws and mathematical principles themselves are taken as (currently) justified. So how do we justify these? What justifies *these* fundamental laws and principles as acceptable?

For obvious reasons, we shall limit our discussion to *scientific* theory and set aside the many (interesting) issues surrounding the acceptability of mathematics used in science. The philosophy of mathematics is yet another area of philosophical enquiry with a heritage stretching back to the ancient Greek philosopher Pythagoras and beyond. The notion of mathematical truth, in particular, has been the subject of considerable study — most notably in the late 19th and early 20th centuries with key players like Bertrand Russell, David Hilbert and Kurt Gődel. Our focus, though, is squarely on scientific hypotheses, laws and theories.

7.3.2 Hypotheses, Laws and Theories

Science is made up of many hypotheses and laws, and groups of them that work together to make up scientific theories — evolutionary theory, electromagnetic theory, and so on. Let's just stop for a minute to get clear on our terms here. As potential scientists you will come across the terms 'hypothesis', 'law' and 'theory' a lot, and the way they are used in science is sometimes different to how they get used "on the street", so some clarification may help here.

— **An Hypothesis:** a (scientifically testable) claim used to predict or explain some particular phenomenon or event.

  E.g. hypothesising that, since a ball began to move, it was acted on by some force. The hypothesis is that the ball was acted on by some force.

— **A Law:** a (scientifically testable) claim describing *a general regularity* in nature used to predict or explain some particular phenomenon or event.

  E.g. All bodies remain at rest unless acted upon by some force. $E = mc^2$.

— **A Theory:** a set of interconnected laws and principles working together to form a *model* (typically involving significant idealisation) used to explain the general regularities themselves.

  E.g. Einstein's theory of general relativity (including $E = mc^2$, etc.) is a model that explains gravitational laws. Darwin's theory of evolution (including principles of natural selection, etc.) is a model that explains biological diversity.

Thus theories explain how the world works in general and why it works as it does by providing a model of the system in question. Hypotheses and laws are simply used to tell us why some particular thing happened but might themselves stand in need of explanation.

NB: Later in these lectures we'll pin down, more exactly, what we mean by "scientifically testable". For now, the above distinctions should be sufficiently clear.

Some more examples:

1. "Gravity is caused by undetectable particles exerting forces" may appear to be an hypothesis but is not (it is not testable — as we shall see later when discussing testability).
2. "The postman is sick" is an hypothesis (that one might offer to explain the lack of mail), but is not a law (it is not general).
3. "$F = ma$" is a law (describing a general regularity: when a force acts on an object it is caused to accelerate by an amount which when multiplied by the mass of the object, equals the force applied), but is not a theory. The relationship between the three quantity remains unexplained.
4. "Einstein's theory" explains general features of the observable world expressed by gravitational laws, etc.

Common Misuses and Abuses

Note that these key terms are not always used this way outside of science.

"Theory": People sometimes mean a mere *guess* or speculation lacking any support. (For example, "That is just a theory".) The former US President Ronald Reagan was famously reported as saying "Evolution is a theory, a scientific theory only . . . ". He meant it was just a guess, arguing that it was no more reasonable to believe than creationism (the view that the world was created by God rather than evolved).

"Law": People sometimes mean a general regularity in nature that has been *proven.* (For example, "But that is a law. It cannot be wrong.")

Scientists generally use the terms 'theory' and 'law' in a way that is neutral concerning whether they have no support, some support, or very strong support. More generally, the terms 'hypothesis', 'law' and 'theory' do not indicate a difference in how well established or proven a scientific claim is; they indicate the kind of claims in question — a specific claim that is general (law) or not, or an explanatory set of claims (theory). When we want to indicate that we are speaking of the *currently supported or accepted* view, we

often speak of *the* theory of evolution, or the law of gravity. (Here we indicate that one from among many candidates is accepted.)

7.3.3 The Task Ahead

Scientists typically believe that their way of thinking about the world involves some method, "the scientific method", and this rational method gives a way of justifying hypotheses, laws and theories (the results of scientific activity) as scientific knowledge. So what is this "method" and how does it produce scientific knowledge?

Notes:

7.4 Science and Inductive Reasoning

A popular view of scientific method is that science begins with particular observations of natural phenomena. From observation one logically arrives at general principles — scientific laws — by an inference known as *induction.*[1]

For example, imagine Boyle (1627-1691) studying the behaviour of a gas at constant temperature. He observes the following numerical measures of its volume at different pressures:

Pressure	Volume
1	12
2	6
3	4
4	3

From an examination of these few measures he infers the general law that the product of the pressure and volume is constant (given constant temperature) — Boyle's Law.

Justification of scientific laws is thus by way of a special kind of generalising argument — using "inductive reasoning" to infer the law-like structure of the universe. Our belief

[1] "Induction" here is not to be confused with the mathematical inference known as "mathematical induction".

in the laws of science is therefore rational since based on logical argument from evidence. ... Or so the story goes according to the inductivist.

In the same way, Avicenna (the 10-11th century Arabic philosopher, scientist, physician and mathematician) gave the following example of inductive inference, referring to the purgative effects of scammony[2] (not to be tried at home!):

	It is observed that ingestion of scammony is followed by
	the discharge of red bile
	This observation is repeated under circumstances in which
	other possible causes of the discharge of red bile have been excluded
Hence,	all scammony according to its nature withdraws red bile.

We seem to use this kind of inference all the time in daily life. Why expect that the fire will burn me? Because it has numerous times in the past. Why think that food will nourish me? Because it has numerous times in the past.

The main principle underlying this use of "inductive inference" is the idea that *what occurs frequently does not do so by chance.*

In this way, from observations of lunar eclipses and other phenomena concerning light one inductively infers general principles or laws — e.g. that light travels in straight lines; that opaque bodies cast shadows; and that certain configurations of opaque and luminous bodies place one opaque body in the shadow of another. (See the left side of Figure 7.1.)

These general principles or laws themselves can then serve as assumptions, along with some other facts concerning particular conditions etc., enabling one to *deduce* statements about phenomena. Continuing with the example above: the laws concerning opaque bodies and light, along with the fact that the earth and the moon are opaque bodies and the fact that on such-and-such a date the earth will pass across the line between the luminous sun and moon, can be used to deduce that there will be a lunar eclipse on that date. In this way we can make *predictions* about events not yet observed. (See the right side of Figure 7.1.)

Alternatively, by showing how an observation can be deduced from the laws one can progress from an already observed fact that the lunar surface has darkened to an understanding of *why* this took place. In this way, deducing the observation from laws about

[2] A twining plant having a stout taproot, *Convolvulus scammonia*, found in Syria, Asia Minor, Greece, etc. The dried milky juice was used as a medicine from ancient times.

the nature of light and opaque bodies, along with particular facts, we can provide an *explanation* of events already observed.

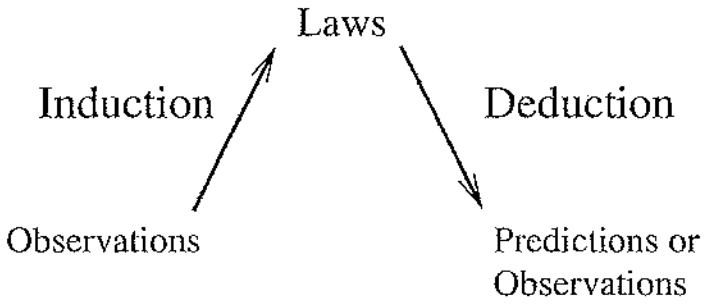


Figure 7.1: Induction and deduction.

On this account, "induction" is the crucial process used to arrive at laws. It allows us to infer some general law-like property or relation from a number of particular, observed cases or events. Our experience of some novel scientific phenomenon (for example, the effect of some drug on the human nervous system) is usually very undiscriminating; we do not initially see the general principles at work — we begin with a confused mass. However, over time, with sufficiently many repeated occurrences of the phenomenon, we are able to infer the general principles underlying the phenomenon — we reason our way to a universal law-like feature of the universe. By examining many cases we can inductively infer a formal pattern.

Key Point: Inductivist models of scientific thinking claim scientific laws are justified by *induction* from observations of a particular phenomenon.

Notes:

7.5 The Renaissance: experimentation and mathematics

Notice the use made of an actively pursued set of experimental results in the second premise of Avicenna's inductive inference. Later, in the 16th and early 17th century, the Renaissance development of instruments like the telescope, microscope and accurate clocks meant that much more sophisticated experiments could be undertaken (e.g. in the microscopic realm and the astronomical realm). And more extensive use was made of increasing amounts of experimental evidence that was becoming available. The emphasis on active experimentation to acquire new and relevant data for theorising became increasingly significant.

Another significant Renaissance development was the increasing use of mathematics and mathematical modelling in scientific theorising.[3] In the Renaissance the idea of a "clockwork universe" developed — a conception of the universe as like a giant mechanical clock (typically with God the great Watchmaker who set it all in motion) governed by laws amenable to precise mathematical modelling, and there was a shift from qualitative analysis to quantitative analysis and measurement. Scientists began to theorise about motion with idealised mathematical models invoking frictionless planes, perfectly spherical bodies, etc. This reached its greatest expression in Newton's Laws of Motion, whose mathematical simplicity and wide applicability confirmed a view of the cosmos as essentially mathematical in nature.

A century before this great Newtonian triumph, Galileo had put the point clearly:

> Philosophy [i.e. science] is written in that vast book which stands forever open before our eyes, I mean the universe; but it cannot be read until we have learnt the language and become familiar with the characters in which it is written. It is written in the mathematical language, and the letters are triangles, circles and other geometrical figures, without which means it is humanly impossible to comprehend a single word.[4]

[3] It seems hard to imagine science without sophisticated mathematics involved but, remember, the calculus of infinitesimals — developed by Gottfried Leibnitz and Isaac Newton, and necessary for the modelling of motion, acceleration, population growth, etc. — and probability theory — developed by Blaise Pascal — were not developed until the 17th century.

[4] *Il Saggiatore (The Assayer).* Quoted in A.C. Crombie's *Grosseteste and Experimental Science*, Oxford (1953) p. 285.

Image 7.2: Galileo, looking a touch haggard. (Source: en.wikipedia.org)

The book of nature is written in the language of mathematics, and that is why you need to get some mathematical skills under your belt before getting anywhere in science. If you don't like the maths, blame the Renaissance![5]

Key Points: In the Renaissance, technological innovation led to more *active experimentation* and an emphasis on the pursuit of observational evidence. Renaissance science also emphasised the *mathematical* structure of the scientific universe.

QUESTION: Could we do science purely *qualitatively* — i.e. using notions like *strong*, *weak*, *hot*, *cold*, *near*, *far*, *fast*, *slow*, etc. — instead of the Renaissance push to quantitative measurement? If not, why not?

Notes:

[5]For some examples of this increasingly mathematical approach to science see: H. Kearney, *Science and Change 1500-1700*, Weidenfeld & Nicolson (1971), pp. 66-7; A.C. Crombie, *Science, Optics and Music in Medieval and Early Modern Thought*, Hambledon Press (1990), p. 325.

7.6 A Common View of Science

Consistent with the account of scientific method described in the last lecture, a popular view of science is described by Alan Chalmers as follows:

> Scientific knowledge is proven knowledge. Scientific theories are derived in some rigorous way from the facts of experience acquired by observation and experiment. Science is based on what we can see and hear and touch, etc. Personal opinion or preferences and speculative imaginings have no place in science. Science is objective. Scientific knowledge is reliable knowledge because it is objectively proven knowledge.[6]

The notion of "proof" referred to here will, for the inductivist, be proof by inductive inference. On this "inductivist" interpretation of the popular view then, scientific reasoning is inductive and scientific method yields reliable knowledge through the application of inductive reasoning from the "facts of experience".

There are problems here though:

(i) The idea that scientific knowledge is *proven* knowledge can mislead us to think of such knowledge as certain. It is not.

(ii) The idea that science is objective and that "personal opinion or preferences and speculative imaginings have no place in science" is misguided.

(iii) The reliability of induction as a form of proof is difficult to justify.

Let us turn firstly to the problems concerning (ii).

7.6.1 Discovery versus Justification

With a little thought it seems obvious that there is a distinction to be made, first clearly drawn by John Herschel in the 1830s, between the means by which scientific theories are *discovered* — the context of discovery — and the means by which they are to be *justified* — the context of justification. How we discover a theory — i.e. the means by which we come to have the theory in our minds — is one thing, whereas establishing a theory as rationally acceptable — i.e. justifying it — is another thing.

The common view described above seems to wrongly characterise science by ruling as illegitimate that highly imaginative and creative aspect of science whereby practitioners "cook up" theories for consideration and testing. Who ever came up with the idea that

[6] A. Chalmers, *What Is This Thing Called Science?*, (1976) p. 1.

there could be "dark matter" or "dark energy" and what in heaven's name were they on when they thought up the idea? Who cares? What is relevant is simply whether or not such an idea (indeed, whether any scientific law, hypothesis or theory) can be justified. For, irrespective of how someone came up with the idea, what matters from the point of view of understanding the universe is whether or not such an idea can be justified.

With this in mind, we do not need to give any account of the process whereby we *discover* scientific theories. We do not need to claim that the process is in any way rational or reliable, let alone suggest that it proceeds by way of induction. An account of scientific method as essentially inductive then need only be committed to the view that scientific laws are *justified* by induction. Personal opinion and preferences, and speculative imaginings play a key role in science. Science is, in this sense, clearly subjective. Imaginative theorising by individual subjects puts hypotheses on the table for consideration which might otherwise never have been considered (these hypotheses are not "objectively given" to us), and some of the greatest honours in science go to those who have used their subjective imagination in ways that are ingenious, and which have produced theories which have subsequently come to be seen as justified. But how are they justified? This, according to those advocating an inductive scientific method, is by way of inductive inference. And so to problem (i).

Key Point: The discovery of scientific hypotheses — a process entirely separate from justification — is a *creative* process that infuses science with subjectivity.

Notes:

7.6.2 Inductive Inference and Fallibilism

What *exactly* is the logic of inductive inference? Consistent with Avicenna's example, inductive inference, according to one of its most famous advocates, the philosopher J.S. Mill (1806-1873), consists in inferring from a finite number of observed instances of a phenomenon, that it occurs in *all* instances of a certain class that resemble the observed instances in certain ways.[7] For example, from the fact that ingestion of scammony is observed on a number of occasions to produce red bile we infer that its ingestion always produces red bile. Similarly, from the fact that John, Peter, etc. are all mortal we infer, by induction, that all humans are mortal. Let's look at this in more detail.

[7] J.S. Mill, *A System Of Logic*, Vol. I, p. 354.

Observational (singular) statements

The simple inductivist account claims that science starts with observation. The scientist, with normal, unimpaired senses records what she sees without prejudice. Impartial reports as to how the world is are justified by the use of the senses. Statements reporting these particular facts — often referred to as *observational statements* — serve as the basis for the derivation of scientific laws. Examples of some simple observation statements are:

> At midnight on Jan. 1 1975, Mars appeared at position x in the sky.
>
> Mrs Smith struck her husband.
>
> The water boiled at 100 degrees Celsius at sea level (at place p and time t).

Observing what is the case will establish such statements as true or false at a particular place and time. Such statements are *singular statements*; they describe a particular event or state of affairs at a particular place and time.

Universal (general) statements

Scientific statements however describe general patterns in nature. For example:

> Planets move in ellipses around their sun. (Astronomy)
>
> Animals in general have an inherent need for some kind of aggressive outlet. (Psychology)
>
> Water always boils at 100 degrees Celsius at sea level. (Physics)

These statements refer to *all* events of a particular kind at any place and time; they are not about particular events or states of affairs but suitably general. The laws and theories of science involve general statements of this kind; they are *universal statements.*

The problem then for those who think that science starts with particular observation statements is to explain how one can justifiably arrive at universal statements from particular ones. For example: just because water boiled when heated to 100^0C at sea level by person x_1 at place p_1 and time t_1, and by person x_2 at place p_2 and time t_2, and by person x_3 at place p_3 and time t_3, and ... and by person x_n at place p_n and time t_n, how can we thereby justifiably infer that the result holds in general, for all future times, places and persons? How can the general be justified on the basis of the particular? How can a set of observations about how the universe is now justify claims as to how the universe is in general?

Inductive reasoning

The inductivist reply is that under certain conditions we can generalise. So long as the following conditions are met, we may legitimately generalise to an appropriate universal statement:

1. The number of observation statements forming the basis of a generalisation is large.
2. The observations are repeated under a wide variety of conditions.
3. No accepted observation statement conflicts with the derived universal law.

Condition (1) helps rule out anomalous cases (e.g. a defective measuring instrument) and stops one jumping to conclusions prematurely. Condition (2) implies that it is not enough to increase our base of singular statements by simply repeating tests on the same subject under the same conditions. To rule out the possibility of the observed phenomenon being due to some hidden factor we ought to test for the phenomenon under as varied conditions as possible. For example, the claim 'All liquids contract when frozen' would seem a justified generalisation if water was not considered; testing liquids under a wide variety of conditions will include testing the liquid water which is unusual in that it expands when frozen. Now obviously if water is observed to expand when frozen then the universal law 'All liquids contract when frozen' is not justified. Hence condition (3) is necessary.

This kind of reasoning — from a finite list of singular statements to a universal statement, from some to all — is called *inductive reasoning* and the process of reasoning thus is called *induction.*[8] The simple Inductivist position can be summed up by saying: science is based on inductive inference.

> *Principle of Inductive Inference:*
> If a large number of As have been observed in the past, under a wide variety of conditions, to possess the property B without exception we can infer that all As have the property B.

Scientific knowledge is built up from and justified by a secure base of particular observation statements by induction.

Science then, on this account, is justified by its use of induction in inferring, from particular observations, general laws and theories; *scientific statements can be inductively justified by experience.* Scientific statements, based on observational and experimental evidence (i.e. the facts) are contrasted with statements of other kinds — those based on pure logic or mathematics, authority, tradition, prejudice, or any other foundation. Scientific statements are derived in a rigorous and objective manner from objective facts. Science is a body of such knowledge and scientific progress then is the piecemeal addition of laws and theories to that body of knowledge; the accumulation of facts, and new laws

[8]NB: there are other forms of inductive reasoning, like inferring from the fact that water has always boiled at 100^0C at sea level in the past that it will do so when I next boil it. The problems we go on to discuss apply equally to these other forms but for simplicity we shall concentrate on the simple form presented here.

and theories arrived at via induction form the ever-growing observational base. This cumulative conception of scientific knowledge is sometimes called 'the bucket theory'.

Fallibilism

So the notion of proof that the inductivist relies on is inductive "proof". But it is important to realise that inductive proof falls short of certainty. Call it "proof" if you want, but it would be wrong to think inductive inference is anything like mathematical proof. When we prove Pythagoras's Theorem, we justify it as true, and because of the nature of mathematical proof we then take it to be established with *certainty*. There is no "probably" about it. This is not the case with inductive proof. Inductive inference cannot establish laws as certain. At best, it makes them *highly likely*.

For example, if water has been observed, on numerous occasions, to boil at 100^0C at sea level without exception, at best that only makes it *likely* or *probable* that all water boils at 100^0C at sea level. It is always left open to further contradictory evidence (evidence that we must recognise may be "out there", for all we know). We must recognise that our scientific investigations yield results that are clearly *fallible*. No-one in their right mind these days would claim that a currently "proven" scientific law is established as certain and so beyond revision. Since we cannot assume we have all the relevant data, scientific laws are always to be considered open to revision. In fact, we are constantly revising our scientific understanding of the world on the basis of new evidence and this involves admitting that we didn't have things "quite right" previously — a polite way of saying we were, in fact, *wrong* in what we previously thought!

If you think about it, everything we claimed to know about the scientific structure of the world in the past has been shown to be wrong. Scientists were once confident, for example, that Newton's laws of motion were absolutely certain, but new scientific evidence found in the early twentieth century led to developments in physics that resulted in their rejection in favour of more general relativity theory. Of course, you may want to say that Newton's laws weren't "wrong", they were just too general — they are a correct account of motion at low speeds. But that is to admit that, as general laws of motion (what they were put forward as describing), they *were* wrong. They are not true. More recently, scientists claim to have new evidence that suggests that they cannot account for some 80 percent of matter in the universe! They now speak of "dark matter" and "dark energy" — the stuff they can't yet detect but suppose is there. What revisions of our scientific laws will this lead to? We'll see. The point is, we can never discount the possibility of new evidence forcing revisions of our scientific understanding.

Science is an activity of constant testing of what we think we know — our *fallible* claims to knowledge — and constant searching for new information that might further confirm or refute what we think we know.

Bearing this in mind, the corresponding Principle of Inductive Inference should reflect this. It now says (and perhaps this what what many had in mind all along when it comes to induction):

> *The Weakened Principle of Inductive Inference:*
> If a large number of As have been observed in the past, under a wide variety of conditions, to possess the property B without exception we can infer that all As probably have the property B.

The common-sense account of science mentioned earlier that sees science as providing objectively *proven* knowledge is mistaken if 'proof' is read as 'proved with certainty'. Inductive proof can, at best, establish laws as reasonable-given-the-evidence.

Accepting this fallibilist shift then, we may ask why we ought to accept the weakened principle? Can such a principle be justified? And so to problem (iii) identified earlier.

Key Points: Simple induction proceeds from the *singular* to the *general.* Legitimate inductive inference is subject to certain *conditions* (1 - 3 above). Unlike mathematical proof, scientific proof is *fallible* — induction only justifies scientific laws as *probable.*

Notes:

7.6.3 The Problem of Induction

There are two ways we might try to justify the principle of induction: (a) we might try to justify it mathematically or logically — i.e. *a priori*; or (b) we might try to justify it scientifically, by experiment — i.e. *a posteriori.* It might not be obvious that these are the two kinds of defence one can offer so let me elaborate.

A priori justification

Consider the inferential principle "If no As are Bs and all Cs are Bs then no As are Cs". This can be justified on the basis of purely logical considerations; it is self-evidently true. Just find yourself a nice peaceful darkened room to think about it and pure thinking

alone, logic alone, should convince you of its truth. So too with "2 + 2 = 4". You don't need to know anything about how the world actually happens to be, no particular facts, to justify it. Such claims can be justified *a priori*.

Knowledge arrived at in this way, knowledge that can be established prior to knowledge of how the world happens to be, is called *a priori* knowledge. (It's stuff you can come to know just sitting in a dark cupboard with the door closed!) The most obvious examples of *a priori* knowledge include knowledge of mathematical principles or knowledge of logical principles. In this respect, mathematics (in particular) is quite distinct from science. Mathematics is an *a priori* discipline. Science is not *a priori* ··· you cannot develop scientific knowledge of the world isolated from any experience of how it happens to be, you must observe or otherwise actively engage with the world to see what, as a matter of fact, is the case.

A posteriori justification

Science is an *a posteriori* discipline. Pharmacology, for example, involves principles or claims that cannot be justified *a priori*. (If you are sitting in a dark cupboard, you'll need to open the door and have a look at the world outside to decide whether they are true!) Take a more ordinary example: "SCIE1000 students are the smartest students on campus". Pure reason, working alone prior to experience of how the world happens to be, cannot justify this claim. Sitting in a dark room with no idea of the outside world, one cannot prove it. We can only justify the claim given some knowledge of how the world is, after we have investigated the particular facts "on the ground", as it were. Such claims can only be justified *a posteriori*.

This kind of knowledge, knowledge that can only be established given some knowledge of how the world happens to be, is called *a posteriori* knowledge. Scientific knowledge is typically counted as the paradigm of *a posteriori* knowledge.

Here are some more examples of claims falling into the two categories of knowledge, justified in one of the two ways:

A PRIORI KNOWLEDGE	A POSTERIORI KNOWLEDGE
$2 + 2 = 4$	There are 2 houses of Parliament
All cats are cats	All cats land on their feet
Squares are rectangles	No celestial body has a square orbit
If it's not not 2016 then it's 2016	It is 2016

So what about the induction principle then? Does it fall under either heading?

The Principle of Induction? Provable *a priori* or *a posteriori*?

The Principle cannot be justified *a priori*. When we try to determine logically (*a priori*)

how probable a specific law is given the number of observations we have made it seems always to be low — near zero. This is because no matter how many observations I make, compared with the general claim that things will *always* be as the law describes, my number of observations seems to pale into insignificance next to the potentially infinite number of possible situations I am making claims about. My evidential support for some law or theory always appears insignificant compared with the full strength of my general claim. Not only could I always be wrong (i.e. science is fallible, as discussed), but it seems that I can never even show a law to be probable on purely *a priori* grounds!

Consider an analogy. Suppose you are out on a lake in heavy fog. Everywhere you can see in your vicinity is water. This is how things are in your bit of the world at your particular time and you are utterly ignorant of everything else. You have only ever lived in a boat on water, with the rest of the world obscured by fog. How confident could you be in inferring, by induction, that the universe in general, throughout space (i.e. everywhere) and time (i.e. always), is water? Not very. How things seem here and now is not a very reliable guide to how they are everywhere and always. Yet, the (weakened) inductivist is essentially in this position, it seems. From a finite number of observations over (at best) a few hundred years, in our little local part of the cosmos, how confident can we be that we can know with a high probability how things are in the cosmos in general — as when we claim to have confidence in a scientific law? Not very. Our justification is very weak, as it involves judging the nature of the cosmos in its potential infinity from evidence of a small, finite part.

(A lot of effort has gone into trying to develop a notion of 'probability' and an inductive logic that will enable us to logically, *a priori*, calculate or estimate the degree of support theories attain given the body of evidence in their favour but that story continues and I leave it to those interested to read further ...[9])

Ok, so perhaps we can justify the induction principle *a posteriori*. That is, if we can find out enough about how the world is then perhaps we can discover (*a posteriori*) that this general principle is, as a matter of fact, true. The problem here is that this would mean trying to scientifically justify the principle — i.e. justify it experimentally (instead of mathematically or logically). But the very thing we are trying to justify is inductive scientific proof! We can't use a scientific proof to prove anything until we have a story about what scientific proof is. We can't, in particular, use a scientific proof to show that our method of scientific proof (the principle of induction) is justified.

So, we have a problem justifying the inductivist's account of what science is.

Key Points: Knowledge divides into the *a priori* and the *a posteriori*. Induction seems impossible to justify under either category.

[9]See: W. Salmon, *The Foundations of Scientific Inference*, (1966).

Notes:

****

So there is a problem trying to justify the principle of induction. This is known as "The Problem of Induction" or "Hume's Problem" since it was the Scottish philosopher David Hume who brought the problem into sharp relief in his *Treatise On Human Nature* (1739).[10] The inductivist who claims scientific knowledge based on observation seems unable to defend their use of the principle of induction.

Image 7.3: David Hume. (Source: en.wikipedia.org)

[10]You might like to look at Alan Musgrave's very readable account of the problem of induction and some responses in his *Common Sense, Science and Scepticism*, Cambridge University Press (1993), esp. Ch. 8 & 9.

7.7 Popperian Science

7.7.1 The Hypothetico-Deductivist Account of Science — Falsificationism

The philosopher of science Karl Popper certainly thought that induction (and the presumption of the uniformity of nature) was unjustifiable. He argued that we shouldn't pin our hopes on induction as an account of how science works because we'd then be relying on an unacceptable principle of induction. Popper claimed that science didn't, in fact, need induction at all. We can explain how science works and why it is rational without any need to rely on anything as suspect as induction. His account of science and scientific method has been widely accepted by the science community and has become known as "the hypothetico-deductive method", sometimes also called "falsificationism". And, properly understood, it contrasts in a number of respects with the "common-sense view" mentioned earlier.

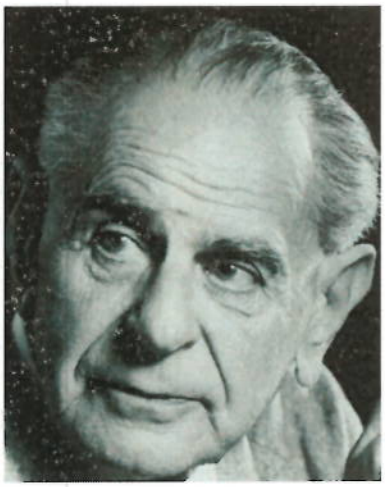


Image 7.4: Karl Popper. (Source: en.wikipedia.org)

According to Popper, science proceeds by conjectures (hypotheses) and refutations (deductive inferences that show the hypothesis to be false) — hypothesis and deduction (hence the name *hypothetico-deductivism*. Rather than trying to *prove* universal laws that constitute scientific knowledge by induction from observational evidence, as the inductivist would have us believe, science is in the business of proposing bold conjectures as laws describing what we see around us and then subjecting these conjectures to stringent tests to see if the law can withstand attempts to *falsify* it (hence the alternative name *falsificationism*). If it can withstand the tests then, though we are not in a position to claim it as true (when are we ever in such a position? — this is the force of the problem of induction) we may claim it as the best "law" currently available and so rational to believe, at least for the time being.

The popular notion that science is a body of established fact is entirely mistaken; conjectured scientific laws are, at any given time, those which have not yet been shown to

be false and, on balance, are the best we currently can conceive of to account for the nature of the world around us.

Notes:

7.7.2 Inductive Proof vs. Falsification

Popper's proposed solution to the problem of induction derives its force from a logical difference between the (supposedly inductive) proof of a universal claim and the (clearly deductive) falsification or disproof of a universal. Consider the claim 'All metal expands when heated'. No number of instances of metals expanding when heated can be sufficient to prove the claim true yet a single instance of a metal not expanding when heated will be enough to prove the claim false — i.e. "falsify" it. The claim says that all As are B yet, if we can find a single A that is not B then the universal claim will be shown to be false. Unlike inductive proof, this is uncontroversial.

The falsity of a universal statement can be conclusively inferred from certain singular (particular) statements. Thus there *is* a clear and uncontroversial logical relation between singular and universal statements: singular statements, though they cannot inductively prove universal ones, can falsify them. It is this justifiable logical relation that Popper relies on to explain how it is that observation and scientific data relates to scientific laws and theories.

Laws are not justified by being proved by the data, they are justified by being not disproved by the data.

Falsifiability

Let us look more closely at just what we mean when we describe some statement as falsifiable. The following claims are all falsifiable in the Popperian sense:

1. 'It never rains on Wednesdays.'
2. 'All metals expand when heated.'
3. 'Heavy objects fall straight downwards if not impeded.'
4. 'When a ray of light is reflected from a plane mirror, the angle of incidence is equal to the angle of reflection.'

Claim (1) can be falsified if the world is observed to be such that it rains on some Wednesday. Claim (2) can be falsified if at some particular time a metal is observed not to expand under heating. Claim (3) can be falsified if a heavy object, which was not impeded in any way, was observed not to fall straight downwards. Claim (4) would be falsified if some ray of light at some time were observed to be reflected from a plane mirror in such a way that the angle of incidence was different from the angle of reflection.

The definition of "falsifiability" then is the following: *an hypothesis is falsifiable if there exists a logically possible observation statement or set of observation statements that are inconsistent with it — i.e., which if established as true would falsify the hypothesis.*

Scientific laws then (given the requirement that they be falsifiable) are testable in spite of their being unprovable; they can be tested by systematic attempts to falsify them.

Examples of claims that are not falsifiable are:

Logical Truths — e.g. 'Either it is raining or it is not raining.'

Definitional Truths — e.g. 'All bachelors are unmarried males.'

Mathematical Truths — e.g. '$2 + 2 = 4$'

Certain Modal Truths — e.g. 'Luck is possible in sporting situations.'

This latter example is the stock-in-trade of many fortune-tellers and newspaper astrologists. Such claims can never be shown to be false because they are not capable of being falsified.

As further examples of seemingly unfalsifiable claims, consider the following:

1. 'The cosmos doubled in size overnight.'
2. 'God created the earth 6,000 years ago complete with fossil record.' (To test our faith perhaps — in this way, it is argued, we can consistently argue for creationism.)
3. 'The world came into existence only five minutes ago, complete with a "history".'

And what about:

4. 'Survival of the fittest'?

Falsifiability as the Demarcation of Scientific Statements

The falsificationist's view is that scientific hypotheses must provide information as to how the world is, and so therefore how it is not. In other words. scientific hypotheses

must have some information content in the sense of ruling out certain possibilities. In fact, scientific statements, in general, whether hypotheses, laws or simply observation statements, must have information content in this sense. They must, in this sense, be "testable". Claims that are true or false regardless of how the world is tell us nothing about the world itself, are not falsifiable (i.e. not testable), and thus are not scientific (though they may appear to be scientific). Because scientific statements (including scientific hypotheses) make definite claims about the world they have informative content and so must be falsifiable.

This simple fact is used to test whether statements count as "scientific" or not. If I tell you that you may be lucky in sport today I might appear to be making a prediction about your future. In this sense it may appear that I am making a scientific claim about the future. Or, consider the claim that the electron may curve anti-clockwise in the cloud chamber and not clockwise. These claims are not falsifiable; they are not testable. (They only say that something *might* happen, and the mere *possibility* is not falsified by its *actually* not happening.) They rule nothing out. (Anything is *possible*!)

Key Point: In order to be scientific, an hypothesis must be *falsifiable.* We cannot inductively prove such statements but we can deductively *disprove* them.

Notes:

7.7.3 How Scientific Knowledge Advances

Scientific hypotheses then are by their very nature falsifiable. And the scientific method proceeds by putting forward such hypotheses, however arrived at. Then subjecting them to stringent testing to see if they are, in fact, false. Observations are compared to the consequences predicted by the hypothesis. If the observations conflict with the prediction then the (falsifiable) hypothesis is actually *falsified.* If not, if they pass stringent testing against our observations, then the hypothesis can be provisionally accepted (and if it is suitably *general* then it will be a provisionally accepted *law*) until and unless it is later falsified. If a law is falsified then we can establish this deductively, and we then go on

look for new bold hypotheses that will explain all that the old rejected law will explain and moreover which will explain the observation that led to the rejection of the old law.

There is no room for induction in this picture. Deriving predictions needs only deductive logic, and so too inferring that an hypothesis is false given our observations. Nothing on this picture "proves" hypotheses; they are simply useful conjectures that have been put forward and not yet been shown false.

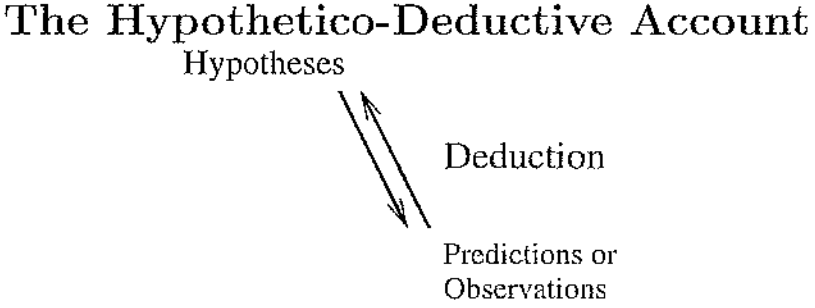


Figure 7.2: The Hypothetico-deductive account.

According to the hypothetico-deductive account then all our knowledge is, of its nature, provisional and will always be. At no stage are we in a position to be able to "prove" what we know to be true — it is always possible that it will turn out to be false. It is a simple fact of our intellectual history that nearly everything we have ever claimed at any time to know has later turned out to be false. A good example is Newton's Laws; they must have seemed so secure for two centuries until the Relativistic-turn early last century.

It is a mistake to try to prove a proposed law or theory to be true; to do so is to look for more assurance than will ever be available. What we can do is justify our preference for one candidate law (i.e. hypothesis) or theory (i.e. a collection of hypotheses) over another. (For example, the other was falsified, or was not scientific, or though scientific and not yet falsified was less informative so less useful than its preferred rival.) As theories are falsified we look for new bold conjectures that will explain all that the old rejected theory will explain and moreover which will explain the observation that led to the rejection of the old theory. *We learn from our mistakes; science progresses by trial and error.*

The popular notion that science is a body of established fact is entirely mistaken. Nothing in science is permanently established, it is changing all the time and not through the addition of new facts as the inductivist would have us believe. If we are rational then we will base our decisions on "the best available knowledge at the time", which is exactly what science provides us with.

Key Point: Popper's hypothetico-deductive account, falsificationism, describes science as a process of proposing hypotheses then deducing consequences from them for testing.

Hypotheses are accepted as provisional unless they are falsified by observation.

Notes:

7.7.4 Hypothetico-Deductivism and the Problem of Induction

So, it is claimed, at no point does induction play a role in assessing the status of scientific knowledge; the problem of induction does not arise. To be sure, we may as a matter of psychological fact invoke inductive methods to think up a conjectured law or hypothesis just as we might have hit upon it in a moment of sublime inspiration, blind drunkenness, or a dream but the psychological means whereby the law or hypothesis was arrived at tells us nothing about the its status (as acceptable or unacceptable). We do not have to face the problem of justifying *how we came up with* our conjecture, whether it was by inductive inference or our dream-inspired method; the way we think up some law or hypothesis is not something requiring justification. What *does* require justification is why we might persist with such a claim and take it as something we can work with rather than abandoning it — this is done by seeing if it can pass those tests applied to it, not by any use of the problematic principle of induction.

Key Point: Since hypothetico-deductivism does not rely on induction to justify scientific laws, there is no problem of induction.

Notes:

7.7.5 Problems

"Great" I hear you say. Problem solved. We can now describe scientific method as essentially hypothetico-deductive, thus establish our belief in scientific laws and theories as rational, and go on to explain how it might be that (returning to where we began) no rational person could deny human-induced global warming. Using (now) rationally justified scientific laws we can show how the hypothesis of human-induced global warming best explains the data we are confronted with, and so by inference to best explanation, rationally infer that such a phenomenon exists. All the evidence, along with rationally justified laws, points to the hypothesis and so we are rationally justified in accepting it.

But — and there is nearly always a 'but' in philosophical debate, as there often is in scientific debate — a hypothetico-deductive account does raise some interesting questions. Firstly, some have doubted that the hypothetico-deductive method actually avoids use of inductive justification of scientific laws. Recall the claim in §7.7.2: "laws are not justified by being proved by the data, they are justified by being not disproved by the data". Why should the mere fact that an hypothesis has not been disproved by the data give reason for believing it? Secondly, Popper presents cases where *one* theory is being tested against our experimental data, but hypotheses are tested in groups. When we "test" a theory, we are assuming a lot of other theories in the background. So if we find anomalous results should we reject the theory being "tested" or one of the other auxiliary hypotheses operating in the background? Lastly, even when anomalies *are* detected, we frequently do not go on to reject the theory thought to be at issue. Often we retain theories that have been successful and proved themselves powerful and look for ways to reconcile ourselves to the anomaly.

This last feature of actual scientific practice has led to more sophisticated accounts of science that see science as developing through stages. Thomas Kuhn's 1960s study of science and scientific method — *The Structure of Scientific Revolutions* — was a landmark work in this direction. Focussing heavily on how science actually functions, with a close study of the history of science, Kuhn introduced the notion of a "paradigm" into popular intellectual culture (for example, "Subvert the dominant paradigm"). His book was one of the most cited books of the twentieth century. According to Kuhn there are periods of "normal science" where people work to develop a particular scientific "paradigm" in a more-or-less Popperian way. Newtonian science formed one paradigm, for example, before being replaced by Einstein's revolution. Scientists commit to these paradigms and often persist with them even in the face of contrary evidence. Eventually, the contrary evidence mounts until the pressure is so great that the paradigm eventually breaks down. At this point there may be radical changes in our understanding of central scientific concepts (e.g. mass). So-called "revolutionary science" occurs and over time a new paradigm is established.

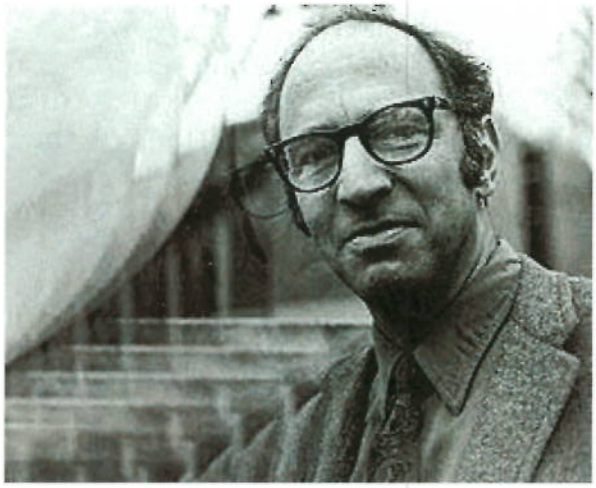


Image 7.5: Thomas Kuhn. (Source: en.wikipedia.org)

There is also the later work of the radical Paul Feyerabend (see his entertaining 1975 essay "How to Defend Society Against Science", for example – http://www.galilean-library.org/manuscript.php?postid=43842) who argued that there is no such thing as *the* scientific method. He claims that close study of the history of science shows that there is *no* set of rules, *no method*, that describes "scientific progress".

We won't pursue this further. But there is a lot more to be said here. Those interested might look at Alan Chalmers' very readable book *What Is This Thing Called Science?*, available from the Library – much of the notes so far follow this text. (There are a number of editions but most after edition 1 are very good.) A very good introduction to some of the historical material is J. Losee's *A Historical Introduction to the Philosophy of Science*, 3rd ed. (1993), and D. Oldroyd's *The Arch of Knowledge*, University of New South Wales Press (1986). Further material is also mentioned in the notes and footnotes.

Notes:

7.8 Exercise

Create your own glossary by writing down definitions of the following terms:

(a) Induction

(b) Singular statement

(c) Universal statement

(d) Hypothesis

(e) Fallible

(f) Weak induction

(g) A priori

(h) A posteriori

(i) Hypothetico-deductive method

(j) Falsifiable

(k) Falsified

7.9 Is Most of Biomedical Science False?

One of the key things to remember from all that we have said so far is that our scientific knowledge is *fallible*. That is, the claimed cause-effect relationships can never be known with certainty. But the discussion, so far, has been rather abstract and heavily focussed on strict (i.e. universal) law-like connections (e.g. Snell's Law or Newton's Laws). What does a view of science as fallible really amount to in practical terms? And what about law-like generalisations typically encountered in the medical sciences where strict law-like connections are never proposed – where, instead, they speak of *general tendencies* in the population to react in particular ways to particular medications (what are commonly termed "statistical generalisations")?

For example, when it is hypothesised that "cholesterol causes heart disease" no one thinks that a couple of cases of people with high cholesterol without heart disease would refute or falsify the general claim. (Compare this with possible counterexamples to Snell's Law that really would refute it.) Similarly, when medical researchers assert that "the drug statin reduces cholesterol" they do not think their hypothesis is refuted by a subject in a study who takes statin without any decrease in their cholesterol levels. It is recognised that these hypotheses are really of statistical form – they are not 100% general: "*Most* (i.e. some merely high percentage of) people taking statin will have a correlated reduction in cholesterol" and "*Most* people with high cholesterol levels will have a correlated increased incidence of heart disease", not *all*.

The causal connections here are so complicated that there will always be exceptions to any hard-and-fast rule correlating one with the other – whereas in the cases of laws of classical physics, for example, we take it that the causal connections hold without exception, for example, Snell's Law: the angle of incidence of a light beam on a plane surface is *always* equal to its angle of reflection. The most we can claim using statistical generalisations (and it is taken to be enough to guide medical intervention) is that there is a strong (i.e. statistically likely) tendency in the human population to react to statins by a reduction in cholesterol, and there is a strong tendency of those with high cholesterol to suffer heart disease.

What fallibilism tells us, of course, is that even these statistical generalisations cannot be known with certainty. But, remarkably, one recent scientist has claimed that, contrary to popular belief, the strength of our knowledge claims here are actually so weak that

> "most published [biomedical] research findings are probably false".[11]

[11] See the claim by Stanford epidemiologist Prof. John Ioannidis discussed in "Unreliable Research: trouble at the lab", *The Economist*, Oct. 19th 2013 – with a condensed version available at: http://www.economist.com/news/leaders/21588069-scientific-research-has-changed-world-now-it-needs-change-itself-how-science-goes-wrong. Ioannidis's much-cited paper itself can be found at: http://www.plosmedicine.org/article/info:doi/10.1371/journal.pmed.0020124.

When you see a report "Studies have shown that ...", he argues that most of the time these studies have not shown anything of the sort!

7.9.1 Can that really be right?

To admit to being fallible is one thing (News flash: We're not God!), but suggesting that our biomedical "knowledge" is most likely false is quite another. If we have reason to think it is probably false then it can't be rational to believe it, can it? Most studies, when claimed to have shown ..., have in fact not shown it at all, it seems.

The alarming claim of "probable falsity" rests on some key assumptions about testing hypotheses in the biomedical sciences. And then with some fairly basic statistical inferences we do, indeed, get Ioannidis's very surprising result. But we shall see that appearances can be deceptive here. (1) There is a *qualified* result that is, admittedly, quite counter-intuitive – which just goes to show how bad our intuitions are when it comes to statistical reasoning (as we saw in the breast cancer testing case that Peter discussed); but (2) this qualified result is liable to be misunderstood, and fallibilists (whether you think science works by induction or falsification) ought not be surprised by the situation.

So, what are the assumptions behind the argument for "probable falsity"? Well, the first thing to notice is that, when we are testing a statistical generalisation, our data set might *accidentally* produce a seemingly positive response in a population to a drug (e.g. statin). That is, the apparent effect in the drug trial (e.g. general lowering of cholesterol) was just a coincidence (the trial showed a lowering despite the statin having no effect). Such a result could sometimes be expected merely by chance, so we need to ensure that our data shows the expected effect occurring at a level higher than that predicted by chance alone.

If chance alone could explain the observed effect then there is no need to explain it by claiming that the drug in question was the cause! For example,

A: "So you took aspirin and your headache went away. Hah! Chance could explain that."

B: "Yeah, but we have observed this result several hundred times now so I think we need some better explanation. I know: aspirins (mostly) relieve headaches."

So the role of chance is recognised and attempts are made to minimise its effect in skewing results arrived at by testing. We require what are called "statistically significant" results – results that are so unlikely to have come about by chance that we are justified in taking them as evidence that an effect is occurring. But it is admitted that we cannot eliminate chance all together. Despite our best efforts, we recognise that we might sometimes infer a positive result from a statistically significant data set (where we think chance is so

very unlikely – e.g. seeing a strong positive result in a very large test population – that it is reasonable to suppose that a real effect is occurring), but we are wrong. Despite the result being so unlikely due to chance that it is reasonable to take it as evidence of a real effect, there is no real effect. Such a result is termed a "false positive". We have falsely supposed that a positive result is established.

For example, in the statin/cholesterol case, we may have reasonably inferred that the drug *does* reduce cholesterol – claiming a positive result – but (freakishly) it happened that the drug did not do that (the apparent effect was pure accident), and the positive claim is false. We have made an error – a "type I error", as it is called.

It is standard in biomedical testing to take all necessary steps in experimentation and testing to reduce this probability of "false positive" claims, of type I errors, to 0.05 (i.e. 5%). (NB: the choice of 5% is not fixed in stone and there is considerable discussion in this field as to whether higher or lower rates might be preferred. But 5% is the typical percentage of error tolerated.) That is to say, among the false hypotheses, we will wrongly claim up to 5% as true.

On this basis, surprisingly many people seem to think that our biomedical testing is highly reliable in that we can restrict type I errors to no more than 5% of our positive claims. Of course, we are, it seems, allowing that up to 5% of our hypotheses regarding positive effects will be wrong. Among the claimed positive results from biomedical testing – that a drug X_1 produces (general) effect Y_1 (e.g. statin generally reduces cholesterol), and drug X_2 produces (general) effect Y_2, etc., we can expect that up to 5% such positive claims are false, with the remaining 95% true. But that sounds ok:

It seems there is a 95% probability that any claimed (biomedical) result is a true positive result.

We cannot be certain about *all* our results, but 95% of them are nonetheless true.

Ioannidis points out that this is fallacious. (And the reasoning mirrors that in the breast cancer testing case.) To see this, assume that in testing we have:

– false positive or "type I error" rate = 5% (so of the false hypotheses 5% will nonetheless be declared true);

Assume also (as Ioannidis does) that the sensitivity or "power" of our testing is 80%; that is, of the true hypotheses 80% will be declared true – so of the true hypotheses 20% will nonetheless be declared false:

– false negative or "type II error" rate = 20%.

(NB: just as with tolerance of type I errors, what counts as an acceptable type II error rate is a matter of considerable debate. In some situations, say where we are trying

to find fatal diseases like breast cancer, a 20% type II error rate seems unacceptably high – after all, missing 20% of patients with this fatal disease seems problematic. But increasing the power or sensitivity of the test to avoid type II errors will typically also increase the prevalence of type I errors. This may, itself, also be a problem.)

And, finally, assume (along with Ioannidis) that the prior probability of any hypothesis under consideration being true (i.e. the probability that it is true prior to any data one way or the other) is 10%. That is, of all the hypotheses considered "interesting" enough to investigate, only 1 in every 10 are likely to be true.

Then some simple arithmetic shows that:

In fact there is only a 64% probability that any claimed (biomedical) result is a true positive result.

This is illustrated in Figure 7.3.

In fact, Ioannidis continues, for many studies the sensitivity or power is reduced to about 40% and in these cases there is less than a 50% probability that any claimed result is a true positive result. This is why Ioannidis says most published research findings are false.

On the good side, an hypothesis apparently "confirmed by studies" goes from a prior probability of 10% to 40%-64% (that's an increase in likelihood between 400% and 640%!) but such hypotheses are still likely enough to be false that it is grossly misleading to claim that "studies have established such hypotheses as true". So why would we? (It is often helpful to rid ourselves of errors by understanding how they arise.)

Apart from the obvious answer that we have simply made *some error of reasoning* when we suppose that studies (under the conditions described) are capable of establishing hypotheses as 95% certain, there are two particular errors that are likely at work here.

Error 1.

Firstly, the following argument is, for many, a good one:

"Were the hypothesis false then it is almost certain (95%) that we would reject it. So, since we have accepted it, it is almost certain (95%) that it is true."

The premise is true – it is just what we established when showing a type I error rate of 5% – so we wrongly infer on this basis that the conclusion is true. (NB: if the degree of certainty here was 100%, then the argument would be a good one. If there are no false positives then all accepted hypotheses are true! Can you prove this?) That is,

since pr(hypothesis is rejected / hypothesis is false) = .95, it may be thought to follow that:

pr(hypothesis is true / hypothesis is accepted) = .95.

But that is simply wrong!

More generally: $pr(A/B) \neq pr(notB/notA)$.

Key point: We should not confuse the accuracy of a test – i.e. the probability of a test to avoid type I errors (false positives) – with the probability that the test yields true positives.

Assume that, of 'interesting' hypotheses, 10% will be true. Imagine that 1000 are submitted for testing; 100 of these will in fact be true.

Now assume that the tests have a false positive rate of 5%, and a false negative rate of 20%. Of the 900 false hypotheses, 855 will be identified correctly as false, and 45 as true. Of the 100 true hypotheses, 80 will be identified correctly as true and 20 as false.

Testing will identify 125 hypotheses as true, but only 80 are in fact true: 36% of the 'new true' hypotheses are false. Also, 875 hypotheses are identified as false, of which 20 are true (an error rate of 2.3%). Of course, negative results are unlikely to be published.

Figure 7.3: How a small proportion of false positives can prove to be very misleading.

Error 2.

Nor should we confuse the likelihood of a test to avoid type II errors (false negatives), 80% remember, with the likelihood of the test to deliver up true positives. This is another, second, potential confusion that follows from yet another argument that many people think is a good one:

"Were the hypothesis true then it is almost certain (80%) that we would accept it. So, since we have accepted it, it is almost certain (80%) that it is true."

This involves a version of what some statisticians have called "the fallacy of the transposed conditional".[12] Just because it is 80% certain that if an hypothesis is true then we would accept it, it doesn't follow that it is 80% certain that if we accept it then it is true. That is,

Just because pr(hypothesis is accepted / hypothesis is true) = .8, it doesn't follow that:

pr(hypothesis is true / hypothesis is accepted) = .8.

More generally: $pr(A/B) \neq pr(B/A)$.

Key point: We should not confuse the power or sensitivity of a test – i.e. the probability of a test to avoid type II errors (false negatives) – with the probability that the test yields true positives.

Notes:

7.9.2 How to Increase Confidence in Our Results

So, on the basis of studies/trials large enough to meet the conditions outlined above in respect of type I and type II errors (conditions that many biomedical tests are typically required to meet), when they produce statistically significant results many will claim that the trials "show" the effect they were testing for. Such trials are often published as "establishing" the hypothesis in question, and reports will frequently state the hypothesis as if it has been "shown" or (even more alarmingly) "proven" – e.g. "Results show X

[12] See: Ziliak and McCloskey, *The Cult of Statistical Significance: how the standard error costs us jobs, justice, and lives*, University of Michigan Press (2008), Ch. 14, "Medicine Seeks a Magic Pill".

causes Y"; "Studies have proved that P causes Q". But, as we have seen, this is not justified. So how can we, at least, increase our confidence in our results?

Accentuate the Negative

One way to make confident claims is to focus on the negative – i.e. report the *negative* results. If we look back at our example where the confidence level in positive results is less than 50%, the confidence level in negative results is over 93% (only 60 of the 915 negatives are false negatives).

But there is a social/psychological barrier in the scientific community to such a focus on the negative. *The Economist* (p. 4):

> "[R]esearchers and the journals in which they publish are not very interested in negative results. They prefer to accentuate the positive ... Negative results account for just 10-30% of published scientific literature, depending on the discipline."

For Popper and those who think science progresses using falsification, of course, we can only ever be confident in the negatives (though we cannot be completely confident because of false negatives). *To ever think that we are establishing "the new True", as the positive results in the graphic shown earlier were described, is a serious error.* The positive results simply yield hypotheses that have not (yet) been falsified.

Inductivists, of course, *can* make positive claims with varying degrees of confidence, but even they should accept that we can only have very limited confidence in a positive result after testing. *To think that we are establishing "the new True" is a serious exaggeration.* The positive results simply yield hypotheses that have some degree of confirmation.

Repeat Testing

An obvious way, on an inductivist account, to increase confidence levels in results is to get more data, i.e. repeat the tests. In the Ioannidis example analysed above, passing the first trial raised the probability of the hypothesis under consideration from 10% to 40-64%. A repeat trial (using the same stringent type I and II error rates) can raise it even further, in principle eliminating 95% of the false positives arrived at after trial 1, with a resulting probability of a positive being a true positive being between 94% and 97%.

But there are social/psychological barriers in the scientific community here too. *The Economist* again:

> "[R]eplication is hard and thankless. Journals, thirsty for novelty, show little interest in it ... Most academic researchers would rather spend time on work

that is more likely to enhance their careers. This is especially true of junior researchers, who are aware that overzealous replication can be seen as an implicit challenge to authority.

There are ways, too, to make replication difficult. Reproducing research done by others often requires access to their original methods and data. A study published last month in *PeerJ* by Melissa Haendel, of the Oregon Health and Science University, and colleagues found that more than half of 238 biomedical papers published in 84 journals failed to identify all the resources (such as chemical reagents) necessary to reproduce the results."

7.9.3 Fallibilism - again!

What all of the foregoing shows, I think, is that the frequent claim that:

"Studies have shown ..."

should really be read as:

"Studies have (so far) been unable to rule out ..."

or, perhaps more optimistically if you are an inductivist:

"Studies have (so far) provided some support for ...".

Exaggerating the findings of studies does not help the search for truth (even if it sounds a lot more impressive). More and careful studies help.

Key Point: The logic of scientific method is one thing, its application by researchers is another. There are social and psychological factors that undermine even the cautious optimism of the inductivist that our study of scientific method recommends. Factoring in these "human elements" we might sometimes move from cautious scepticism to downright disbelief.

Working out how the world might work is hard enough in itself but, with key drivers in the scientific research arena seemingly encouraging quick routes to positive results, we should learn to be very cautious in our acceptance of results. As always, a sceptical attitude towards our abilities to fathom how things are is likely to lead to the most reliable of results.

Notes:

Part 4: High

Let us drink and be merry, dance, joke, and rejoice,
With claret and sherry, theorbo and voice!
The changeable world to our joy is unjust,
All treasure 's uncertain,
Then down with your dust!
In frolics dispose your pounds, shillings, and pence,
For we shall be nothing a hundred years hence.

We'll sport and be free with Moll, Betty, and Dolly,
Have oysters and lobsters to cure melancholy:
Fish-dinners will make a man spring like a flea,
Dame Venus, love's lady,
Was born of the sea;
With her and with Bacchus we'll tickle the sense,
For we shall be past it a hundred years hence.

Then why should we turmoil in cares and in fears,
Turn all our tranquill'ty to sighs and to tears?
Let 's eat, drink, and play till the worms do corrupt us,
'Tis certain, Post mortem
Nulla voluptas.
For health, wealth and beauty, wit, learning and sense,
Must all come to nothing a hundred years hence.

Coronemus nos Rosis antequam marcescant (1675), Thomas Jordan (c1612 – 1685).

Image 7.6: *The Bitter Tonic* (c1630), Adriaen Brouwer (1605 – 1638), Staedelsches Kunstinstitut und Staedtische Galerie, Frankfurt am Main. (Source: en.wikipedia.org)

Image 7.7: (overleaf) *The Garden of Earthly Delights* (1503 – 1504), Hieronymus Bosch (c. 1450 – 1516), Museo del Prado, Madrid. (Source: en.wikipedia.org)

Recall that there are five broad parts to this course. This is Part 4, "Drugs".

In Parts 1 and 2 of the notes we considered the importance of modelling in science, and how mathematics and computing are used when developing and applying models. In Part 3 we discussed what constitutes science and knowledge, and how this can be communicated.

This part of the notes includes two chapters. In both cases the context is pharmacokinetics: the study of what happens to a drug inside the body. The mathematical content which underpins our study of pharmacokinetics is calculus: rates of change, and areas under curves.

In the first chapter we study instantaneous and exact rates of change. Earlier we said that in science we are only interested in studying phenomena whose values change over time (after all, if something never changes then there is no particular need to study it: the values will remain unchanged forever). The primary mathematical tool that considers rates of change is the *derivative.* You will have studied these in detail at school; in SCIE1000 we do not focus on how to *find* derivatives, but we instead focus on how to *use* and *interpret* them.

In the second chapter we study areas under curves (AUCs). The primary mathematical tool for analysing AUCs is the *integral.* We will cover several methods of finding AUCs, but will mainly use the *trapezoid rule.* More importantly, you will need to know how to use and interpret the results.

You may be surprised to see the important roles played by derivatives and AUCs in pharmacology. For example, once a drug is administered, key determinants of the impact of the drug are its maximum blood concentration C_{max}, the time at which this occurs t_{max}, and the total exposure of the body to the drug, which is the AUC. Other related phenomena include the *bioavailability* of drugs administered by different routes, the *bioequivalence* of two different drugs, and even such well-known concepts as the *Glycaemic Index* (GI) of foods.

Chapter 8: **Sex, drugs and rates of change**

I met with a gal and we went on a spree
She taught me to smoke and to drink whuskey.
Cigareets and whuskey and wild wild women
They'll drive you crazy, they'll drive you insane,
Cigareets and whuskey and wild wild women
They'll drive you crazy, they'll drive you insane.
And now I'm feeble and broken with age
The lines on my face make a well written page.
I'm leavin' this story how sad but how true
On women and whuskey and what they will do.

Artist: *Jim Croce.* Song: *Cigarettes, whiskey and wild women.*

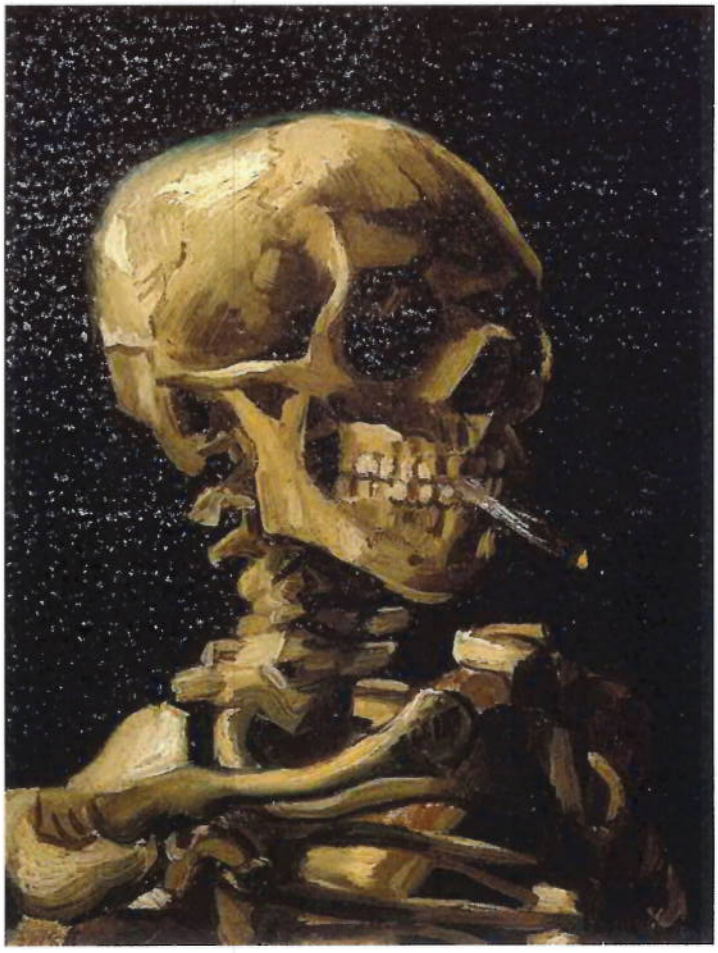


Image 8.1: *Skull with a burning cigarette* (c1885), Vincent van Gogh (1853 – 1890), Van Gogh Museum, Amsterdam. (Source: en.wikipedia.org)

8.1 Pharmacokinetics and rates of change

***Question** 8.1.1*

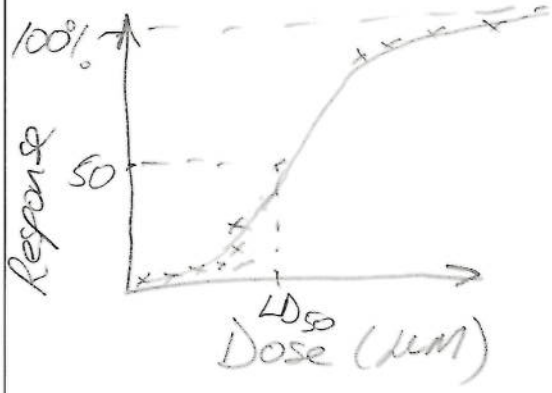
What are pharmacodynamics and pharmacokinetics?

Pharmacology (drugs)

Change

Pharmaco-Dynamics (PD)

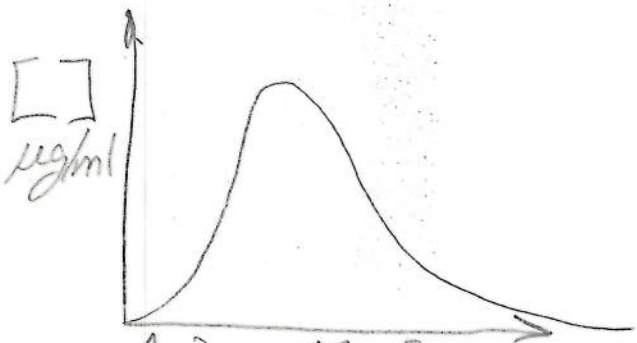
- Drug on body
- Dose-Response curve



Slope = ROC

Pharmaco-Kinetics (PK)

- Body on drug



Power — Exponential decay

surge function

$c = at^{p} e^{-kt}$

AUC = Total exposure

Some drug-related terminology

Broadly speaking, a *drug* is any externally derived chemical substance introduced into an organism that affects the function of that organism. Drugs may enhance physical or mental well-being, and include both medicinal and so-called recreational drugs.

Pharmacology studies the properties of drugs and their effects on living organisms.

Pharmacokinetics studies what happens to drugs inside the body, particularly the extent and rates of **absorption**, **distribution**, **metabolism** and **excretion**.

Pharmacodynamics studies the interaction of drugs with the cells of the body. (We will not cover pharmacodynamics in SCIE1000.)

Drug concentrations

After the administration of a drug, key determinants of its impact on the body are the drug **concentration** in the bloodstream, which is commonly measured as mass per volume (such as mg/L), and the **time** over which that concentration occurs. Concentrations can be measured at various times after drug administration and plotted on a *drug concentration curve*.

- Mathematics and functions are particularly important when modelling the *change* in drug concentrations over time, as they help to predict the *impact* of the drug and the *timing* of subsequent interventions.

Case Study 18: Zoloft and depression

- Depression is one of the most common mental health problems.
- Unlike many health problems, depression (and other mental illnesses) can occur more frequently in young adults than in older adults; see Figure 8.1.

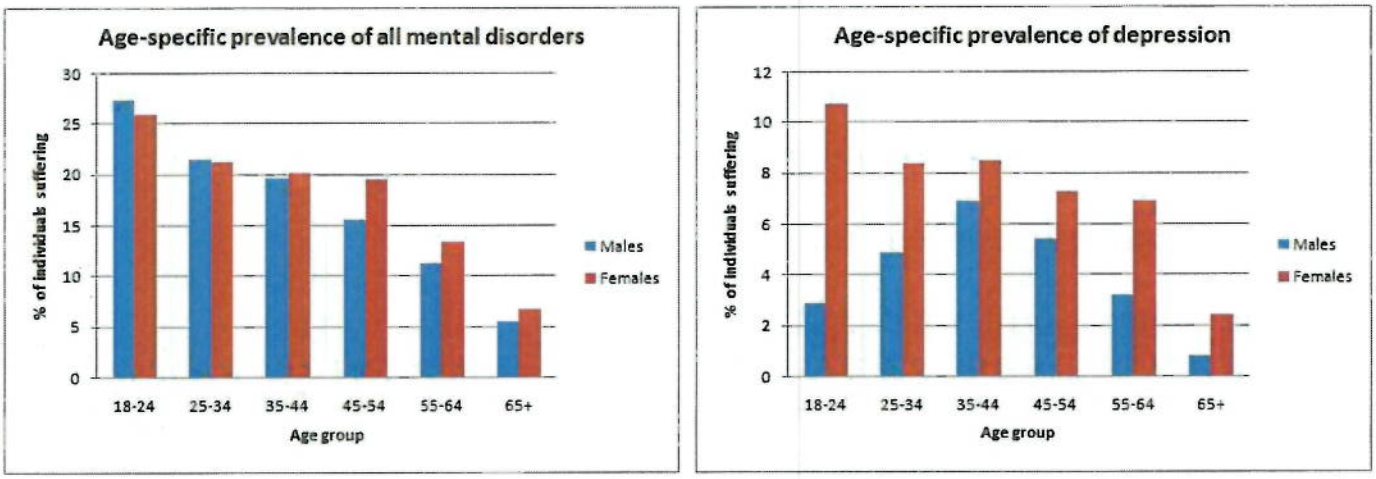


Figure 8.1: Age-specific prevalence of mental disorders and depression in Australian adults. (Source: National Survey of Mental Health and Wellbeing, 2007, Australian Bureau of Statistics.)

- There are multiple treatments available for depression, including a variety of therapy-based treatments, and pharmacological interventions.
- *Zoloft* (and a number of generically branded equivalents) is the brand name of the drug *sertraline hydrochloride*, which is an antidepressant of the SSRI class (Selective Serotonin Reuptake Inhibitor).
- The Consumer Medicine Information fact sheet explains that SSRIs "... are thought to work by blocking the uptake of a chemical called serotonin into nerve cells in the brain. Serotonin and other chemicals called amines are involved in controlling mood".
- Zoloft is the most commonly prescribed antidepressant in Australia, and one of the most prescribed drugs overall on the Australian Pharmaceutical Benefits Scheme.
- Zoloft is taken orally as a pill. The usual dosage ranges from 25 mg per day to 200 mg per day.
- Zoloft has a number of comparatively mild side effects (including insomnia, loss of appetite, and some sexual impairment), and is generally believed to be both effective and well tolerated.
- However, there has been media controversy in recent years about the possible adverse impacts of SSRIs on a small number of people.

Question 8.1.2

Drug concentration curves (for sertraline or other drugs) allow pharmacologists to observe, measure and analyse factors including:

(a) the peak drug concentration C_{max};

(b) the time t_{max} at which C_{max} occurs;

(c) the *half-life* $t_{1/2}$ of the drug, which is the time taken for the concentration to fall to half of its previous value;

(d) the maximum rates of drug absorption/removal and when these occur;

(e) the "total exposure" of the body to the drug.

Figure 8.2 shows the average blood sertraline concentrations for 11 young women involved in a study [45]. Participants received daily oral doses of sertraline over 30 days (to achieve 'steady state' concentrations), then a final dose was administered and blood concentrations monitored. Mark on the graph the values (or possible values) of each of (a) to (e) described above.

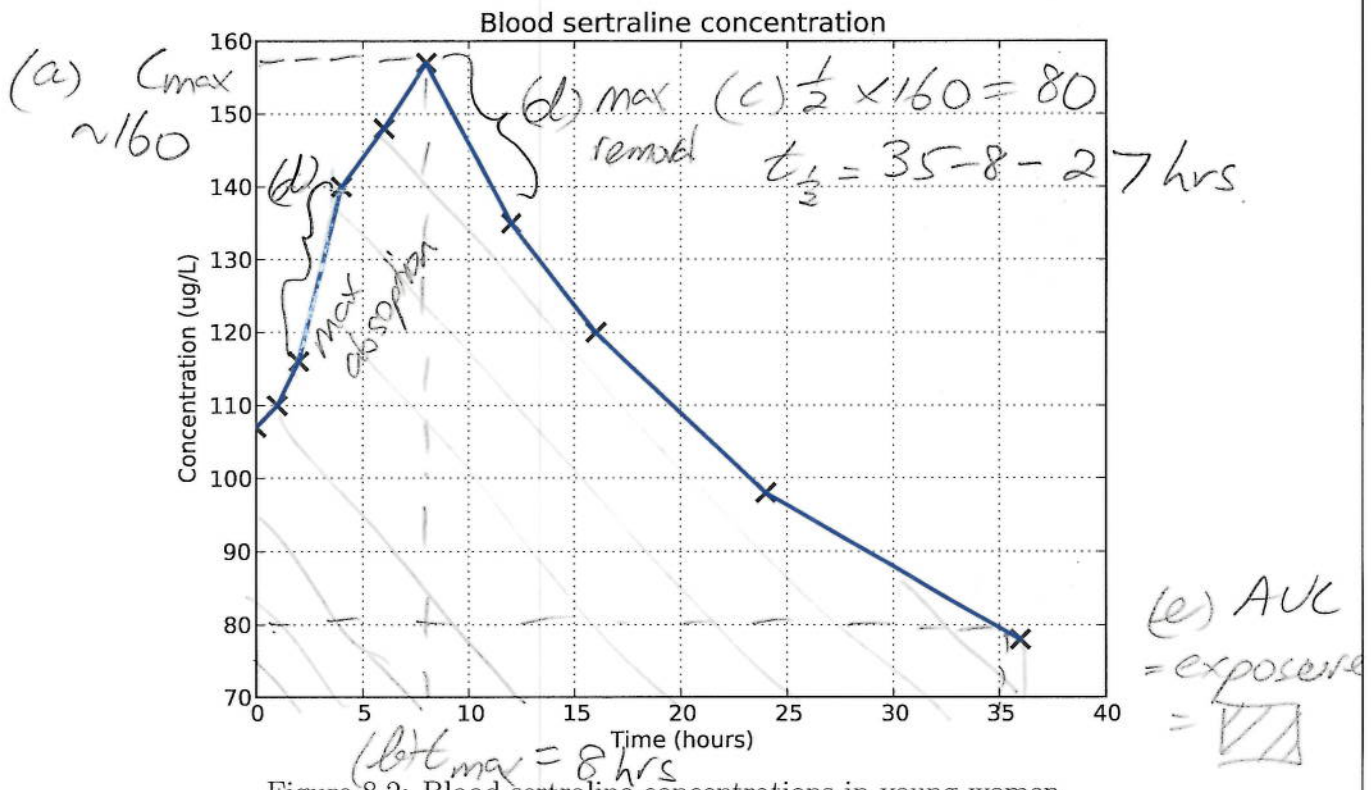


Figure 8.2: Blood sertraline concentrations in young women.

- Compare the information on Zoloft in the following example with some of the features/observations in Example 8.1.2. Also note the use of mathematical rates of change in the example.

Example 8.1.3

(The following is taken from the sertraline fact sheet at *www.pbs.gov.au.*)

"**Pharmacokinetics:** In humans, following oral once-daily dosing over the range of 50 to 200 mg for 14 days, mean peak plasma concentrations (C_{max}) of sertraline occurred between 4.5 to 8.4 hours post dosing. The average terminal elimination half-life of plasma sertraline is about 26 hours. Based on this pharmacokinetic parameter, steady-state sertraline plasma levels should be achieved after approximately one week of once-daily dosing. Linear dose-proportional pharmacokinetics were demonstrated in a single dose study in which the C_{max} and area under the plasma concentration time curve (AUC) of sertraline were proportional to dose over a range of 50 to 200 mg.

Dosage: Adults (18 years and older) The usual therapeutic dose for depression is 50 mg/day. ...patients not responding to a 50 mg/day dose may benefit from dose increases up to a maximum of 200 mg/day. Given the 24 hour elimination half-life of sertraline, dose changes should not occur at intervals of less than 1 week. The onset of therapeutic effect may be seen within 7 days ....

Use in Children and Adolescents aged less than 18 years: Sertraline should not be used in children and adolescents below the age of 18 years for the treatment of major depressive disorder. The efficacy and safety of sertraline has not been satisfactorily established for the treatment of major depressive disorder in this age group.

67-270 × recommended dose!

Overdosage: On the evidence available, sertraline has a wide margin of safety in overdose. Overdoses of sertraline alone of up to 13.5 g have been reported. Deaths have been reported involving overdoses of sertraline, primarily in combination with other drugs ...."

- The general shape of the blood sertraline concentration curve shown in Figure 8.2 is typical of many drug concentration curves. The corresponding functions are sometimes called *surge* functions.

Surge functions

In a **surge** function, the value initially rises rapidly before falling off exponentially over time. A general equation for a surge function is

$$f(t) = at^p e^{-bt}$$

where the values of a, p and b depend on the phenomenon. Figure 8.3 shows the general shape of a surge function.

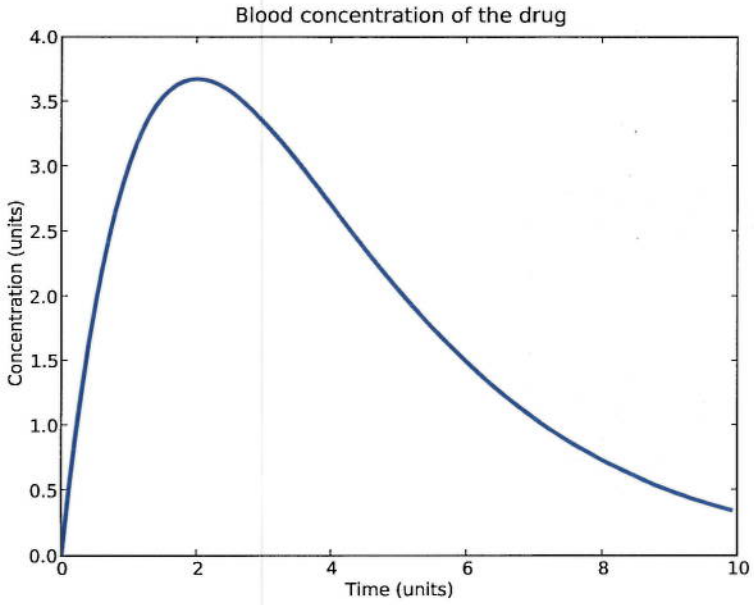


Figure 8.3: General shape of a surge function.

Question 8.1.4

(a) Give some physical reasons why blood concentration curves typically take the form of surge functions.

A Absorption
D Distribution
M Metabolism
E Excretion

Rapid ↑ Power
- ADME happening simultaneously

Slower decay Exponential

(continued over)

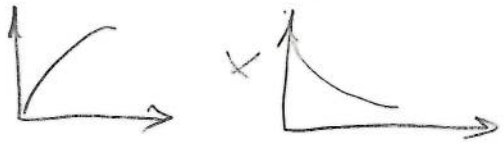
Question 8.1.4 *(continued)*

(b) Explain mathematically why functions of the form $f(t) = at^p e^{-bt}$ have a 'surge function shape'.

$at^p \times e^{-bt}$

① power × ② exponential

t = time



as $t \uparrow$, dominance shifts from ① to ②

- both occur all the time

(c) Soon we will study some examples of surge functions, including blood concentrations of:

- paracetamol: $C_1(t) = 14t^{0.6}e^{-0.5t}$ μg/mL.
- a long-lasting contraceptive: $C_2(t) = 1.4t^{0.15}e^{-0.02t}$ ng/mL.

Without drawing them, briefly discuss how the graphs of C_1 and C_2 would appear, including their similarities and differences. (Ignore the differences in concentration units. For C_1, time is measured in hours, and for C_2, time is measured in days.)

Paracetamol

- Power higher ∴ faster ↑ (eg. for headache)
- Exponential higher ∴ faster ↓ (doesn't need to effective for very long)
- Higher dose (μg) vs (ng) relative
- Constant multiplier 14 > 1.4 ∴ higher C_{max}

∴ paracetamol is fast acting for headache ✓

contraceptive is long-lasting @ a low dose.

End of Case Study 18: Zoloft and depression.

- Pharmacokinetics is particularly concerned with the *rate* at which the drug concentration *changes*.
- The concept of one quantity changing as another quantity changes, and the rate at which the change occurs, is crucial to understanding and modelling many processes in science, engineering, social sciences and economics.

***Example** 8.1.5*

In addition to explaining the dynamics of drug concentrations in the body, analysing rates of change is important for solving problems such as:

- landing a space capsule on the moon with minimum fuel usage;
- predicting the spread of ash from a volcanic eruption;
- modelling earthquakes and tsunamis, and predicting which areas will be affected, and when;
- predicting future populations of two interacting species;
- estimating the impact of a vaccination program on the spread of a disease;
- predicting the impact of a constricted artery on blood flow;
- minimising risk in a share portfolio;
- determining the time taken to reach equilibrium in a chemical reaction; and
- predicting the time at which a student will attain a certain threshold level of knowledge about a topic.

- We will cover two methods for analysing rates of change:
  - *average* rates of change; and
  - *exact* rates of change.

8.2 Average rates of change

- The *average rate of change* measures the average change between two observed values of some phenomenon.
- In science, average rates of change are usually measured over time, such as 60 m s^{-1}.

Average rate of change

Let (x_1, y_1) and (x_2, y_2) be two points. The **average rate of change of y with respect to x** between these points is the *slope of the straight line joining the points.* As we saw earlier, the slope equals the change in y values divided by the change in x values, so:

$$\frac{\text{change in } y}{\text{change in } x} = \frac{\Delta y}{\Delta x} = \frac{y_2 - y_1}{x_2 - x_1}.$$

(Note that Δ is the Greek capital letter "Delta", and usually means "the change in the value of".)

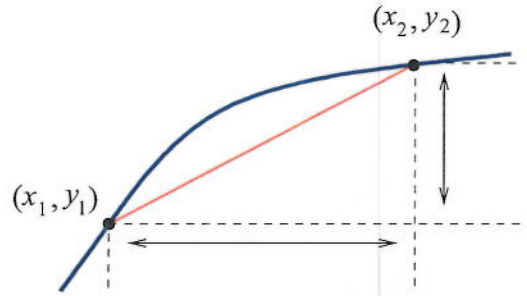


Figure 8.4: Average rate of change.

Example 8.2.1

The concentration of atmospheric CO_2 has risen by about 70 ppm over the last 50 years. Hence the average rate of change over this time is:

$$\frac{70 \text{ ppm}}{50 \text{ years}} = 1.4 \text{ ppm year}^{-1}.$$

Case Study 19: **Cigarettes**

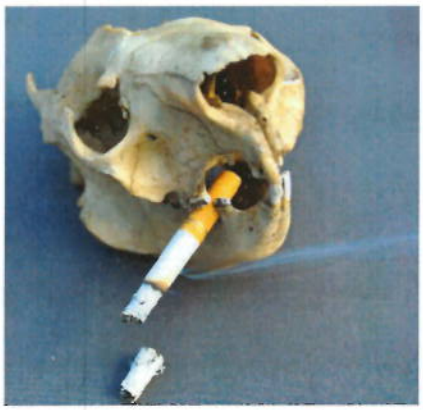


Photo 8.1: Cool, sophisticated smoker! (Source: PA.)

- Nicotine is a highly addictive, poisonous alkaloid found in a number of plants, including tobacco.
- After inhaling tobacco smoke, nicotine typically enters your blood stream within five seconds, and reaches your brain after about 10 seconds.
- In addition to nicotine, tobacco products also contain a large number of other compounds (including heavy metals, poisons and radioactive materials), many of which are toxic or known carcinogens.

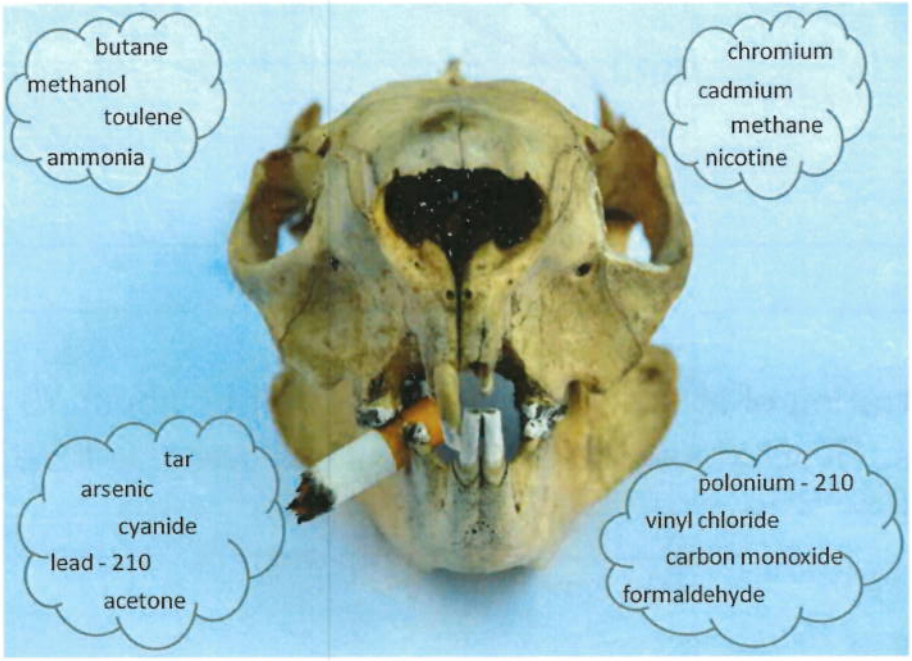


Photo 8.2: "Smoking kills": the joy of cigarette smoke. (Source: PA.)

- Smoking is a risk factor for many types of cancer. Figure 8.5 shows the proportions of various cancers in Australians directly attributable to smoking; see [1].

Cancer site	Male %	Female %	Cancer site	Male %	Female %
Lung	89	70	Larynx	69	60
Oral cancers	52	42	Renal pelvis	51	43
Oesophagus	50	41	Anus	39	29
Bladder	38	28	Vulva	–	32
Pancreas	23	16	Penis	21	–
Kidney	17	12	Stomach	12	8

Figure 8.5: Proportions of various cancers that are directly attributable to smoking.

- From 1991 to 2001, the male incidence rate for cancers attributable to smoking fell by an average of 1.4% per year, while the rate for females rose by 0.7% per year.
- Over the same period, mortality rates fell by 1.9% per annum for males and rose by around 0.1% per annum for females.
- Smoking is also a major risk factor for morbidity and mortality from other causes.
- As we saw earlier, in 1960 the Framingham Heart Study showed that smoking increases the risk of heart disease.
- Around 17% of all deaths from heart disease are due to smoking.
- Approximately 16% of Australians smoke, and there are around 15000 smoking related deaths in Australia each year.

***Question* 8.2.2**

Figures 8.6 and 8.7 show measured blood nicotine concentrations $N(t)$ after smoking a cigarette, starting at time $t = 0$ minutes (see [3]). (Note that measurements were taken at discrete time intervals; we show connecting lines in the graph only for clarity.)

(continued over)

Question 8.2.2 (continued)

t (min.)	0	3	6	10	14	20	35	65	95
$N(t)$ (ng/mL)	5	11	15.4	13.4	12.8	11.3	9.8	8	7

Figure 8.6: Measured blood nicotine concentrations $N(t)$ after smoking a cigarette.

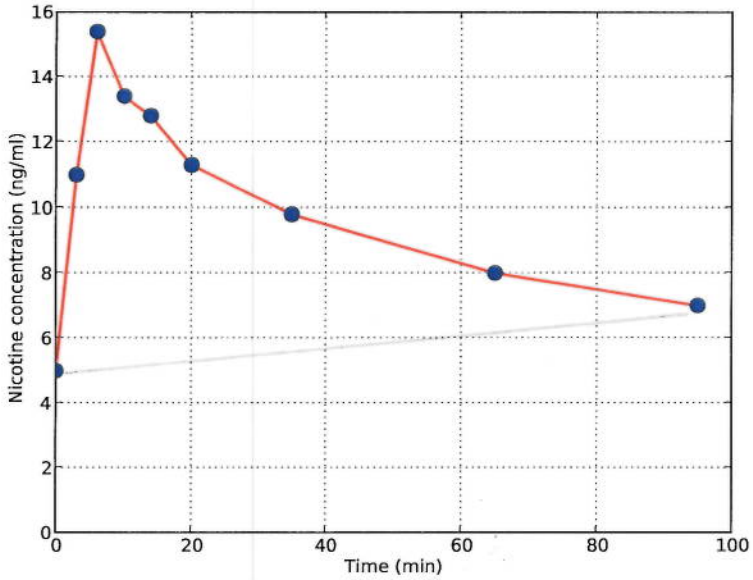


Figure 8.7: Graph of measured blood nicotine concentrations after smoking.

(a) Describe and explain the main features of the shape of the graph.

- Rapid initial ↑
- A [peak]
- Slow decline

∴ surge fn $y = at^1 e^{-bt}$

(b) Find the **total** change in concentration from $t = 0$ to $t = 95$ min.

$\Delta [\ ] = 7 - 5 = 2$ ng/ml

(c) Find the **average** rate of change in concentration from:

(i) $t = 0$ to $t = 95$ min; and $m = ROC = \frac{7-5}{95-0} = 0.021$ ng/ml/min

(ii) $t = 10$ to $t = 95$ min. $ROC = \frac{7-13.4}{95-10} = \frac{-6.4}{85} = -0.075$ ng/ml/min

End of Case Study 19: Cigarettes.

TIP: Leave the units as they are!

Note: ng/ml/min $= \frac{ng}{ml \cdot min} \neq \frac{ng\ min}{ml}$

$6/3/2 = \frac{6}{3 \times 2} \neq \frac{6 \times 2}{3}$

8.3 Derivatives and rates of change

- Rather than measure the average rate of change between two points, in many situations it is more useful to measure the *exact* rate of change at a point. The mathematical term for an exact rate of change is *derivative*.
- In SCIE1000, you will not be finding derivatives (in general), but will need to interpret and use them.

Derivatives

If $y = f(x)$ is a function, then the derivative y' is a new function that gives the exact rate at which y is changing with respect to x.

The **value** of the **derivative** at any point describes the behaviour of the **function** at that point. At any point:

- if y' is **positive** then the function y is **increasing**;
- if y' is **negative** then the function y is **decreasing**; and
- if the function y has a **local maximum** (peak) or **local minimum** (trough) at a point, then y' **equals zero** at that point.

The *derivative of the derivative*, or *second derivative*, is denoted f''.

Question 8.3.1

While smoking tobacco, the body absorbs many chemical compounds in addition to nicotine, including cyanide (which is highly toxic to humans). Figure 8.8 shows blood cyanide concentrations after smoking a cigarette, starting at time $t = 0$ minutes; see [30].

t (min.)	0	5	10	15	20	25	35	65
conc. (μg/mL)	0.11	0.43	0.21	0.16	0.14	0.15	0.125	0.1

Figure 8.8: Measured cyanide concentrations in the blood of a person after smoking a cigarette.

(continued over)

Question 8.3.1 (continued)

The function $C(t) = 0.1 + 0.3t^{0.6}e^{-0.17t}$ μg/mL is a reasonable model of the measured blood cyanide concentrations. Figure 8.9 shows a plot of $C(t)$, along with the measured data values.

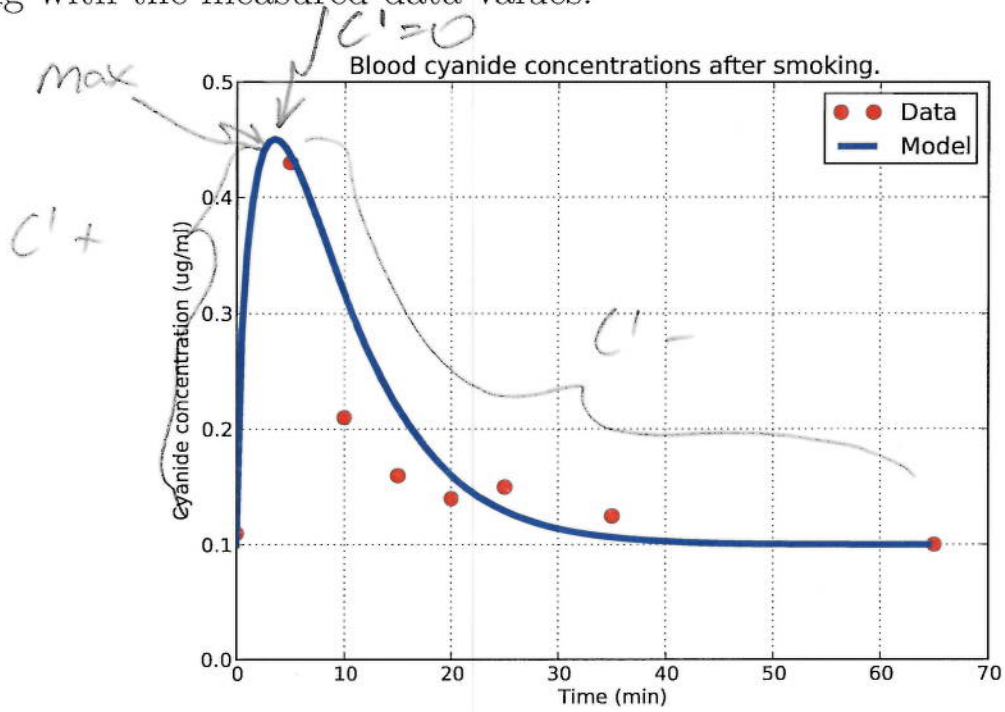


Figure 8.9: Measured and modelled blood cyanide concentrations after smoking a cigarette.

(a) What is the physical meaning of C'?

C' = ROC of [cyanide] over time = rate that cyanide is eliminated from body.

(b) On the graph:

(i) mark with a cross any points at which $C' = 0$;

(ii) label any local maxima/minima with the word 'max'/'min';

(iii) identify all regions where C' is positive/negative.

(c) What is happening physically when C' is:

(i) positive? AD > ME [cyanide] ↑

(ii) zero? AD = ME [cyanide] unchanged

(iii) negative? AD < ME [cyanide] ↓

8.4 Derivatives and Newton's method

***Question** 8.4.1*

The blood concentration of an injected long-lasting female contraceptive (medroxyprogesterone acetate or MPA) in ng/mL can be modelled by the function $C(t) = 1.4t^{0.15}e^{-0.02t}$. The graph of $C(t)$ is:

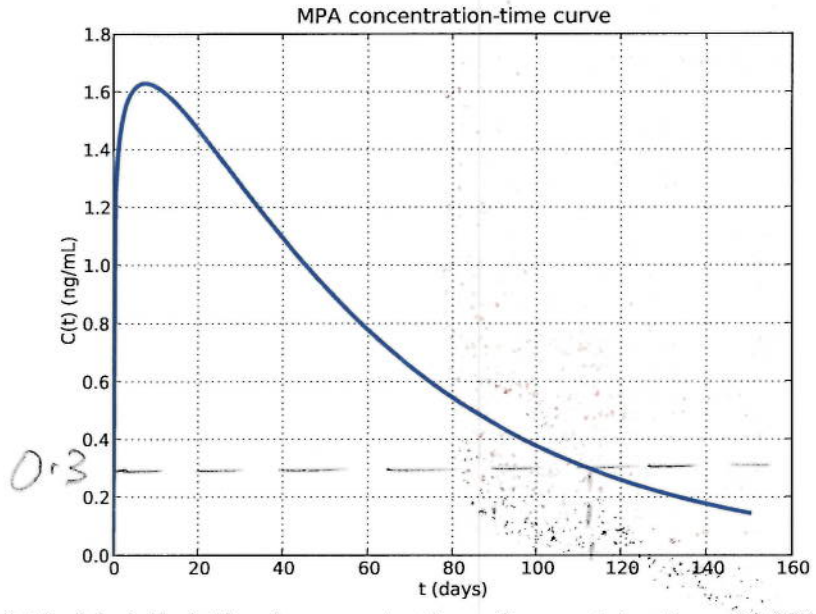


Figure 8.10: Modelled blood concentration after an injection of MPA.

(a) The minimum blood concentration for reliable contraception is 0.3 ng/mL. Estimate the time at which reliable contraception ceases.

$\sim$112 days $\approx$ 16 weeks

(b) Rewrite Part (a) as an equation to be solved.

$C(t) = 0.3 = 1.4\,t^{0.15}\,e^{-0.02t}$

(c) How could the equation in Part (b) be solved?

t appears twice. Cannot be solved analytically. Can use trial + error or an iterative procedure such as NEWTON'S METHOD

- Some equations are difficult or impossible to solve **exactly**. An alternative is to find an **approximate** solution, using *solution-finding* algorithms, which involve repeatedly applying similar mathematical steps or *iterations*.
- Usually, a *numerical error* is associated with the approximate solutions. These errors can often be reduced by performing more iterations.

- *Newton's method* is an iterative solution-finding algorithm which uses an *initial estimate* of a solution and a derivative to find a solution. Newton's method does not always *converge* to a solution, but will usually converge if the initial estimate is 'good enough'.
- Note that Newton's method only solves equations of the form $f(x) = 0$. Before applying Newton's method, the equation may need rearranging. For example, to use Newton's method to solve the equation in Part (b) of Question 8.4.1, we instead solve $C(t) - 0.3 = 0$.

Newton's method

Informal description: To solve $f(x) = 0$:

1. Choose an initial estimate of the solution.
2. Calculate a new estimate using the old estimate and the derivative. (The new estimate is hopefully better than the old.)
3. Stop if the new estimate is sufficiently accurate or if too many steps have been taken. Otherwise, return to Step 2.

Formal description: To solve $f(x) = 0$:

1. Let x_0 be an initial estimate of a solution of f that is 'sufficiently close' to an actual solution of f. At the ith iteration ($i = 0, 1, 2, \ldots$), x_i is the current approximation of the actual solution.
2. Calculate the next estimate: $\displaystyle x_{i+1} = x_i - \frac{f(x_i)}{f'(x_i)}$
3. **(a)** If the value of x_{i+1} is sufficiently accurate then stop; x_{i+1} is the estimated solution.

   (b) If x_{i+1} is not sufficiently accurate after a certain number of steps then stop, because the method is probably not converging to a solution. Choose a 'better' value for x_0 and start again.

   (c) Otherwise, return to Step 2.

- Newton's method is based on equations of straight lines!

- Let the initial estimate of an unknown solution of $f(x)$ be x_0. Newton's method calculates the next estimate x_1 by extending a line from the point $(x_0, f(x_0))$ to the x-axis, with the slope of the line equal to the value of the derivative f' at the point x_0; see Figure 8.11.

***Question* 8.4.2**

Use the gradient of a straight line to develop Newton's method.

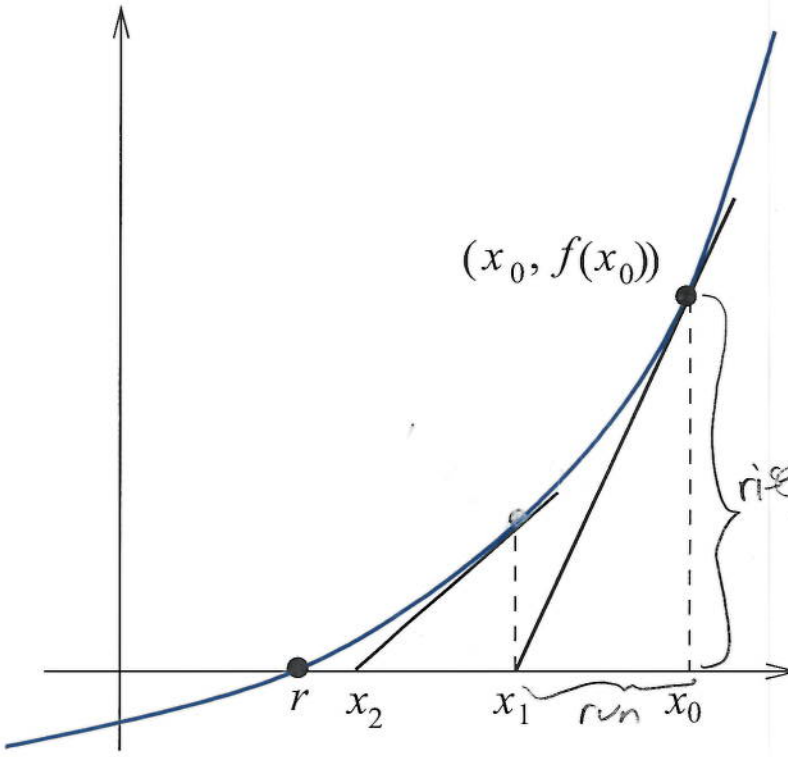


Figure 8.11: Two steps of Newton's method.

$$\text{Slope} = \frac{\text{rise}}{\text{run}}$$

$$y' = f'(x_0) = \frac{f(x_1) - f(x_0)}{x_1 - x_0}$$

$$= \frac{0 - f(x_0)}{x_1 - x_0}$$

$$f'(x_0) = -\frac{f(x_0)}{x_1 - x_0}$$

rearrange for x_1

$$\Rightarrow x_1 - x_0 = -\frac{f(x_0)}{f'(x_0)}$$

$$\Rightarrow \boxed{x_1 = x_0 - \frac{f(x_0)}{f'(x_0)}}$$

- If x_0 is sufficiently close to the solution, then the new approximation x_1 will be better than x_0.

- These steps continue until either Newton's method finds a sufficiently accurate approximation to the solution, or the process has taken too many steps without converging to a solution.

***Example** 8.4.3*

Use Newton's method to estimate $x = \sqrt{12}$.

To solve $x = \sqrt{12}$, we can solve $x^2 = 12$, or $x^2 - 12 = 0$.
Hence, let $f(x) = x^2 - 12$, so we need to solve $f(x) = 0$.

To apply Newton's method, we first need to find the derivative and then choose an initial estimate of the solution:

- Because $f(x) = x^2 - 12$, we have $f'(x) = 2x$.
- We know that $\sqrt{12}$ is between 3 and 4, so we will use $x_0 = 3$ as the initial estimate of the solution. (We could choose other estimates but $x_0 = 3$ is likely to be "close" to the solution.)

Now we have everything we need to use Newton's method. Recall that the equation for finding the next estimate of the solution is:

$$x_{i+1} = x_i - \frac{f(x_i)}{f'(x_i)}.$$

Performing three steps of Newton's method gives the results shown in Figure 8.12, with the last column showing the sequence of approximations to the solution.

i	x_i	$f(x_i) = x_i^2 - 12$	$f'(x_i) = 2x_i$	x_{i+1}
0	3	−3	6	3.5
1	3.5	0.25	7	3.4642857
2	3.4642857	0.001275	6.92857	3.4641016

Figure 8.12: Using three iterations of Newton's method to find $\sqrt{12}$.

After three steps, the estimate of $\sqrt{12}$ is $x_3 = 3.4641016$. Note that:

- The estimated solution barely changed between x_2 and x_3.
- The estimate of the solution is quite accurate; in fact, x_3 is correct to seven decimal places.

8.5 Pleasures of the flesh and derivatives

Image 8.2: "Who does not love wine, wife and song will be a fool for his lifelong!". (Kimmel and Voigt, New York, 1873. Source: en.wikipedia.org.)

The mistress drinks, the master drinks,
the soldier drinks, the priest drinks,
the man drinks, the woman drinks,
the servant drinks with the maid,
the swift man drinks, the lazy man drinks,
the white man drinks, the black man drinks,
the settled man drinks, the wanderer drinks,
the stupid man drinks, the wise man drinks,
The poor man drinks, the sick man drinks,
the exile drinks, and the stranger,
the boy drinks, the old man drinks,
the bishop drinks, and the deacon,
the sister drinks, the brother drinks,
the old lady drinks, the mother drinks,
this man drinks, that man drinks,
a hundred drink, a thousand drink.
Six hundred pennies would hardly
suffice, if everyone
drinks immoderately and immeasurably.
However much they cheerfully drink
we are the ones whom everyone scolds,
and thus we are destitute.
May those who slander us be cursed
and may their names not be written in
the book of the righteous.

Artist: Carl Orff, *When we are in the tavern* (from *Carmina Burana*).

- We will now apply derivatives to some drug concentration graphs.

Case Study 20: **Whisky (back to BAC)**

- A standard drink contains 10 g of alcohol.
- Usually, the measure of Blood Alcohol Concentration (BAC) is the percentage of total blood volume that is alcohol (or equivalently, grams of alcohol per litre of blood). In Australia the legal blood alcohol content for driving is 0.05%, or 0.5 g/L.

- Unlike many other drugs, the rate of alcohol metabolism is roughly constant (called a *zero-order* reaction in Chemistry).
- The rate of metabolism is usually independent of the BAC because typical levels of alcohol consumption saturate the metabolising capacity of enzymes within the liver.

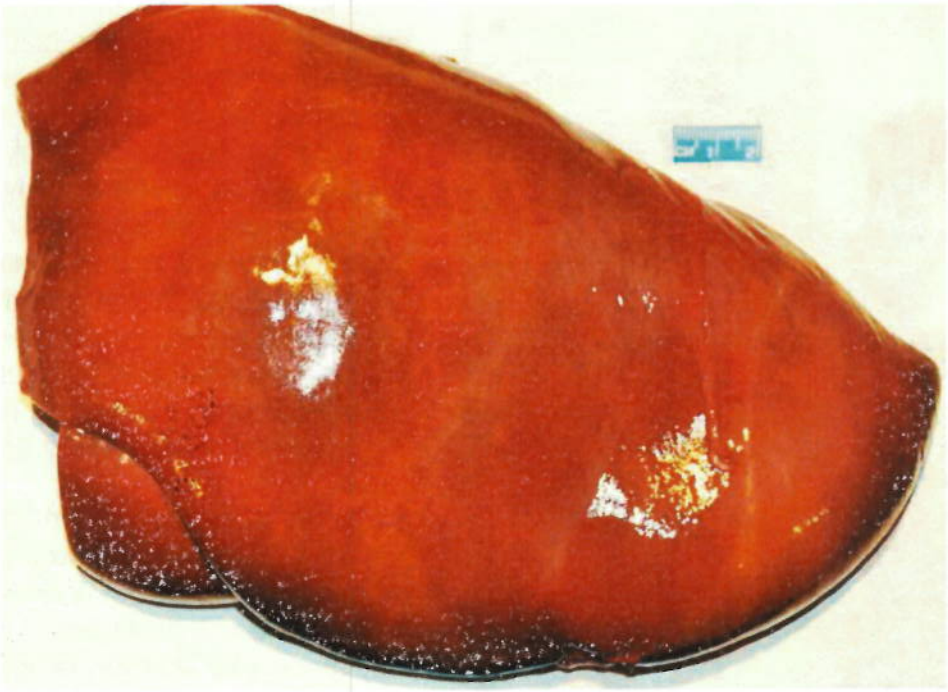


Photo 8.3: Calf liver. (Source: PA.)

- The exact rate of metabolism varies between individuals, influenced by factors such as age, mass (weight) and gender.
- A graph of BAC from the time drinking commenced will show a rapid initial rise during the absorption phase, prior to a decline in concentration during the elimination phase.
- Because the rate of alcohol metabolism tends to be constant, a graph of BAC from the time of peak concentration shows a linear decline until metabolism is almost complete.

Photo 8.4: Only in Poland. (Source: PA.)

Question 8.5.1

Figure 8.13 shows some BAC measurements (see [55]). Let $B(t)$ represent the straight line modelling the BAC from $t = 1$ h to $t = 6$ h.

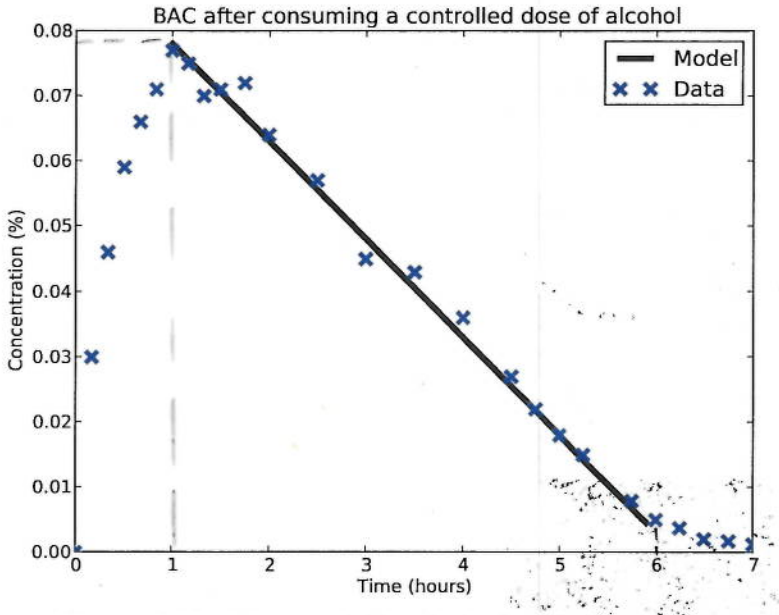


Figure 8.13: Measured blood alcohol concentrations.

(a) Find an equation for $B(t)$.

$y = mx + c$

$m = \text{slope} = \frac{\text{rise}}{\text{run}} = \frac{0.004 - 0.078}{6 - 1} = -0.015\,\%/\text{hr}$

$B(t) = -0.015t + c \Rightarrow B(1) = 0.078 = -0.015 \times 1 + c$

$\Rightarrow c = 0.078 + 0.015$

$= 0.093$

$\boxed{B(t) = -0.015t + 0.093}$

(b) Find $B'(t)$ (include units).

$-0.015\,\%/\text{hr}$

(c) Interpret, in words, what $B'(t)$ represents.

Clearance rate of alcohol in % per hour

ROC is constant because enzymes are saturated

(continued over)

***Question 8.5.1** (continued)*

(d) Figure 8.14 shows measured BACs after researchers administered four different controlled doses of alcohol to study participants (see [55]). Find the value of B' for each of the concentration graphs, and interpret your answer. What does this mean, and is it consistent with "real life"?

$A: B' = \frac{0.003 - 0.018}{1.5 - 0.5} = \frac{-0.015}{1} = -0.015$ %.hr

$B: B' = \frac{0.007 - 0.042}{3 - 0.5} = \frac{-0.035}{2.5} = -0.014$ %.hr

$C: B' = \frac{0.007 - 0.064}{4.8 - 0.9} = \frac{-0.057}{3.9} = -0.015$ %.hr

$D: B' = \frac{0.006 - 0.078}{6 - 1} = \frac{-0.072}{5} = -0.014$ %.hr

Same clearance rate no matter how much you drink!

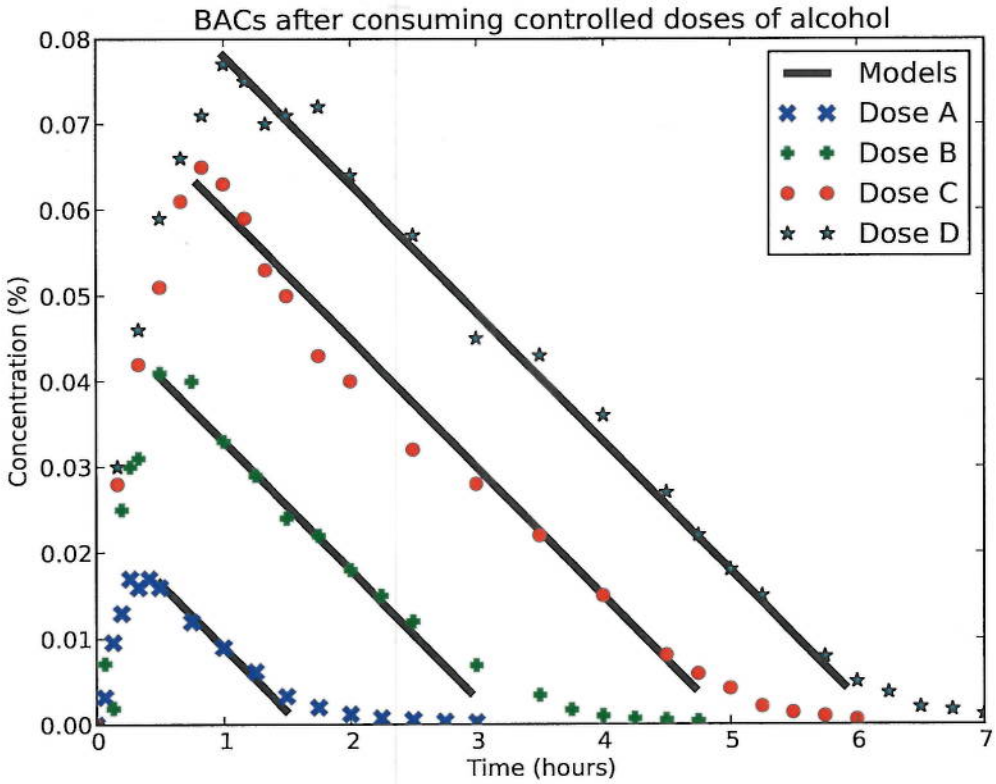


Figure 8.14: Measured BACs after administration of four different controlled doses of alcohol.

Question 8.5.2

In a recent experiment for SCIE1000, a commercial breathalyser was used to monitor the BACs of two females who each consumed approximately 4 standard drinks (38 g of alcohol) over an hour. Figure 8.15 shows the experimental data.

Photo 8.5: Breathalyser. (Source: MG.)

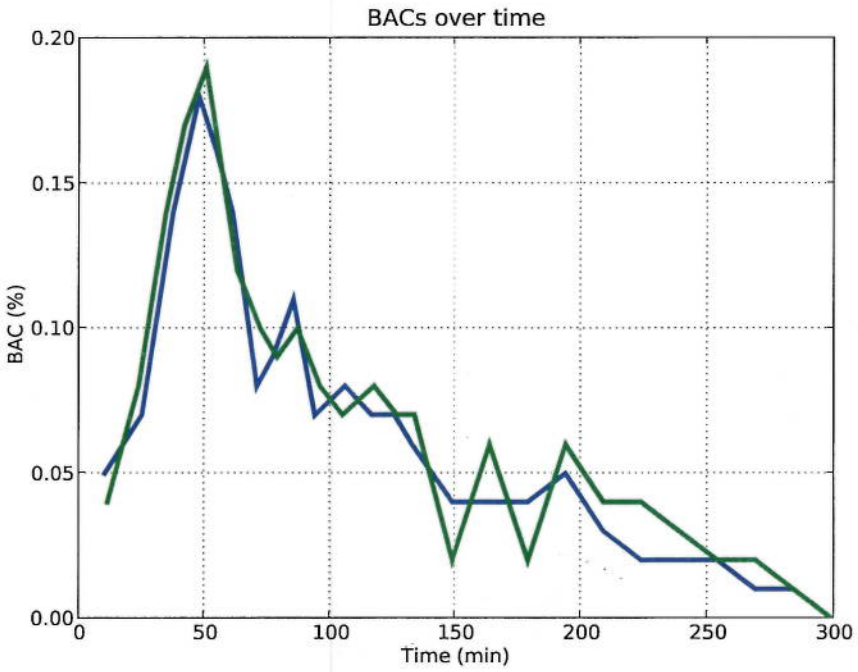


Figure 8.15: Measured blood alcohol concentrations.

Comment on the graphs in Figure 8.15, highlighting both the expected and unexpected features of the shapes of the graphs.

- Unexpected early decline after peak
- Breath alcohol different to BAC.

End of Case Study 20: Whisky (back to BAC).

Case Study 21: **Methamatics**

- Unlike alcohol, which is typically consumed in large volumes, the usual dose sizes of most other drugs are much lower. Thus, most drugs do not saturate metabolic enzymes, so are metabolised by the body at a rate proportional to the current concentration of the drug in the bloodstream (called *first-order* reactions in Chemistry).
- Hence, the typical blood concentration functions of such drugs are based on exponentials, and it is useful to talk about their half-lives. Figure 8.16 shows the approximate half-lives of some substances.

drug	$t_{1/2}$	drug	$t_{1/2}$
caffeine (coffee)	5 h	nicotine (cigarettes)	2 h
codeine (analgesic)	3 h	progesterone (oral contraceptive)	40 h
adrenaline	2 min	3,4-Methylenedioxymethamphetamine (ecstasy)	8 h
testosterone	3 h	γ-Hydroxybutyric acid (GHB)	40 min
diacetylmorphine (heroin)	8 min	morphine	2.5 h
sertraline (Zoloft)	26 h	sildenafil (Viagra)	3 h

Figure 8.16: Some common drugs and hormones, and their approximate half-lives.

Question 8.5.3

The synthetic central nervous system stimulant methamphetamine (also known as 'ice', 'speed' and 'meth') is a relatively common recreational drug that is typically smoked, 'snorted', injected, swallowed orally or taken as a rectal suppository. It also has approved medical uses, for example in the treatment of attention-deficit hyperactivity disorder, ADHD.

Methamphetamine causes the release of chemicals such as *dopamine* and *serotonin* in the brain, leading to feelings of intense excitement and euphoria, intensified emotions, increased energy, reduced appetite, increased libido and a feeling of extreme well-being. The 'high' can last from 2 to 12 hours. It is also highly addictive. Side effects include depression, psychosis, extreme risk taking, and damage to almost every major organ system in the body.

(continued over)

Question 8.5.3 (continued)

Methamphetamine can be synthesised from *pseudoephedrine*, which is a common nasal decongestant in medications for allergies and colds. Australia now heavily regulates the sale of such preparations to prevent their use in the production of methamphetamine.

In [6], researchers administered methamphetamine to study participants, and recorded subsequent blood concentrations of the drug in ng/mL. A model of the concentration during the elimination phase for one participant at time $t > 3$ hours after dosing is given by $M(t) = 47e^{-0.06t}$; see Figure 8.17.

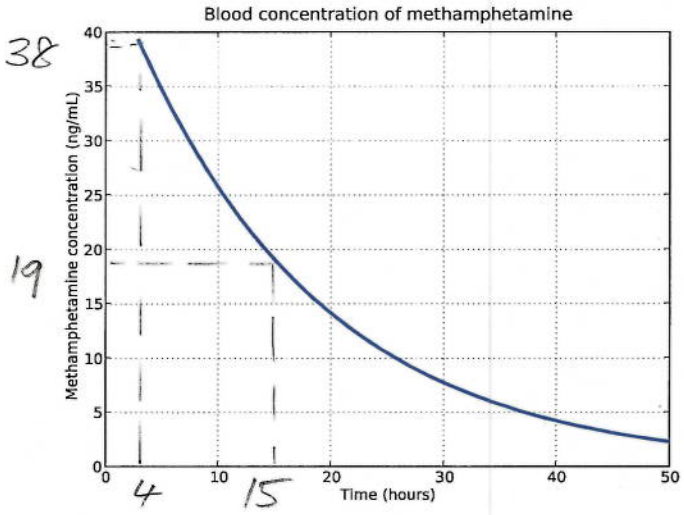


Figure 8.17: Modelled blood concentration of methamphetamine during the elimination phase.

(a) Find the units of $M'(t)$.

$M'(t) = \frac{\text{rise}}{\text{run}} = \text{ng/ml/hr}$

(b) Find the half-life of methamphetamine for this person.

From graph: $t_{\frac{1}{2}} = 15 - 4 = 11$ hrs

Let current point be t_0

From equation: $M(t) = 47e^{-0.06t}$

$$\frac{M_{new}}{M_{old}} = \frac{1}{2} = \frac{47e^{-0.06(t_0 + t_{\frac{1}{2}})}}{47e^{-0.06t_0}} = e^{-0.06t_{\frac{1}{2}}}$$

$$\Rightarrow \ln\left(\tfrac{1}{2}\right) = -0.06t_{\frac{1}{2}} \Rightarrow t_{\frac{1}{2}} = \frac{\ln\left(\frac{1}{2}\right)}{-0.06} = 11.55 \text{ hrs}$$

End of Case Study 21: Methamatics.

Case Study 22: **Wild, wild women**

Photo 8.6: A pair of paltry testes, before and after dissection. (Source: PA. (Just the photo!))

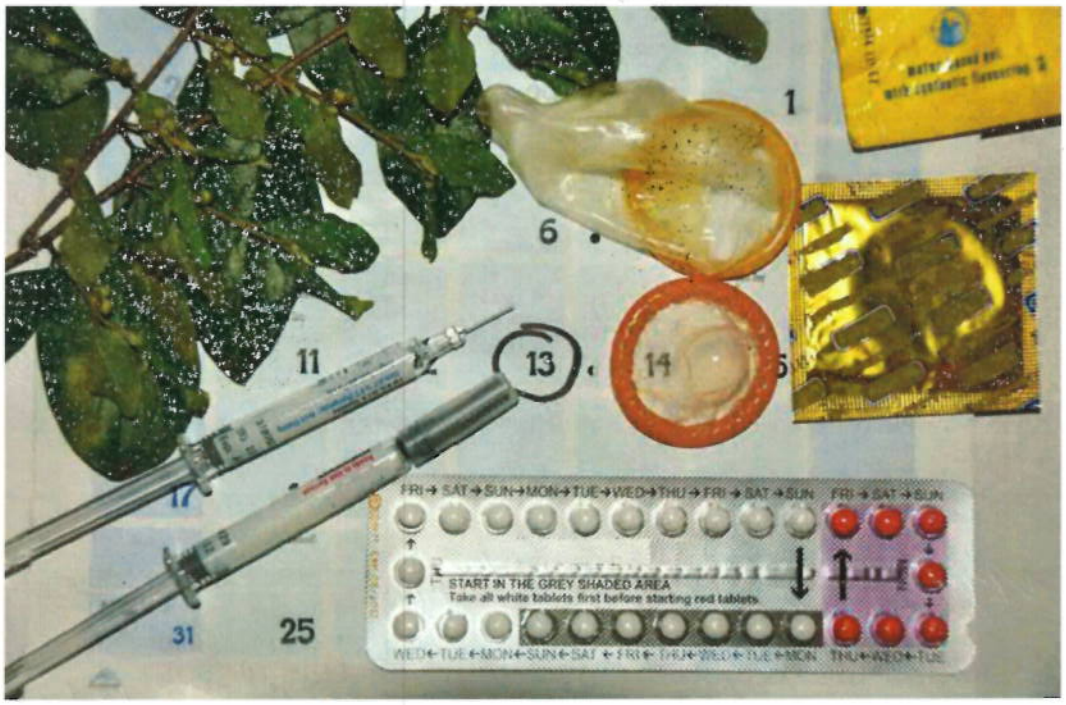


Photo 8.7: Various types of contraceptive including: oral contraceptive, condoms, injected contraceptives and traditional herbal methods. (Source: PA.)

- Each of the many different methods of contraception has advantages and disadvantages.
- Figure 8.18 compares the effectiveness of various methods of contraception, based on data given in [11].

Method	Typical use	Ideal use	1-year
Chance	85	85	
Spermicides	26	6	40
Periodic Abstinence	25	1 – 9	63
Cap			
Parous Women	40	26	42
Nulliparous Women	20	9	56
Sponge			
Parous Women	40	20	42
Nulliparous Women	20	9	56
Diaphragm	20	6	56
Withdrawal	19	4	
Condom	14	3	61
Oral pill	5	0.1	71
IUD	0.1 – 2.0	0.1 – 1.5	80
Depo-Provera IM 150 mg	0.3	0.3	70
Female Sterilisation	0.5	0.5	100
Male Sterilisation	0.15	0.10	100

Figure 8.18: The expected percentage of women who will experience an unintended pregnancy when using various methods of contraception for a year, through either typical use or ideal (very careful) use. Also shown is the average percentage of women continuing to use that method of contraception after one year.

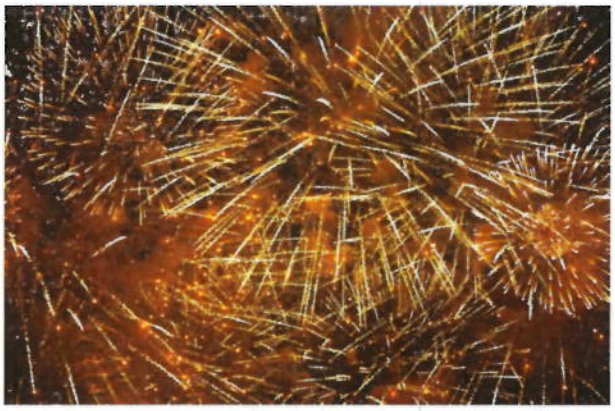


Photo 8.8: Fireworks, Hong Kong. (Source: PA.)

- Depo-subQ Provera 104 is a long-term female contraceptive administered as an injection every 12 – 13 weeks.
- The active ingredient in a standard 0.65 mL dose is 104 mg of the artificial female hormone medroxyprogesterone acetate (MPA), which is similar to progesterone.

- The contraceptive works by inducing changes to the female reproductive system, inhibiting egg release and creating a hostile environment to sperm.
- It is 99.7% effective, which is very high when compared to many other forms of contraception.
- Commonly quoted benefits are convenience and reliability.
- Studies have identified some side effects, including breakthrough bleeding, reduced libido, weight gain and potentially, reduced bone density.

***Example** 8.5.4*

Figure 8.19 shows some pharmacokinetic parameters of MPA after a single subcutaneous injection of Depo-SubQ Provera 104 in healthy women. The data are based on results in [11], with a sample size of $n = 42$ women.

	C_{max} (ng/mL)	t_{max} (day)	C_{91} (ng/mL)	AUC_{0-91} (ng day/mL)	$AUC_{0-\infty}$ (ng day/mL)	$t_{1/2}$ (day)
Mean	1.56	8.8	0.402	66.98	92.84	43
Min	0.53	2.0	0.133	20.63	31.36	16
Max	3.08	80.0	0.733	139.79	162.29	114

Figure 8.19: Pharmacokinetic parameters of MPA.

In Figure 8.19:

- C_{max} = peak blood concentration;
- t_{max} = the time at which C_{max} occurs;
- C_{91} = blood concentration at 91 days;
- AUC_{0-91} = the area under the concentration-time curve over 91 days;
- $AUC_{0-\infty}$ = the area under the concentration-time curve over an indefinite time period; and
- $t_{1/2}$ = half-life of MPA.

Example 8.5.5

A patient receives an injection of Depo-subQ Provera 104. After the dose, the concentration $C(t)$ of MPA in her blood (in ng/mL at time t in days) is modelled by $C(t) = 1.4t^{0.15}e^{-0.02t}$. Figure 8.20 shows the graph of $C(t)$.

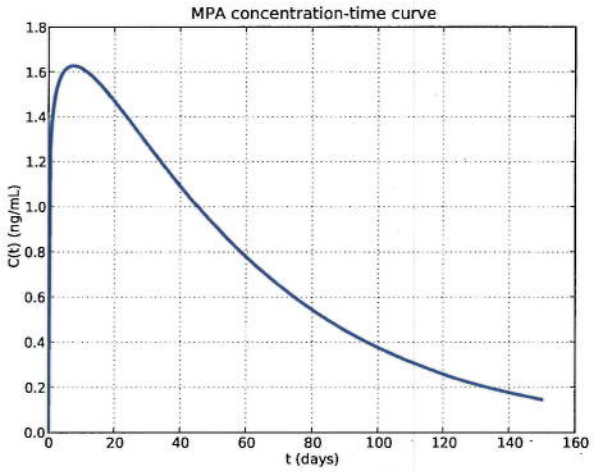


Figure 8.20: A model of the blood concentration of MPA.

- Now we will demonstrate how to use derivatives and Newton's method to determine the timing of a follow-up injection. This process is ideally suited to being performed by a computer.
- Reliable contraception ceases when $C(t) = 0.3$ ng/mL, so $C(t) - 0.3 = 0$.
- Hence the function $f(t)$ for Newton's method is: $f(t) = 1.4t^{0.15}e^{-0.02t} - 0.3$.
- The derivative is: $f'(t) = 1.4e^{-0.02t}\left(0.15t^{-0.85} - 0.02t^{0.15}\right)$ Product rule
- We can use $t_0 = 50$ as the initial estimate for the solution. (Remember that you usually have a choice of many different initial values.)
- The following program was run, and on the fourth step gave: $t_4 \approx 112.440$.
- Further steps did not give significant changes, so the blood concentration of MPA decreased to 0.3 ng/mL at around 112 days, or about 16 weeks.
- For reference, the time recommended by the manufacturer for follow-up injections is 12–13 weeks, which provides a reasonable safety margin.

Program 8.1: Using Newton's method for contraception

```
# Use Newton's method to find the follow-up time for a
# contraceptive injection.
from pylab import *

# Define the function and its derivative.
def func(t):
    return 1.4 * pow(t,0.15) * exp(-0.02*t) - 0.3

def funcDash(t):
    val1 = 1.4 * exp(-0.02*t)
    val2 = 0.15 * pow(t,-0.85) - 0.02*pow(t,0.15)
    return(val1 * val2)

# Initialise variables
val = eval(input("What is the initial estimate? "))

# Loop through steps of Newton's method.
i=0
while abs (func(val)) > 0.0001:
    val = val - func(val) / funcDash(val)
    i = i+1
    print("Step ",i,":",round(val,3))

print("Estimated time is:",round(val,3),"days")
```

System

Input

Newton's Method

System.

Output.

Here is the output from running the above program:

```
What is the initial estimate? 50
Step  1 : 89.769
Step  2 : 108.467
Step  3 : 112.302
Step  4 : 112.44
Estimated time is: 112.44 days
```

End of Case Study 22: Wild, wild women.

Case Study 23: **Drink deriving**

- In practice (particularly in legal cases), models of BAC use the *Widmark formula*, developed in 1932. The equation is:

$$B = \frac{A}{rM} \times 100\% - Vt$$

where B is the BAC at time t since commencing drinking, A is the amount of alcohol consumed in grams, V is the rate at which the body eliminates alcohol measured in % per time period, M is the body mass in grams and the *Widmark factor* r estimates the proportion of body mass that is water.

- The precise value of r depends on factors such as gender, age and percentage body fat. Reasonable estimates are $r \approx 0.7$ for adult males and $r \approx 0.6$ for adult females. The typical value for V is 0.015 % hr^{-1}.

Question 8.5.6

(a) What is the physical meaning of the term rM in the Widmark formula?

rM = H_2O content of body (into which alcohol mixes)

Alcohol is soluble in H_2O.

Volume of distribution

(b) Why do females typically have a lower value of r than males?

♀ typically have more fat than ♂. And fat has less H_2O than muscle

(c) Verify that the units in the Widmark formula are consistent.

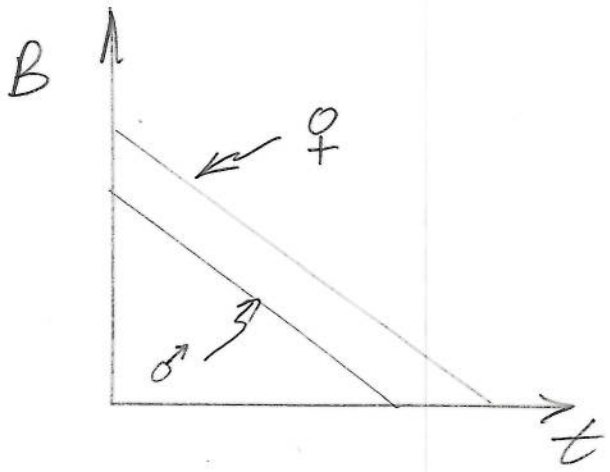
$B = BAC = \%_0 = \frac{g}{g} \times \%_0 - \%_0/\text{hr} \times \text{hr}$

$\%_0 = \%_0$.

(continued over)

***Question** 8.5.6 (continued)*

(d) The Widmark formula is: $B = \dfrac{A}{rM} \times 100\% - Vt$. Sketch rough graphs of B for "typical" men and women.



(e) Use the Widmark formula to justify Australian government guidelines that to remain below the legal driving BAC limit, within the first hour "men should drink at most two drinks and women at most one".

Clearance rate for ♂ × ♀ = −0.015%.hr

What is B_{max} from equation for B?

✓ ♂ 2 drinks, 80kg, $B_{max} = \dfrac{A}{rM} = \dfrac{20}{0.7 \times 80000} \times 100\% = 0.036\%$

× ♀ 2 drinks, 60kg, $B_{max} = \dfrac{A}{rM} = \dfrac{20}{0.6 \times 60000} \times 100\% = 0.056\%$

✓ ♀ 1 drink, 60kg, $B_{max} = \dfrac{A}{rM} = \dfrac{10}{0.6 \times 60000} \times 100\% = 0.028\%$

What is B_{max} for ♂ 3 drinks, 80kg?

Question 8.5.7

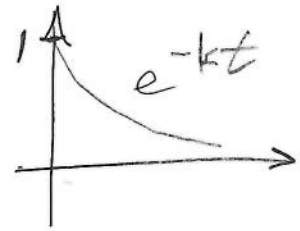
In the Widmark formula, the absorption term assumes that the body absorbs alcohol **immediately** after consumption. The following variation, from [42], takes into account absorption time.

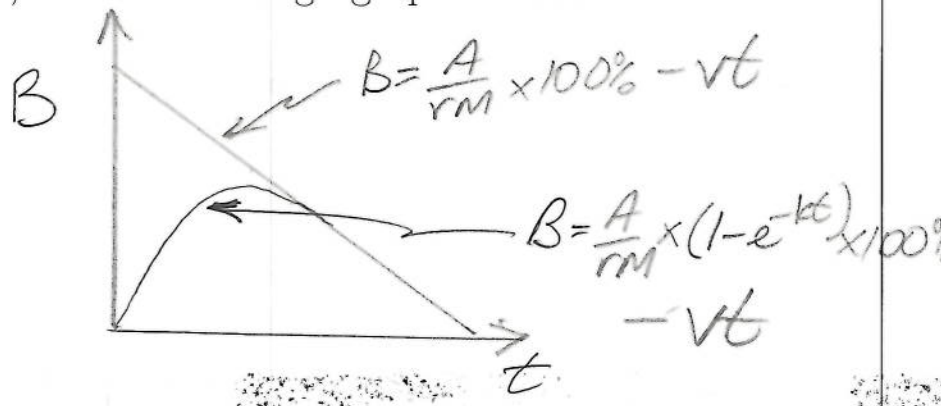
$$B = \frac{A}{rM} \times \left(1 - e^{-kt}\right) \times 100\% - Vt$$

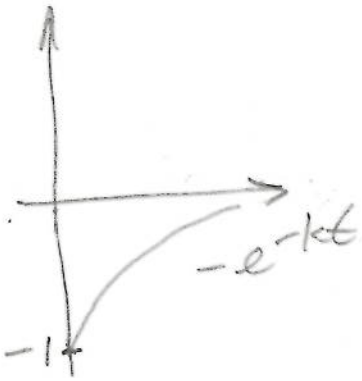
where k is the rate at which the body absorbs alcohol.

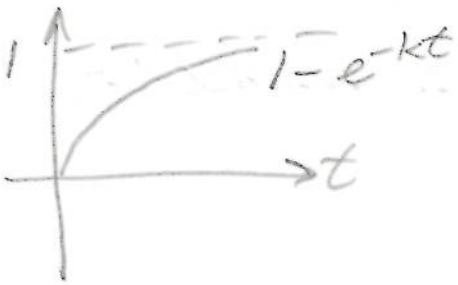
(a) The "standard" Widmark formula is: $B = \frac{A}{rM} \times 100\% - Vt$.

Reconcile the two versions, and sketch rough graphs of each.









- Standard model assumes immediate uptake
- Modified version considers A + D
- Like a surge function sort of but linear decline (alcohol overwhelmes enzymes)

(continued over)

Question 8.5.7 *(continued)*

(b) Recall that $B = \dfrac{A}{rM} \times \left(1 - e^{-kt}\right) \times 100\% - Vt$.

If t is measured in hours, what are the units of k?

e^{-kt} can't have units so k units are $1/hrs$

(c) What factors could influence the value of k for:

(i) a given person, at different times? food (competition for AD)

(ii) different people? Quite consistent among people

Some races don't have the enzymes

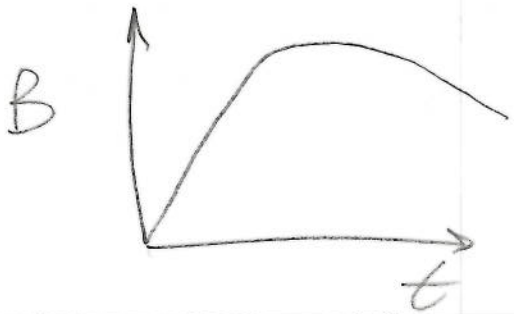
(d) Let t_{max} be the time at which BAC reaches its maximum value B_{max}.
For larger values of k, will t_{max} be larger, smaller or unchanged? Why?

$\therefore$ $k \uparrow$ $t_{max} \downarrow$
faster adsorption

(e) For larger values of k, will B_{max} be larger, smaller or unchanged? Why?

$k \uparrow$ $B_{max} \uparrow$ Because faster absorption means not as much time for M & E to occur

(f) Explain briefly how to find the values of t_{max} and B_{max}.



when $B' = 0$

(continued over)

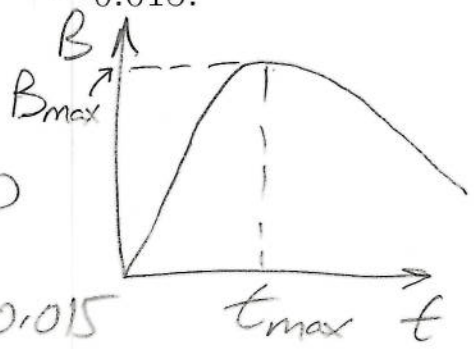
***Question 8.5.7** (continued)*

(g) When consuming alcohol with food in the stomach, $k \approx 2.3$, but when the stomach contains no food, $k \approx 6$. When a "typical" man of mass 80 kg consumes 4 standard drinks with food in his stomach, we have

$$B(t) \approx 0.0714(1 - e^{-2.3t}) - 0.015t$$

$$B'(t) \approx 0.164e^{-2.3t} - 0.015.$$

Find t_{max} and B_{max} for this man.



t_{max} & B_{max} when $B'(t)=0$

$B'(t_{max})=0=0.164e^{-2.3t}-0.015$

$\Rightarrow \frac{0.015}{0.164} = e^{-2.3t_{max}} \Rightarrow \ln(0.09) = -2.3t_{max}$

$\Rightarrow t_{max} = \frac{-2.397}{-2.3} = 1.04$ hrs.

$B(t_{max}) = 0.0714(1-e^{-2.3\times1.04}) - 0.015\times1.04$

$= 0.05\%$.

(h) If the same man consumes the same amount of alcohol, but on an empty stomach, we have $t_{max} \approx 0.56$ hours and $B_{max} \approx 0.0605$ %. Compare this with your answer to Part (g), and relate this to Parts (d) and (e).

$k\uparrow \Rightarrow t_{max}\downarrow$ (d)

$k\uparrow \Rightarrow B_{max}\uparrow$ (e)

For $k=6$ without food, $B_{max}=0.06$ (check)

$\therefore$ over the limit!

Now we can develop a computer model of food consumption and BACs.

Program specifications: Write a program that plots BAC curves up to 5 hours after alcohol consumption on both a full and empty stomach, for men or women of varying masses and for various levels of alcohol consumption.

Program 8.2: BACs and food consumption

```
# Program to compare BACs for men and women of varying masses
# and levels of alcohol consumption, on full and empty stomachs
.
from pylab import *

alcohol = eval(input("How many g of pure alcohol consumed? "))
weight = eval(input("Person's mass (in kg)? "))
gender = eval(input("Type 1 if male, anything else female: "))
if gender == 1:
    mult = 100 * alcohol / (0.7*weight*1000)
else:
    mult = 100 * alcohol / (0.6*weight*1000)
times = arange(0,5.1,0.1)

# Calculate the BAC at each time.
BACfull = mult * (1 - exp(-times * 2.3)) - 0.015 * times
BACempt = mult * (1 - exp(-times *  6)) - 0.015 * times
i = 0

# BAC cannot be negative.
while i<size(times):
    BACfull[i] = max(BACfull[i], 0)
    BACempt[i] = max(BACempt[i], 0)
    i = i+1

plot(times,BACfull,"k—",linewidth=4, label="Full stomach")
plot(times,BACempt,"b–",linewidth=4, label="Empty stomach")
grid(True)
xlabel("Time (hours)")
ylabel("BAC (%)")
title("BAC curve for a full stomach versus an empty stomach")
legend()
show()
```

Input

System

Output

Figure 8.21 shows the output from running the above program for an 80 kg male consuming four standard drinks:

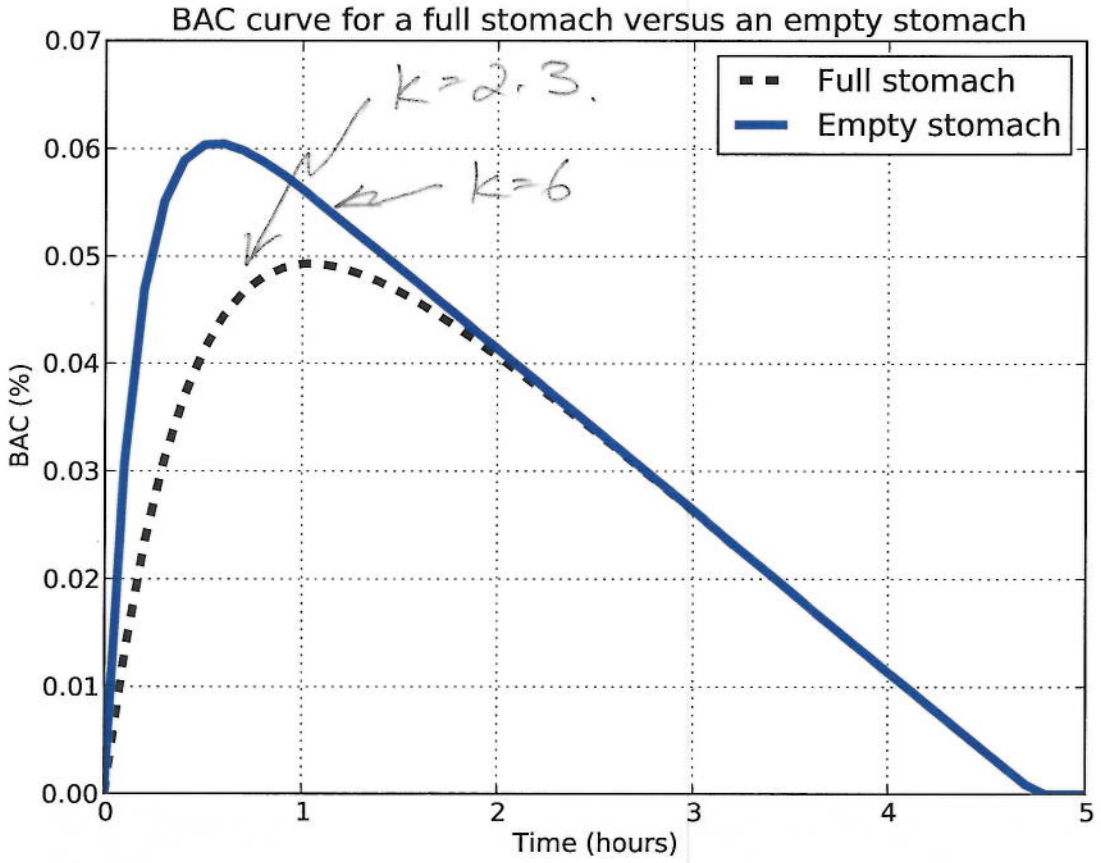


Figure 8.21: Predicted BACs when consuming alcohol on a full stomach compared to an empty stomach.

Question 8.5.8

Briefly compare the graph in Figure 8.21 with your answers to Question 8.5.7. What are some of the physical implications of the graph?

$AUC_{Empty} > AUC_{Full}$

Thus greater exposure if drink on an empty stomach

$\therefore$ eat and drink!

End of Case Study 23: Drink deriving.

Chapter 9: **You, me and AUC**

Lily the Pink she turned to drink
she filled up with paraffin inside
and despite her medicinal compound
sadly pickled-Lily died.

Up to heaven her soul ascended
all the church bells they did ring
she took with her medicinal compound
hark the herald angels sing.

We'll drink a drink a drink
to Lily the Pink the Pink the Pink
the saviour of the human race
for she invented medicinal compound
most efficacious in every case.

Artist: *The Scaffold.* Song: *Lily the Pink.*

Image 9.1: *The Drunks* (1629), Diego Velazquez (1599 – 1660), Museo del Prado, Madrid. (Source: en.wikipedia.org)

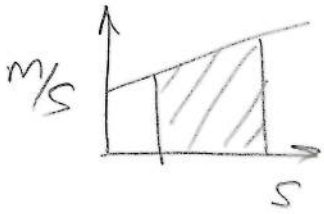
9.1 Areas under curves

- Given a graph, the *area under the curve* or AUC of that graph is the area bounded by that curve, the x-axis and two points on the x-axis.
- The AUC often has a useful physical meaning, which depends on what is being graphed.

Question 9.1.1

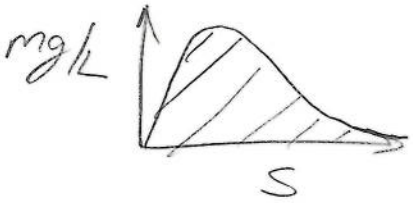
What is meant by the AUC in each of the following.

(a) A graph of velocity versus time.



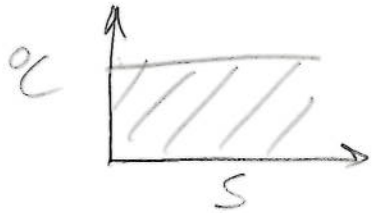
$AUC = m/s \times s = m$ = displacement or distance

(b) A graph of chlorine concentration in water versus time taken for treatment to make the water be safe for drinking.



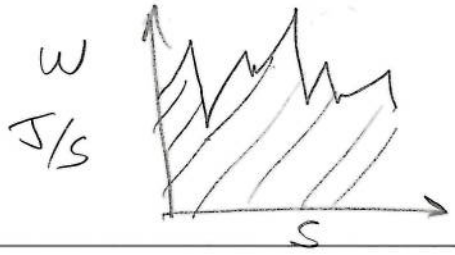
$AUC = mg/L \times s$ = Total exposure to kill pathogens

(c) A graph of cooking temperature versus time required to prevent the parasitic disease *Trichinosis* when preparing pork for consumption.



$AUC = °C.s$ = Total exposure to heat to kill Trichinosis

(d) A graph of electricity consumption in a household versus time.



$AUC = W.s = J/s \times s = J$
= Total energy used.
= kW hrs.

- In pharmacology, the AUC is particularly important.

Question 9.1.2

Figure 9.1 shows a graph with a line fitted to some measured blood alcohol concentrations (see [55]).

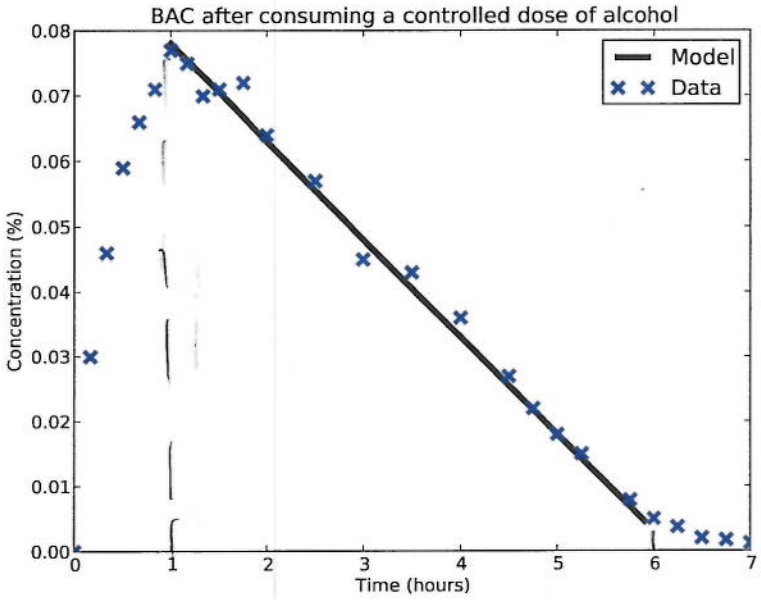


Figure 9.1: A graph of measured BACs.

(a) What are the units of the AUC in the graph?

AUC = % hrs

(b) What does the AUC represent and why is it significant?

Total exposure to alcohol (a toxin)

(c) Calculate the approximate AUC between $t = 1$ hr and $t = 6$ hr.

$$AUC = \frac{L \times B}{2} = \frac{0.08 \times 5}{2} = 0.2 \text{ \% hrs}$$

- Areas under curves are so useful that they have a special name and notation.

AUC and Definite integrals

Given a function $f(x)$, the AUC from the point $x = a$ to the point $x = b$ is called the *definite integral of $f(x)$ from a to b*, written as

$$\int_a^b f(x)\,dx.$$

- There are two common ways of calculating AUCs.
- First, recall from school that if the function $f(x)$ is known, then the *Fundamental Theorem of Calculus* gives an "easy" mathematical approach for finding the AUC between two points a and b on the x-axis:
  - Find an *antiderivative* or *integral* of $f(x)$, say $F(x)$.
  - Substitute the value b into $F(x)$.
  - Substitute the value a into $F(x)$.
  - Subtract the second value from the first one.
  - The answer gives the required AUC.
- More often, AUCs are used in practical applications in which the only available information is a collection of measured data values, and the function $f(x)$ is **not** known.
- In such cases, AUCs are estimated approximately, by summing the areas of geometric shapes of "narrow" width, such as rectangles (called *Riemann sums*), or *trapezoids* (called the *trapezoid rule*).

Question 9.1.3

Give some advantages and disadvantages of each way of finding AUCs.

AUC $\int_a^b f(x)\,dx$ — Exact but need function. Must be integrable

AUC Geometric — Estimate (from data points) but don't need a function.

Question 9.1.4

In Question 9.1.2 Part (c) we used the area of a triangle to calculate the AUC between $t = 1$ hr and $t = 6$ hr for the line shown in Figure 9.2.

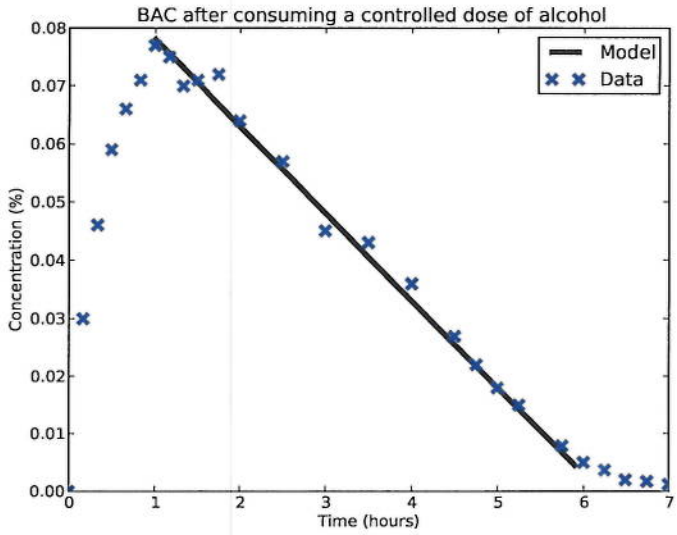


Figure 9.2: Blood alcohol concentrations.

(a) Rewrite Question 9.1.2 Part (c) as a definite integral.

$\int_1^6 B(t)\,dt$

(b) Evaluate Part (a) above using antiderivatives. (Hints: the equation of the line is $B = 0.093 - 0.015t$. Also, $\int (a - bt)dt = at - bt^2/2 + C$.)

$\int_1^6 B(t)\,dt = \int_1^6 0.093 - 0.015t = \left[0.093t - \frac{0.015}{2}t^2 + C\right]_1^6$

$= (0.093 \times 6 - \frac{0.015 \times 6^2}{2} + C) - (0.093 \times 1 - \frac{0.015 \times 1^2}{2} + C)$

$= (0.558 - 0.27) - (0.093 - 0.0075)$

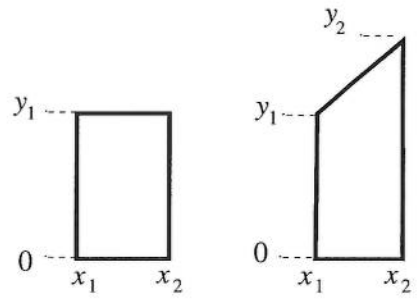
$= 0.2025\%$ hrs

(c) Compare your answer to Part (b) above with Question 9.1.2 Part (c).

Very close!

Question 9.1.5

(a) Derive expressions for the areas of the rectangle and the trapezoid.



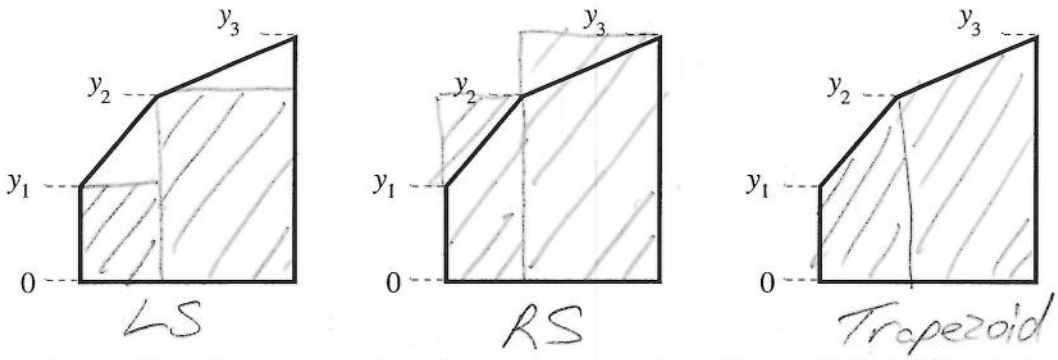
$A_{rect} = L \times B = (x_2 - x_1) \times y_1$ $\quad \frac{B \times H}{2}$

$L \times B$

$A_{trap} = (x_2 - x_1) \times y_1 + (x_2 - x_1)\frac{(y_2 - y_1)}{2}$

OR $(x_2 - x_1)\left(y_1 + \frac{y_2}{2} - \frac{y_1}{2}\right) = (x_2 - x_1)\left(\frac{y_1 + y_2}{2}\right)$

(b) Show how to find the approximate area of the following shape in three ways, using rectangles in two different ways, and using trapezoids.



(c) Comment on the three approaches you used in Part (b). Which is likely to be most accurate? How do the two methods using rectangles relate to the method using trapezoids?

LS < RS, Trapezoid perfect.

When $f(x)\uparrow$, LS underestimates & RS overestimates

" " $\downarrow$, " " over " " & RS under "

What is the average of the LS & RS?

Question 9.1.6

Let $N(t)$ be an unknown function representing the blood nicotine concentration of a person after smoking a cigarette. Figure 9.3 shows some concentrations, measured experimentally (see [3]).

(continued over)

Question 9.1.6 (continued)

t (min.)	0	3	6	10	14	20	35	65	95
Value (ng/mL)	5	11	15.4	13.4	12.8	11.3	9.8	8	7

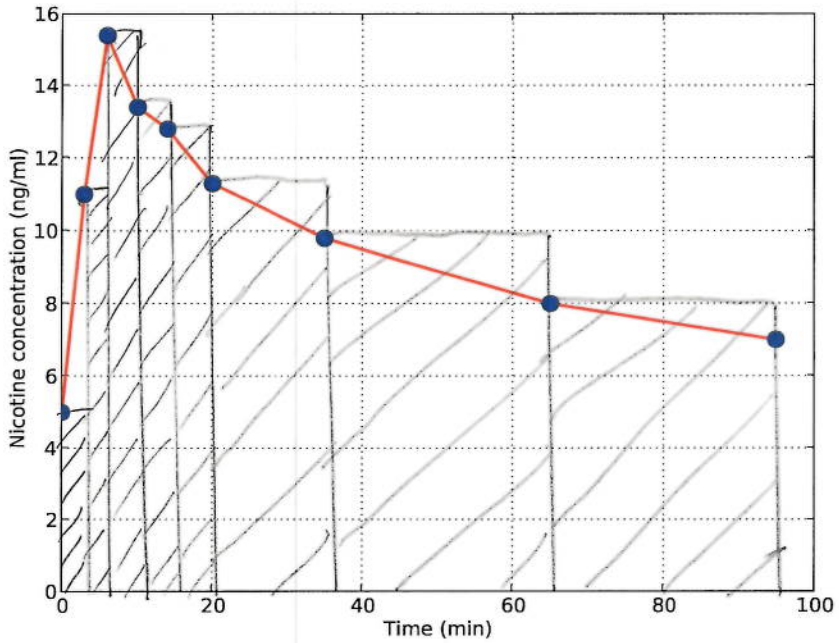


Figure 9.3: Measured blood nicotine concentrations after smoking.

The following calculations use areas of rectangles to estimate partially the AUC of the nicotine concentration curve. Complete the calculations.

LS

$$
\begin{aligned}
(3-0) \times 5 &= 15 \\
(6-3) \times 11 &= 33 \\
(10-6) \times 15.4 &= 61.6 \\
(14-10) \times 13.4 &= 53.6 \\
(20-14) \times 12.8 &= 76.8
\end{aligned}
$$

$(35-20) \times 11.3 = 169.5$

$(65-35) \times 9.8 = 294$

$(95-65) \times 8 = 240$

$\sum$ 943.5 ng min/ml

Question 9.1.7

Repeat Question 9.1.6 but instead use the trapezoid rule.

t (min.)	0	3	6	10	14	20	35	65	95
Value (ng/mL)	5	11	15.4	13.4	12.8	11.3	9.8	8	7

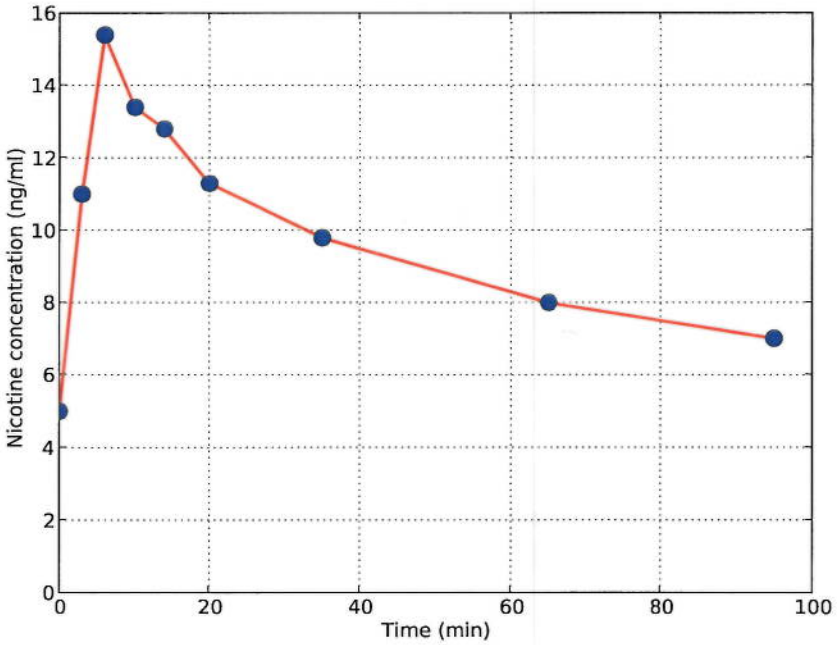


Figure 9.4: Measured blood nicotine concentrations after smoking.

$$
\begin{aligned}
(3-0) &\times (5+11)/2 &&= 24\\
(6-3) &\times (11+15.4)/2 &&= 39.6\\
(10-6) &\times (15.4+13.4)/2 &&= 57.6\\
(14-10) &\times (13.4+12.8)/2 &&= 52.4\\
(20-14) &\times (12.8+11.3)/2 &&= 72.3
\end{aligned}
$$

$(35-20) \times (9.8+11.3)/2 = 158.25$

$(65-35) \times (8+9.8)/2 = 267$

$(95-65) \times (7+8)/2 = 225$

$\Sigma = 896.15$ ng min/ml

We can perform the previous calculations more efficiently using a program.

Program specifications: Write a Python program that estimates the AUC for $N(t)$ using rectangles or the trapezoid rule. The program must output the total AUC and draw a graph showing the shapes used in the sums.

Program 9.1: AUC for nicotine

```
# Use rectangles or trapezoids to estimate AUC for nicotine.
from pylab import *

# Initialise variables
type = eval(input("Type 1 for rectangles, 2 for trapezoids: "))
t = array([0, 3, 6, 10, 14, 20, 35, 65, 95])
concs = array([5, 11, 15.4, 13.4, 12.8, 11.3, 9.8, 8, 7])
area = 0

# Sum the areas of each shape
i = 1
while i < size(t):
    width = t[i] - t[i-1]
    if type == 1:
        height = concs[i-1]
        shapeX = array([t[i-1], t[i-1], t[i], t[i]])
        shapeY = array([0, height, height, 0])
    else:
        height = (concs[i-1] + concs[i])/2
        shapeX = array([t[i-1], t[i-1], t[i], t[i]])
        shapeY = array([0, concs[i-1], concs[i], 0])
    area = area + height * width
# Plot each shape
    plot(shapeX, shapeY, "k-",linewidth=2)
    i = i + 1

# Give the output.
print("The estimated AUC is",area,"ng min / mL")
plot(t, concs, "bo", markersize=8)
xlabel("Time (mins)")
ylabel("Nicotine concentration (ng/mL)")
if type == 1:
    title("Blood concentration of nicotine (rectangles)")
else:
    title("Blood concentration of nicotine (trapezoids)")
show()
```

Input

LHS

Trapezoid rule

System

Output

Here is the output from running the above program twice:

```
Type 1 for rectangles, 2 for trapezoids: 1
The estimated AUC is 943.5 ng min / mL

Type 1 for rectangles, 2 for trapezoids: 2
The estimated AUC is 896.15 ng min / mL
```

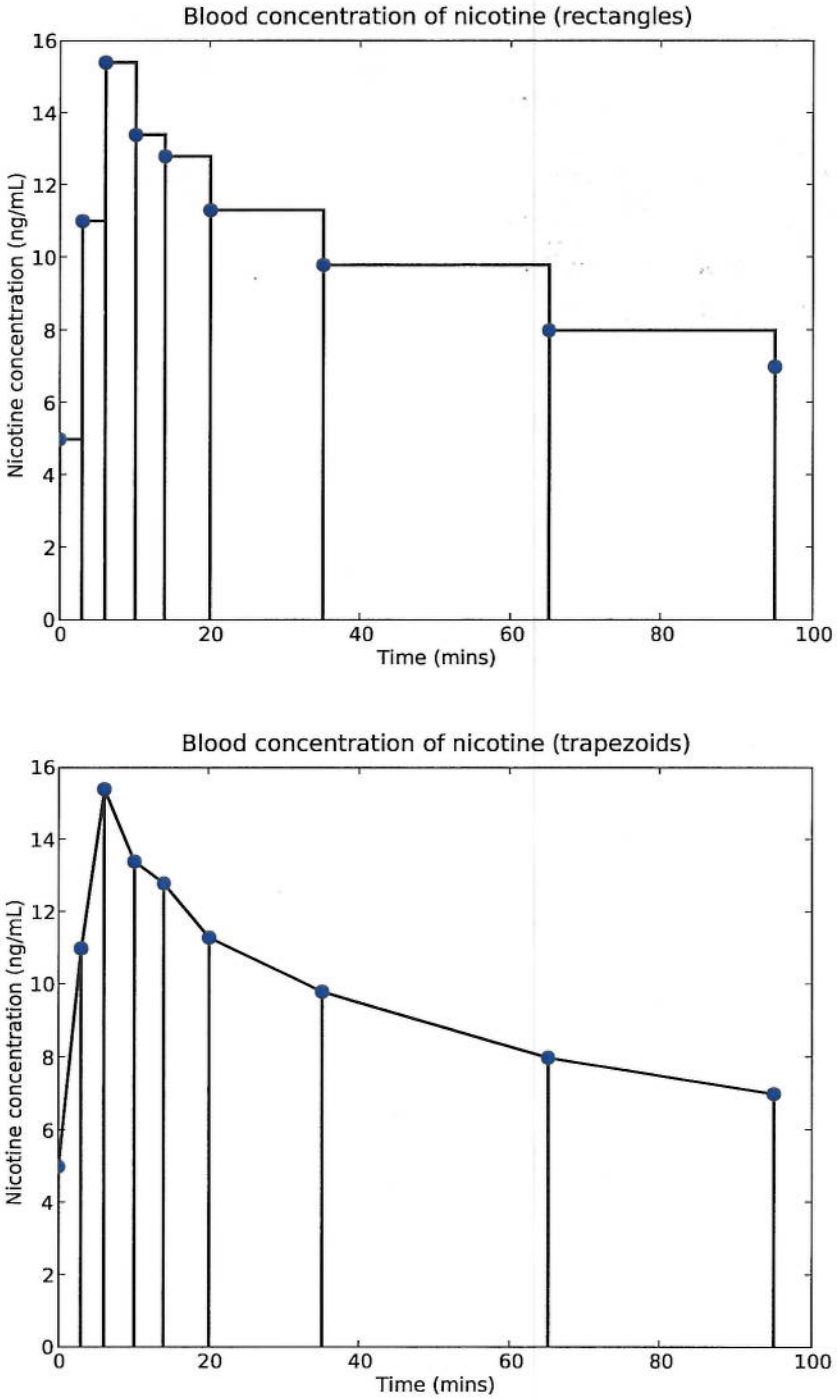


Figure 9.5: Program output using rectangles (top) and trapezoids (bottom).

Case Study 24: **Dying for a drink**

Photo 9.1: Left: mellow and yellow. Right: better red than dead. (Source: DM.)

- In addition to the immediate risks associated with alcohol consumption (such as accidents), the risk of many negative long-term health effects is increased by both the frequency and volume of consumption.

Question 9.1.8

The Widmark formula is used to estimate blood alcohol content (BAC); see Question 8.5.6. For a 'typical' 70 kg man drinking n standard drinks (each containing 10 grams of alcohol), his estimated % BAC at time t in hours since commencing drinking is

$$B = \frac{10n}{490} - 0.015t.$$

(a) Find an expression for the time at which his BAC returns to 0.

when $B=0 \Rightarrow 0 = \frac{10n}{490} - 0.015t$

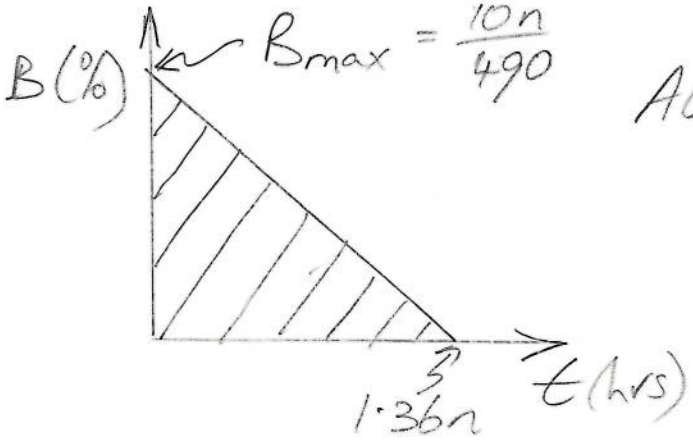
$\Rightarrow t = \frac{10n}{490 \times 0.015}$

$\Rightarrow \boxed{t = 1.36n}$

(continued over)

***Question** 9.1.8 (continued)*

(b) Define the *total exposure to alcohol* E as the AUC of B from $t = 0$ until the BAC reaches 0 again. Find an expression for E for this man.



$$AUC = \frac{1}{2}(B \times H) = \frac{1}{2}\left(1.36n \times \frac{10n}{490}\right)$$

$$= 0.0139n^2$$

(c) Assume that long-term damage to internal organs from consumption of alcohol is proportional to the total exposure to alcohol E (which is simplistic, but not unreasonable). Discuss the impact on E of "one extra drink for the road".

$E \propto n^2$!

So exposure to alcohol goes up with n^2

(continued over)

***Question 9.1.8** (continued)*

(d) A 'typical' man with mass 70 kg consumes two standard drinks every day. A second 'typical' man with the same mass consumes 14 standard drinks once a week, but does not drink at any other time. Estimate the weekly value of E for each man. What are some of the physical ramifications of your answer in relation to binge drinking?

Binge drinking or consistent drinking
– which is worse?

σ_1: Weekly exposure $= 7 \times 0.0139 \times 2^2$
$= 28 \times 0.0139$

σ_2: Weekly exposure $= 1 \times 0.0139 \times 14^2$
$= 196 \times 0.0139$

$\therefore \dfrac{\sigma_2}{\sigma_1} = \dfrac{196}{28} = 7 \times$ more exposure for binge drinking.

(e) For a 'typical' woman of mass 60 kg, $E = 0.0257n^2$. Find the ratio of the values of E for the 'typical' woman and 'typical' man. What does this mean?

$\dfrac{\text{♀ Exposure}}{\text{♂ Exposure}} = \dfrac{0.0257}{0.0139} \approx 1.85$ ∵ less mass + less body H_2O

Now we can develop a computer model for total exposure to alcohol.

Program specifications: Write a Python program that uses the Widmark formula to graph the total exposure to alcohol for a person who consumes from zero to 14 standard drinks.

Program 9.2: Wilful exposure (to alcohol)

```
# Calculate exposure to alcohol.
from pylab import *

mass = eval(input("Enter the person's mass in kg: "))
gender = eval(input("Enter 1 if male, anything else female: "))
drinks = arange(0,15)
if gender == 1:
    water = 1000 * mass * 0.7
else:
    water = 1000 * mass * 0.6
# Estimate E for each number of drinks and plot graph
peakBAC = 1000 * drinks / water
tBAC0 = peakBAC / 0.015
AUC = tBAC0 * peakBAC / 2.0
plot(drinks, AUC, "bo", markersize=8)
grid(True)
xlabel("Number of drinks")
ylabel("Total exposure (% hours)")
title("Total exposure to alcohol")
show()
```

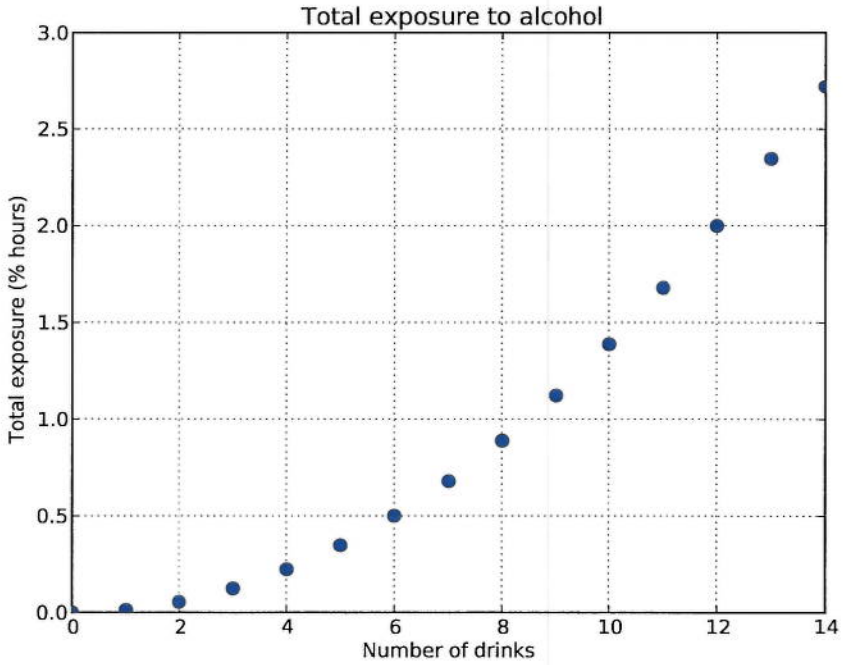


Figure 9.6: Program output showing total exposure to alcohol according to the number of drinks consumed.

End of Case Study 24: Dying for a drink.

9.2 All in the blood

Case Study 25: **Sweet P's**

Photo 9.2: Left: sweet pea. Right: sweet pee. (Source: PA. (Not just the photo!))

- The hormone *insulin* regulates conversion of glucose into usable energy in the body. *Diabetes mellitus* is a group of chronic diseases in which insufficient insulin is produced, or insensitivity to insulin develops. This leads to high levels of blood glucose.
- There are three main types of diabetes: *type* 1 (insulin-dependent diabetes, typically present at birth, representing about 10–15% of all cases), *type* 2 (non-insulin-dependent diabetes, often related to poor lifestyle, 85–90% of all cases) and *gestational diabetes* (developed in 3–8% of pregnancies).
- Typical signs of diabetes include:
  - *polyuria* (excessive urination, often with a sweet taste)
  - *polydipsia* (excessive thirst)
  - *polyphagia* (excessive hunger).
- Once type 1 or type 2 diabetes becomes established, it is usually permanent.
- Diabetes, particularly type 2, is becoming increasingly common in societies with a "western lifestyle". Around 1.7 million Australians have diabetes, although only about half of them are aware of it.

- The Framingham heart study has shown that diabetes significantly reduces life expectancy (by 7.5 years for men aged over 50, and 8.2 years for women).
- Untreated diabetes can cause blindness, kidney failure and cardiovascular disease including blockages in small arteries. Some patients require amputations after blocked peripheral circulation causes the death of soft tissue.
- An *Oral Glucose Tolerance Test* (OGTT) is a common test for diabetes.
- Prior to taking the test, the patient fasts for around 12 hours. During the test, the patient is administered a measured oral dose of glucose, with blood samples taken immediately prior to ingestion of the glucose and at various intervals for 2 hours afterwards.
- The graph in Figure 9.7 compares the measured blood glucose levels for a non-diabetic person with those from a hypothetical diabetic person.

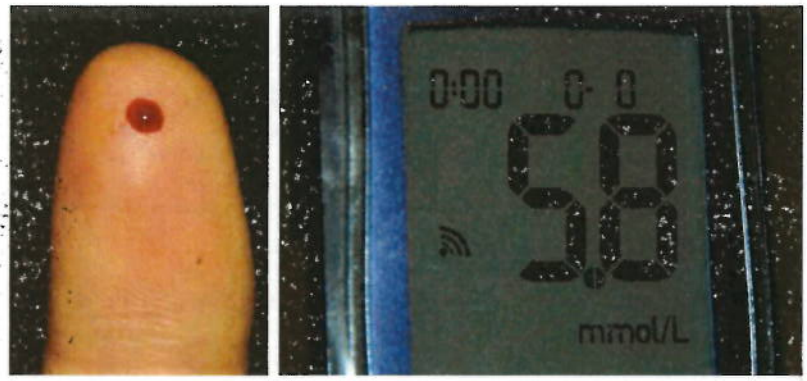


Photo 9.3: Left: bloody finger. Right: measured blood glucose concentration. (Source: PA.)

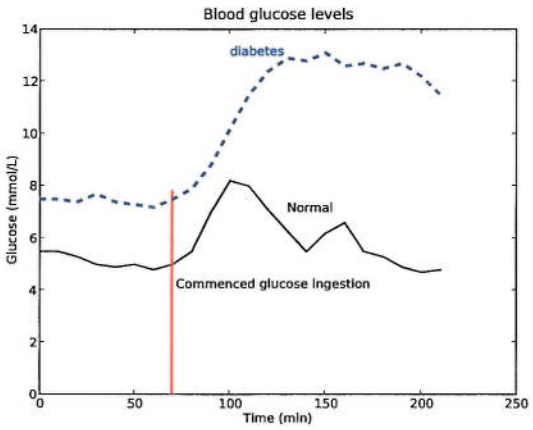


Figure 9.7: Normal blood glucose levels and levels indicative of diabetes.

- Figure 9.8 shows the World Health Organisation guidelines for blood glucose levels indicating various stages of health or disease.

	Blood glucose level (mmol/L)			
time (hr)	Normal	IFG	IGT	DM
$t = 0$	< 6.1	$\geq 6.1, < 7.0$	< 7.0	≥ 7.0
$t = 2$	< 7.8	< 7.8	≥ 7.8	≥ 11.1

Figure 9.8: World Health Organisation guidelines for blood glucose levels as indicators of: Impaired Fasting Glycaemia (IFG); Impaired Glucose Tolerance (IGT or *pre-diabetes*); and Diabetes Mellitus (DM).

Example 9.2.1

Figure 9.9 shows blood glucose level measurements, taken from Peter at ten minute intervals between 6 am and 9:30 am. He had not eaten for 10 hours prior to testing, and at 7:10 am he consumed some food.

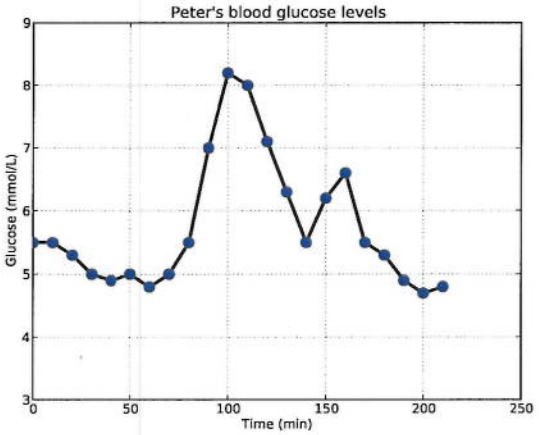
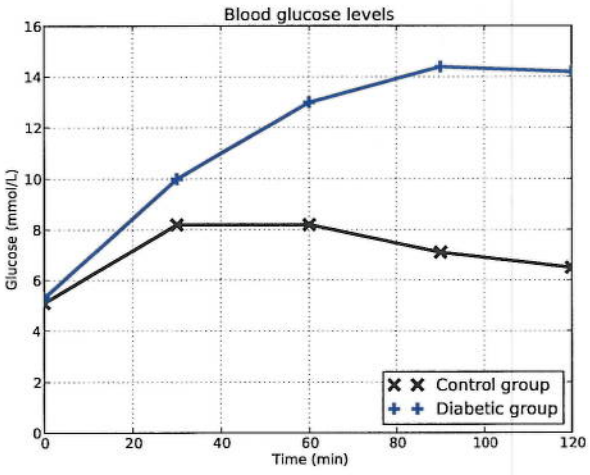


Figure 9.9: A graph of Peter's measured blood glucose levels. Time is given in minutes after 6 am.

- It is possible that an individual blood glucose measurement might be within the normal range at some instant in time, but outside that range over a longer time period. It is often very useful to analyse AUCs as well.

Question 9.2.2

Figure 9.10 shows measured blood glucose levels for a diabetic group and a control group, adapted from [25].



D C

time (mins)	diabetic (mmol/L)	control (mmol/L)
0	5.3	5.1
30	10	8.2
60	13	8.2
90	14.4	7.1
120	14.2	6.5

Figure 9.10: Mean blood glucose levels for diabetic and non-diabetic (control) groups.

(continued over)

***Question 9.2.2** (continued)*

(a) Roughly estimate the AUC for the control group.

$AUC_{control} = 7 \times 120 = 840$ mmol. min/L

(b) Use the trapezoid rule to calculate the AUC for the control group.

$(30-0) \times \frac{1}{2}(8.2+5.1) = 199.5$

$(60-0) \times \frac{1}{2}(8.2+8.2) = 246$

$(90-60) \times \frac{1}{2}(8.2+7.1) = 229.5$

$(120-90) \times \frac{1}{2}(7.1+6.5) = 204$

$\Sigma = 879$ mmol. min/L

(c) The paper [25] found that the "glucose AUC" for the control group is around 265 mmol/L/min. Comment on the units and compare the given AUC for the control group with your values in Parts (a) and (b).

Wrong value & wrong units!

Might have just calculated area above minimum value. i.e. $5.1 \times 120 = 612$

So $879 - 612 = 267$ - close??

(d) How does the concept of measuring the AUC for a blood glucose curve relate to the concept of *total exposure to alcohol* discussed earlier.

AUC is total exposure to glucose. Glucose can be a problematic for the body just like alcohol.

End of Case Study 25: Sweet P's.

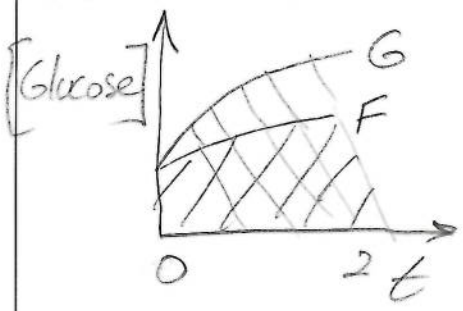
Case Study 26: **Hi GI!**

- The *Glycaemic Index* or GI of foods is often mentioned in marketing campaigns and in association with dietary health claims.
- GIs range between 0 and 100, and indicate the relative extent by which blood glucose levels rise after the consumption of a food. Hence, GI scores are only valid for foods containing carbohydrates.
- Researchers classify foods into the following GI categories.
  - *Low GI* – when GI is ≤ 55, digestion of carbohydrates is slow, with a slow rise and lower peak in blood sugar level. Examples of Low GI foods include cherries, skim milk, apples, chick peas, oranges and carrots.
  - *Medium GI* – when the GI is between 56 and 69, the digestion of carbohydrates occurs at a moderate rate. Examples of Medium GI foods include include boiled potato, honey, ice cream and sultanas.
  - *High GI* – when the GI is 70 or higher, the digestion of carbohydrates is fast, leading to a rapid rise and high peak in blood sugar level. Examples of High GI foods include mashed potato, white bread, cornflakes, watermelon and steamed white rice.
- The many claimed health benefits of Low GI diets include weight loss and improved weight control, improved management of diabetes, reduced risk of cardiovascular disease and increased physical stamina.
- Criticisms of focusing on GIs as a dietary tool include:
  - GIs can vary greatly for a given food, depending on how ripe it is, and how it is processed, stored and cooked;
  - the GI of a food may be less important than the *quantity* consumed;
  - measured GIs may not be very exact or reliable. For example, if the GI of a given food is measured at different times of the day then the results can differ quite substantially; and
  - GIs are measured under tightly controlled experimental conditions, so may not translate to typical food consumption conditions.

- Researchers calculate the GI of a food in the following way:
  - Ten healthy people fast overnight. In the morning, each person consumes a controlled dose of the test food, with known total carbohydrate content (typically 50 g).
  - Over a 2 hour period, researchers measure the increase in the blood glucose level above baseline for each participant, produce graphs, and calculate the AUCs for the test food using the trapezoid rule.
  - On a separate day, participants undergo the same procedure, but consume a glucose solution which contains the same amount of total carbohydrate. Researchers calculate the AUC for glucose (the reference food) for each individual.
  - The **definition** of the GI is: divide the AUC of the test food by the AUC for glucose, and multiply by 100%. An average of the individual GI scores represents the overall GI for the test food.

Question 9.2.3

(a) Use a diagram to illustrate calculation of GI for a test food.



$$GI = \frac{AUC_F}{AUC_G} \times 100\%$$

OR

$$GI = \frac{\int_0^2 F(t)\,dt}{\int_0^2 G(t)\,dt} \times 100\%$$

(b) In Question 9.1.8 we calculated the ratio of the exposure to alcohol of a 'typical' woman versus that of a 'typical' man. How is that calculation similar to the method for calculating the GI of a food?

It is a ratio of exposures - the first ratio of exposures to alcohol & the second the ratio of exposures to glucose.

Question 9.2.4

A study [56] recorded the following blood glucose levels when an individual consumed a controlled quantity of bread in one test, then lentils in another test. The increase in the AUCs above baseline glucose levels is 29 units for lentils, and 66 units for bread. Given that the GI of bread is around 70, find the GI of lentils.

Time (mins)	Glucose (mmol/L) Bread	Lentils
0	4.3	4.3
15	4.8	4.8
30	6.2	4.8
45	5.5	4.3
60	5.0	4.4
90	3.9	4.5
120	4.1	4.7

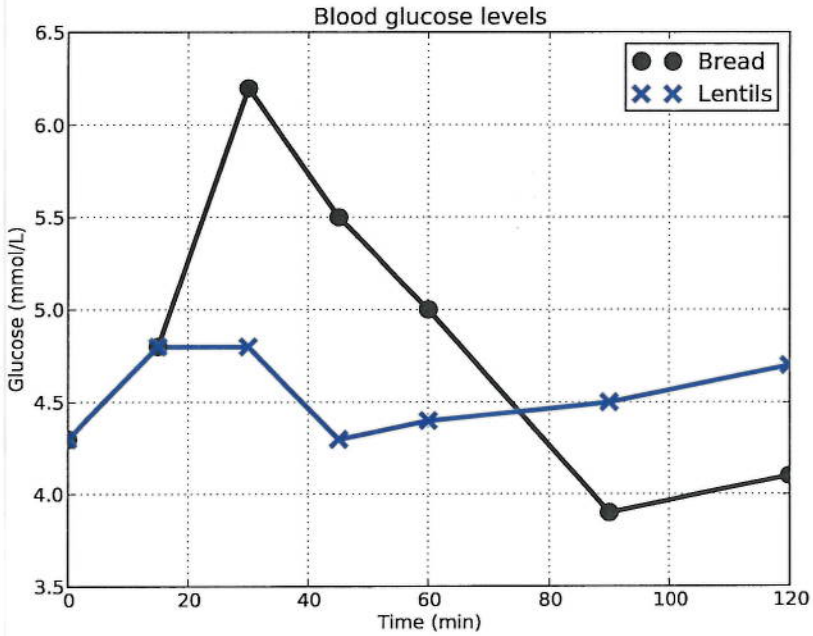


Figure 9.11: Mean blood glucose levels.

$GI_{Lentils} = ?$ $AUC_{Bread} = 66$ $AUC_{Lentils} = 29$ $GI_{Bread} = 70$ $AUC_{Glucose} = ?$

$$GI_{Food} = \frac{AUC_{Food}}{AUC_{Glucose}} \times 100\%$$

So estimate $AUC_{Glucose}$ from Bread & then substitute into Lentils

$$GI_{Bread} = \frac{AUC_{Bread}}{AUC_{Glucose}} \times 100\% \Rightarrow AUC_{Glucose} = \frac{AUC_{Bread}}{GI_{Bread}} \times 100\% = \frac{66}{70} \times 100\% = 94.3$$

$$\Rightarrow GI_{Lentils} = \frac{AUC_{Lentils}}{AUC_{Glucose}} \times 100\% = \frac{29}{94.3} \times 100\% = 30.75$$

OR $\frac{GI_{Bread}}{GI_{Lentils}} = \frac{AUC_{Bread}}{AUC_{Lentils}} \Rightarrow GI_{Lentils} = \frac{GI_{Bread} \times AUC_{Lentils}}{AUC_{Bread}}$

End of Case Study 26: Hi GI!.

$$\Rightarrow GI_{Lentils} = \frac{29 \times 70}{66} = 30.75$$

Case Study 27: **Bioavailability of drugs**

- Drugs can be administered via many routes, including:
  - orally, such as the contraceptive pill;
  - as a gas, such as nicotine from a cigarette;
  - through the skin, such as a nicotine patch;
  - nasally, such as "snorted" cocaine;
  - intravenously, such as chemotherapy drugs for treating cancer;
  - sublingually, such as nitroglycerin used to treat angina; and
  - rectally, such as a paracetamol suppository.
- Different routes of drug delivery are required depending on the drug type, duration and frequency of treatment, and the condition of the patient. Oral administration is common, but other routes may be more convenient for drugs that cause nausea or vomiting, or for patients who cannot swallow.
- After administration of a drug, it typically needs to pass through a number of stages before it enters general circulation and has a chance to act. This can have a substantial impact on the proportion of the dose available to achieve the desired pharmacological impact. For example, the following *first pass effects* reduce the availability of orally-administered drugs:
  - how readily and rapidly the drug dissolves in the digestive tract;
  - whether the drug is damaged by acidic stomach contents;
  - whether the drug is partially metabolised by bacteria in the gut;
  - how much of the drug is absorbed across the intestinal wall;
  - the digestive health of the individual (for example, vomiting or diarrhoea may cause mechanical expulsion of the drug); and
  - how much of the drug is metabolised in the liver prior to entering general circulation (because blood travels from the small intestine to the liver and then to the rest of the body).
- After administering a drug by a given route, its relative **bioavailability** F is the fraction of the dose that enters general circulation compared to a dose administered via a more direct route, usually *intravenously* (IV).

Definition of bioavailability

If $R(t)$ is the blood concentration of a drug after giving a dose by some route and $I(t)$ is the concentration after an IV dose of **the same size**, then the bioavailability F of the drug administered via this route is

$$F = \frac{\int_0^\infty R(t)\,dt}{\int_0^\infty I(t)\,dt}.$$

Question 9.2.5

(a) Explain the meaning of the expression for F.

$\int_0^\infty I(t)\,dt$. Intravenous route is fastest & best other routes are relative to this

$\int_0^\infty R(t)\,dt$. Particular route (slower than $I(t)$)

$0 \to \infty$ look over a long time

(b) How is the method of calculating the bioavailability of a drug similar to the method of calculating the GI of a food?

Bioavailability of a drug route is relative to the maximum possible (Intravenous)

The GI of a food is the effect on glucose levels in the blood relative to the maximum possible (pure glucose)

Question 9.2.6

In [43] and [33], on separate occasions, test subjects were each administered 1000 mg doses of paracetamol. In [43], doses were intravenous and in oral tablet form. In [33], an aqueous dose was administered as a rectal suppository. Figure 9.12 plots the blood concentrations of paracetamol obtained for these subjects over the following six hours.

(continued over)

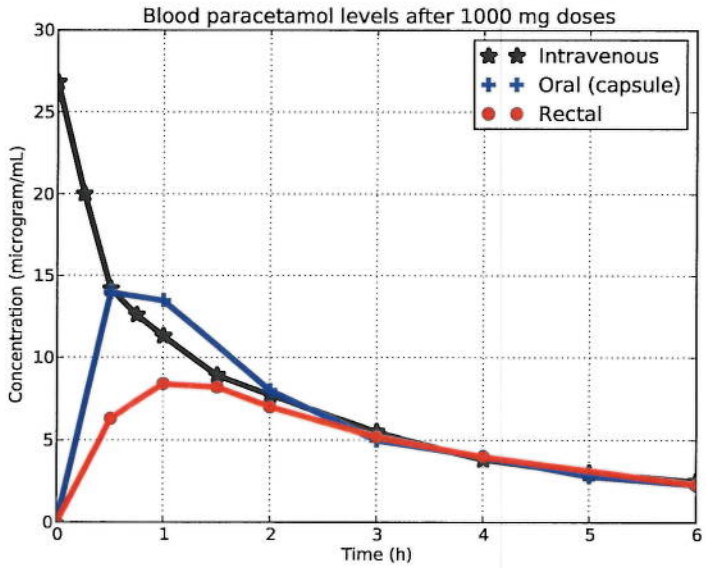


Figure 9.12: Blood concentration curves for paracetamol administered in various ways.

(a) Discuss the shapes of the three curves in Figure 9.12.

Intravenous – instantaneous – no A, D, only M & E
– exponential decay
Oral & Rectal – Some time for A+D to happen. M & E happens too
– surge function shape.

(b) **Estimate** the bioavailabilities of the oral and rectal doses.

Method	# squares	Area. AUC	Bioavailability
IV	8.5	$8.5 \times 5 = 42.5$	$= 1$
Oral	7	$7 \times 5 = 35$	$= \frac{35}{42.5} = 0.82$
Rectal	6	$6 \times 5 = 30$	$= \frac{30}{42.5} = 0.71$

(c) For the IV dose (see [43]), the blood paracetamol concentration in μg/mL at time t in hours after dosing is modelled using the equation

$$I(t) = 13.8e^{-2.55t} + 13e^{-0.28t}.$$

Figure 9.13 plots the measured values and $I(t)$. Using integration, the AUC for this curve is approximately 51.84 μg hr/mL. In [43], the AUC for the oral tablet dose is around 44 μg hr/mL. Calculate the bioavailability of the oral tablet dose.

(continued over)

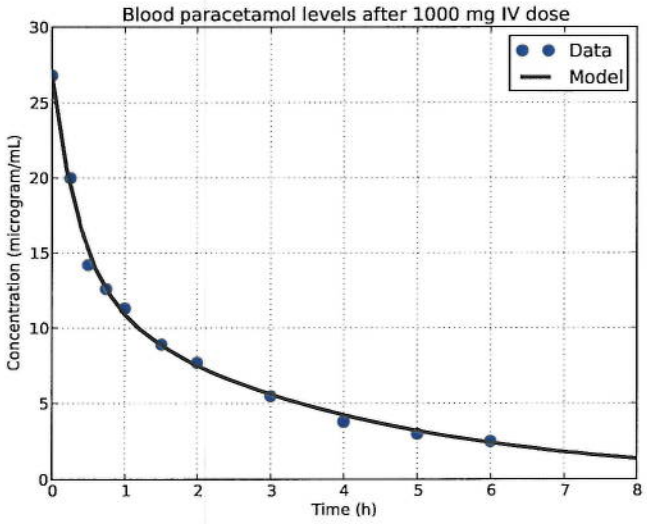


Figure 9.13: The graph of $I(t)$, following an intravenous dose of 1000 mg of paracetamol.

$$\text{Bioavailability}_{oral} = \frac{AUC_{oral}}{AUC_{IV}} = \frac{44}{51.84} = 0.85$$

(d) In [33], the AUC given for the rectal dose is around 2290 μg/mL/min. Comment on the AUC units, then calculate the dose bioavailability.

Wrong units: $\mu g/ml/min \neq \frac{\mu g \cdot min}{ml}$

$$\text{Bioavailability}_{Rectal} = \frac{AUC_{Rectal}}{AUC_{IV}} = \frac{2290 \frac{\mu g \cdot min}{ml} \times \frac{1}{60} \frac{hr}{min}}{51.84 \frac{\mu g \cdot hr}{ml}}$$

$$= 0.74$$

(e) Compare your answers to Parts (c) and (d) with those to Part (b).

Bioavailability $\quad Estimate_{Rectal} = 0.71$ Close!

$Actual_{Rectal} = 0.74$

End of Case Study 27: Bioavailability of drugs.

Case Study 28: **Bioequivalence of drugs**

- Developing new drugs is a lengthy, complex and expensive process.
- Before the new drug can be prescribed and sold, it must pass stringent effectiveness and safety tests administered by the relevant approval body (in Australia, this is the *Therapeutic Goods Administration*, or TGA; in the United States it is the *Food and Drug Administration*, FDA).
- Once a drug is approved, the developing company is typically granted a *patent* over the active ingredient in the drug, providing them with legally enforcible marketing and sales exclusivity for up to 20 years.
- When a drug patent expires (or earlier, if the patent holder agrees), other manufacturers can produce equivalent drugs, typically called *generics*.
- You will almost certainly have purchased (or been offered) generic drugs at pharmacies; they are typically cheaper than the original drugs, and the Australian Pharmaceutical Benefits Scheme(PBS) has encouraged people to consider using generic drugs where available.

***Example** 9.2.7*

Earlier we mentioned the antidepressant drug *Zoloft*, which is a brand name of the drug *sertraline hydrochloride*. This drug was developed by the pharmaceutical company Pfizer, introduced to the market in 1991, and the US patent expired in 2006. In Australia there are now numerous generic versions of sertraline hydrochloride that have been approved by the TGA and are covered by the PBS, including *Xydep*, *Eleva*, *Concorz*, *Setrona*, *Sertracor*, *Terry White Chemists sertraline*, *Chemmart sertraline* and *Sertraline generichealth*.

- To be approved for prescription and sale, a generic drug must contain the same active ingredients as the original "brand name" drug, and the manufacturer has to demonstrate that it is *bioequivalent* to the original.

Definition of bioequivalence

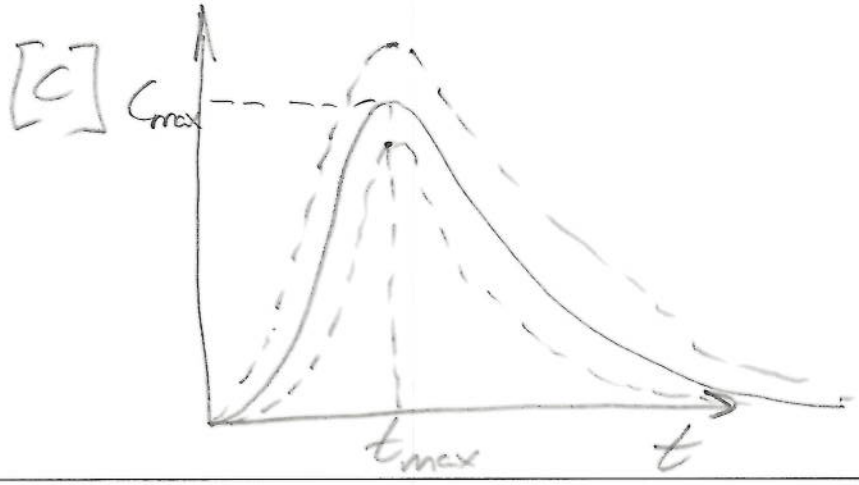
The TGA has adopted the following regulatory definition of *bioequivalence*:

- The drugs must have the same quantities of the same active ingredients.
- The 90% confidence intervals for the ratios of:
  - the areas under the drug concentration curves (AUC ratio); and
  - the maximum blood concentration of the drugs (C_{max} ratio)

  must lie in the range 0.8 to 1.25.
- The times to maximum blood concentration (t_{max}) must be similar.

Question 9.2.8

Using a graph, interpret the AUC and C_{max} ratio tests for bioequivalence.



Example 9.2.9

In [38], researchers compared the bioequivalence of a generic drug and Zoloft. The mean pharmacokinetic ratios, for 24 volunteers, with 90% confidence intervals given in brackets, were: C_{max}: 1.0447 (0.9664 – 1.1293); AUC_{0-96}: 1.0806 (1.0071 – 1.1594); and $AUC_{0-\infty}$: 1.0839 (1.0093 – 1.164). In addition, they found "no significant difference" in the value of t_{max}. Hence the study deduced that the two formulations were bioequivalent.

End of Case Study 28: Bioequivalence of drugs.

Part 5: Life and Death

"O sweeter than the marriage-feast,
'Tis sweeter far to me,
To walk together to the kirk
With a goodly company!

To walk together to the kirk,
And all together pray,
While each to his great Father bends,
Old men, and babes, and loving friends,
And youths and maidens gay!

Farewell, farewell! but this I tell
To thee, thou Wedding-Guest!
He prayeth well, who loveth well
Both man and bird and beast.

He prayeth best, who loveth best
All things both great and small;
For the dear God who loveth us,
He made and loveth all."

The Mariner, whose eye is bright,
Whose beard with age is hoar,
Is gone; and now the Wedding-Guest
Turned from the bridegroom's door.

He went like one that hath been stunned,
And is of sense forlorn:
A sadder and a wiser man
He rose the morrow morn.

The Rime of the Ancient Mariner (1797 – 98), Samuel Taylor Coleridge, (1772 – 1834).

Image 9.2: *The three ages of Man and Death* (c1510), Hans Baldung (c. 1484 – 1545), Museo del Prado, Madrid. (Source: commons.wikimedia.org).

Recall that there are five broad parts to this course. This is Part 5, "Life and death".

Parts 1 and 2 of the notes explored the importance of modelling in science, and how to use mathematics and computing when developing and applying models. Part 3 discussed what constitutes science and knowledge, and the importance of clear communication. Part 4 examined rates of change and areas under curves, in a pharmacological context.

Part 5 of the notes includes three chapters. Here we will investigate changing phenomena in the context of life cycles and populations, and explore how to model these phenomena. The mathematical content that underpins this section is the *differential equation* or DE. Differential equations describe how the value(s) of one or more quantities are changing, and typically involve a function and its derivative(s).

In the first chapter of Part 5, we study the exponential and logistic DEs, and use them to model changes in the sizes of simple populations. We will also learn how to use *Euler's method* to solve DEs numerically.

In the second chapter, we introduce some simple systems of DEs that allow us to model organisms with multiple life stages, and situations in which multiple populations interact, such as predator/prey relationships.

In the third chapter, we use systems of DEs to model the spread of infectious diseases, and investigate the potential impact of pandemics.

Differential equations are very useful, "natural" tools for modelling a huge range of phenomena in science and other areas. For example, if you know that a population is changing in size at a rate proportional to its current value, then you can write and solve a simple DE to represent what is happening. Many natural phenomena display this kind of relationship, so knowing how to write and solve the DE allows the phenomena to be studied.

Chapter 10: **Differential equations and populations**

I see skies of blue, clouds of white,
Bright blessed days, dark sacred nights.
And I think to myself, what a wonderful world.
I hear babies cry, I watch them grow,
They'll learn much more, than I'll never know.
And I think to myself, what a wonderful world.
Yes I think to myself, what a wonderful world.

Artist: *Louis Armstrong.* Song: *What a wonderful world.*

Image 10.1: *The Entry of the Animals into Noah's Ark* (1613), Jan Brueghel the Elder (1568 – 1625), The J. Paul Getty Museum, Los Angeles. (Source: en.wikimedia.org.)

10.1 Introduction to differential equations

- Typically, developing a mathematical model of a phenomenon involves deriving one or more equations that predict the *value* of the phenomenon.
- Sometimes, this is difficult or impossible. Instead, it may be possible to write an equation for how the value is *changing*, then use mathematical techniques to deduce information about the *value*.
- An equation for the rate at which the value of a phenomenon is changing is called a *differential equation*.
- To understand differential equations, it is essential to be clear about the distinction between the *value* of a phenomenon, and the *rate at which that value is changing*.

Differential equations

If y is an unknown function of t, then a **differential equation** is an equation that involves a combination of t, y and/or the derivatives of y.

In all of the examples we will study, the DE will be of the form $y' = \ldots$. (That is, the DE will be an equation for the rate of change of y.)

A particular function y is called a **solution** to a DE if the DE is true when y and its derivative(s) are substituted into the DE.

Thus, a solution to a DE is **another function** which, when **substituted** into the DE, makes the DE **true**.

- Just as with any mathematical model, there are two steps to modelling with DEs: *writing* the equations, and then *solving* them.
- DEs are often very "natural" ways of representing phenomena. That is, it often makes "more sense" to write an equation for a rate of change of some value than to write an equation for the value.

Question 10.1.1

Write a DE to model each of the following.

(a) The straight line distance $D(t)$ travelled by a car increases by 10 m/s.

$D'(t) = 10$

(b) The human population P of Earth is increasing at about 1.1% per annum. Assume this growth rate continues.

$P'(t) = 0.011P$

(c) According to the Australian Bureau of Statistics, during 2013 Australia had: birth rate 1.34%; death rate 0.64%; 504000 people inward migration; 268000 people outward emigration. If these changes continue indefinitely, write a DE for the Australian population $P(t)$ in year t.

$$\begin{aligned} P'(t) &= (B-D)P + I - E \\ &= (0.0134 - 0.0064)P + 504000 - 268000 \\ &= 0.007P + 236000 \end{aligned}$$

(d) The *von Bertalanffy growth model* states that the rate of increase of the length $L(t)$ of a shark of age t years is proportional to an intrinsic positive growth rate r and the difference between a fixed maximum length M and its current length $L(t)$.

$L'(t) = r(M - L(t))$

(e) *Newton's Law of Cooling* states that the rate of change of the temperature $T(t)$ of an object placed in an environment with fixed temperature F is proportional to the relative difference in the temperatures.

$$\begin{aligned} T'(t) &= -k\,\Delta T \qquad F < T \\ &= -k(T(t) - F) \end{aligned}$$

OR $T'(t) = +k(F - T(t))$

- Once a DE has been written for the *rate of change* of a phenomenon, that DE can (sometimes) be *solved* to give the *value* of the phenomenon.
- This (usually) requires an additional piece of measured information, such as the value of the phenomenon at some time, often $t = 0$.

***Example** 10.1.2*

Refer to Question 10.1.1. The following additional information applies in each case.

(a) Initially the car has travelled 0 metres.

(b) The human population of Earth in July 2014 was about 7.24 billion.

(c) The human population of Australia at the start of 2013 was 22.9 million.

(d) In [23], it is shown that:

- the maximum length of a female *sand tiger* or *grey nurse* shark is $M = 295.8$ cm;
- the intrinsic growth rate is $r = 0.11$ yr^{-1}; and
- the length of a typical shark at birth is 110 cm.

(e) An item with temperature 85 °C is placed in a room with constant temperature 25 °C.

- We can solve some DEs *analytically*, using integration and algebra. (We will discuss this more later.)
- However, for more complex cases, this is not possible. Instead, we can often find approximate solutions using numerical algorithms.
- One numerical algorithm for finding an approximate solution to a DE is called *Euler's method*.

10.2 Euler's method

- Euler's method is a fairly simple approach, that you have probably used before without knowing.
- It is based on the following simple observation: if the value of some quantity is changing at a certain amount per time period, then it is possible to estimate the future value at some time using the following:

(future value) =
(current value) +
(estimated change per time period) × (number of time periods)

Question 10.2.1

(See Question 10.1.1 and Example 10.1.2.) The human population of Earth in July 2014 was 7.24 billion, and was expected to grow by about 79.6 million over the next year.

(a) Assuming the population increases by the same number each year, predict the population in July 2017.

$P = 7240 \text{ million} + (79.6 \text{million} \times 3) = 7478.8 \text{ million}$

(b) Instead of growing by a fixed number each year, the global population has a growth rate of about 1.1% per annum. Estimate the population in July 2015, July 2016 and July 2017.

$P'_{2015} = 0.011P = 0.0011 \times 7240 \text{ million} = 79.64 \text{ million}$

$\Rightarrow P_{2015} = 7240\text{M} + 79.64\text{M} = 7319.64\text{M}$

$P'_{2016} = 0.011 \times 7319.64\text{M} = 80.516\text{M}$

$\Rightarrow P_{2016} = 7319.64\text{M} + 80.51\text{M} = 7400.16\text{M}$

$P'_{2017} = 0.011 \times 7400.16\text{M} = 81.402\text{M}$

$\Rightarrow P_{2017} = 7400.16\text{M} + 81.402\text{M} = 7481.56\text{M}$

(continued over)

***Question** 10.2.1 (continued)*

(c) Compare your answers to Parts (a) and (b), and explain the difference.

(a) Same # added each year = Simple Interest

(b) # added dependent on previous one

$y' = ry$ = Compound Interest

$y_{new} = y_{old} + ROC \times$ time period.

- Question 10.2.1 illustrates a general approach called *Euler's method.*
- It proceeds by approximating the unknown function as a series of short **straight lines**, starting from the initial point, each with:
  - **width** equal to a chosen step size h;
  - **slope** equal to the estimated slope calculated using the DE; and
  - **height** equal to the width multiplied by the slope.
- The following is a more formal description of Euler's method.

Euler's method

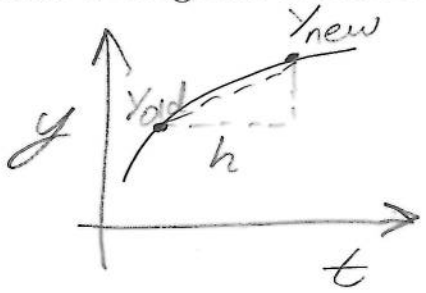
Given an unknown quantity y, a DE of the form $y' = \ldots$, and a value of y (say y_0) at a given time (say t_0):

1. Choose a small *step size* h, and start at the initial point (t_0, y_0).
2. Use the DE to find the (estimated) slope y' at the current point, by substituting the current values of t and y into the DE.
3. Advance the current point to the end point of a short straight line, by setting $t = t + h$ and $y = y + h \times y'$.

   The new point (t, y) is the next approximate function value.
4. If t has reached the desired end-point then stop, else return to Step 2.

Question 10.2.2

Draw a diagram illustrating Euler's method.



$$\text{Slope} = \frac{\text{rise}}{\text{run}} = \frac{y_2 - y_1}{x_2 - x_1}$$

$$y' = \frac{y_{New} - y_{old}}{h}$$

$$\Rightarrow y'h = y_{New} - y_{old}$$

$$\Rightarrow \boxed{y_{New} = y_{old} + y'h}$$

- In the following example we will use Euler's method to solve a simple DE as an illustration of how the method works.
- Whenever you are asked to apply Euler's method, you will be given the value of the initial population, and also what step size to use.

Example 10.2.3

Consider a population of algae growing at a rate of 10% per day, with an initial population of 100 algal cells per mL. Use Euler's method to estimate the population per mL after 5 days, using a step size of $h = 1$ day.

Answer: This population can be modelled by the DE

$$y' = 0.1y \qquad \text{where} \qquad y_0 = 100.$$

(continued over)

***Example** 10.2.3 (continued)*

With a step size of $h = 1$, to find the approximate value of y when $t = 5$, we proceed as shown in Figure 10.1. (Remember that at each step, the new value of t equals the previous value of t plus h.)

Step	t	y	$y' = 0.1y$	$h \times y'$	new t	new y
0	0	100	10	10	1	110
1	1	110	11	11	2	121
2	2	121	12.1	12.1	3	133.1
3	3	133.1	13.31	13.31	4	146.41
4	4	146.41	14.641	14.641	5	161.051

Figure 10.1: Five steps of Euler's method.

So when $t = 5$, $y \approx 161.051$.

Figure 10.2 shows a graph of the approximate solution. The five y values from the last column of Figure 10.1 are marked as asterisks, with straight lines approximating the function between these points.

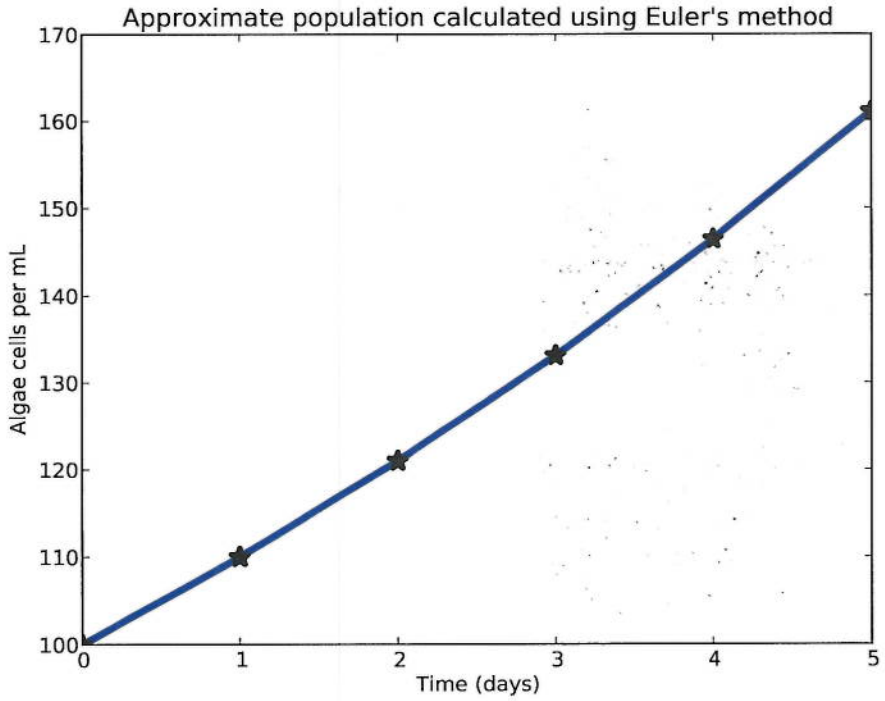


Figure 10.2: Approximate solution to the DE $y' = 0.1y$ with $y(0) = 100$ cells per mL.

There are some important things to know about Euler's method.

- It gives an **approximate** solution, not an exact one. There will be numerical inaccuracies in the answer, particularly over a large range of t values.
- The choice of step size is very important: smaller values will give a more accurate answer, but take longer to calculate.
- Despite these limitations, the method can give very good approximate solutions to quite difficult problems.

Example 10.2.4

In Example 10.2.3 we used a step size of $h = 1$ to solve $y' = 0.1y$.

Figure 10.3 shows approximate solutions with a step size of $h = 2.5$ (bottom curve), $h = 1$ (middle curve) and the exact, true solution (top curve).

As h becomes smaller, the solution becomes more accurate (that is, it moves closer to the true solution).

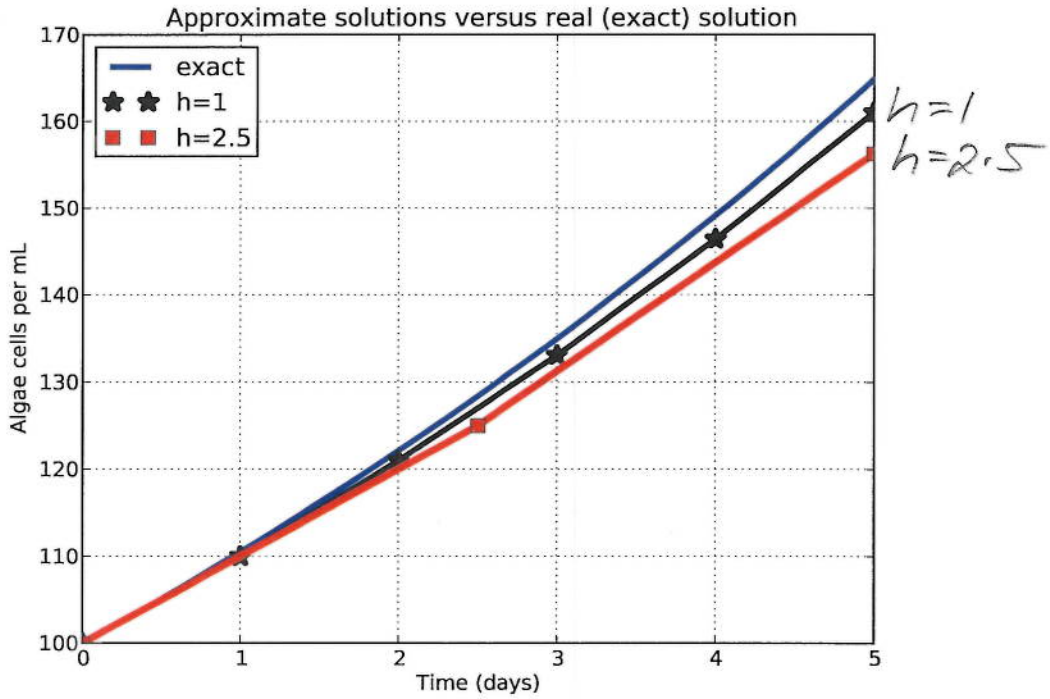


Figure 10.3: Approximate and exact solutions to the DE $y' = 0.1y$ with $y(0) = 100$ cells per mL.

- Euler's method can be used to solve a wide range of differential equations, approximating unknown functions with series of short straight lines.

10.3 The exponential DE

- In the previous section we saw how to use Euler's method to find an approximate solution to a DE. Sometimes we can solve a DE *analytically*, using integration and algebra. This gives an exact solution to the DE.
- In such cases there is no need to use Euler's method to find an *approximate* solution, because the *exact* solution is known.
- Earlier we studied exponential growth and decay. On Page 104 we state "Any phenomenon that has a rate of change proportional to the current amount follows an exponential function".
- This occurs precisely because such phenomena satisfy simple DEs whose solutions are exponential functions.

DE for exponential growth and decay

Any quantity $N(t)$ whose rate of change at any time is proportional to the value of N, with rate of change equal to a constant r per time period, follows the DE

$$N' = rN.$$

The solution to this DE is the exponential function

$$N(t) = N_0 e^{rt},$$

where N_0 is the value of N at $t = 0$.

***Example* 10.3.1**

We can easily show that $N(t) = N_0 e^{rt}$ is the solution to $N' = rN$.

First, differentiate N, giving: $N' = r \times N_0 e^{rt}$.

Then, substitute N into that equation, giving: $N' = rN$.

- Many populations can be modelled effectively using exponential functions, for some periods of time.

Case Study 29: **Poo**

Photo 10.1: Heap of steaming panda poo. (Source: PA.)

- *Escherichia coli* (usually shortened to *E. coli*) are bacteria commonly found in the lower intestine of warm-blooded animals, including humans.
- Most strains of *E. coli* are harmless in the digestive system, or even beneficial to the host. However, some strains produce toxins, and can cause food poisoning, gastrointestinal infections and urinary tract infections.
- One such strain is O104:H4, which caused outbreaks of illness in Europe in 2011. Around 50 people died and more than 4000 became ill. Contamination was traced to a farm that grew organic sprouted vegetables.
- Because *E. coli* can survive outside the body for some time, tests for *E. coli* are often used to identify faecal contamination in environmental samples or foods during hygiene checks.
- Under simplifying assumptions (such as relatively unlimited resources) the rate of increase of a population of *E. coli* at any time is proportional to the population size at that time.
- Hence the population follows an exponential function, and it makes sense to discuss the *doubling time* of the population.
- Under favourable conditions, the doubling time for a population of *E. coli* may be an hour, or even shorter. The rapid growth rate is one reason why good hygiene standards are important in food preparation.

- When studying populations of bacteria, microbiologists commonly count *colony-forming units* (CFU), which is the number of *live* bacterial cells. (Direct counts of individuals include both dead and living cells.)

Question 10.3.2

A population of *E. coli* has a growth constant of 1% per minute. Let $N(t)$ be the population in CFU/g at time t minutes, and assume there are 1000 CFU/g at time $t = 0$. Write and solve a DE to model this population.

DE: $N'(t) = 0.01N$

Solve: $N(t) = N_0 e^{rt} = 1000e^{0.01t}$

Check: $N'(t) = 0.01 \times 1000e^{0.01t} = 0.01N$

Question 10.3.3

A study [50] investigates *E. coli* contamination of pre-cooked meat. Researchers contaminated some ham with 10^7 CFU of *E. coli* and then sliced the ham. The same blade was then used to slice clean ham. The number of CFU were counted on each of 100 slices of the second ham, showing that Slice 1 contained around 580 CFU and Slice 100 contained 9 CFU.

Photo 10.2: Ham with (subtle) faecal contamination. (Source: PA.)

(continued over)

***Question** 10.3.3 (continued)*

(a) How many CFU will be on Slice 1 after 24 hours, assuming it is stored under ideal growing conditions for *E. coli*?

$N_0 = 580$ CFU $\quad r = 0.01$

$N = N_0 e^{rt} = 580 e^{0.01 \times 60 \text{ mins/hr} \times 24 \text{ hrs}}$

$= 580 e^{14.4} = 1.04 \times 10^9$ CFU

= 1 billion after 1 day!

(b) At what time does Slice 100 contain 580 CFU?

$N_0 = 9 \quad r = 0.01$ When does $N = 580$?

$N = N_0 e^{rt}$

$\Rightarrow 580 = 9 e^{0.01t}$

$\Rightarrow \frac{580}{9} = e^{0.01t}$

$\Rightarrow \ln(64.4) = 0.01t$

$t = \frac{4.166}{0.01} = 416$ min (~7 hrs)

(c) Human faeces can contain 10^9 CFU per gram. Given the results of this study, what does this mean for hygiene practices in food preparation?

Wash hands, utensils, machines

Don't poo in kitchen!!!

End of Case Study 29: Poo.

10.4 Limited scope for growth

- Exponential growth models are *unconstrained*, so the growth continues indefinitely with a constant proportional rate of increase, say r.
- This can be accurate over limited time periods, but in reality populations cannot continue to show unconstrained growth forever. Many *constrained* growth models assume that there is a maximum population size.

Carrying capacity of an ecosystem

The **carrying capacity** K of an ecosystem is the maximum population size of a particular organism that is supported by resources within the ecosystem. Resources may include food, water, shelter and sunlight. The carrying capacity for a particular organism often changes over time; for simplicity, we will assume it remains constant.

A population size below the carrying capacity will typically increase towards the carrying capacity, whereas a population size above the carrying capacity (which may occur when, for example, a lake is overstocked with fish) will typically decrease to the carrying capacity.

Question 10.4.1

Let $N(t)$ be the size of a fish population in a certain lake at any time t in months, and let the carrying capacity of the lake be K fish. Draw a rough sketch of $N(t)$ versus time.

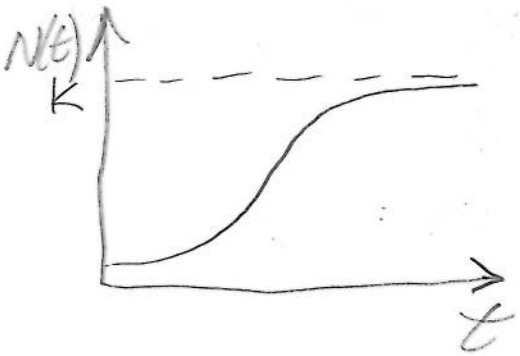
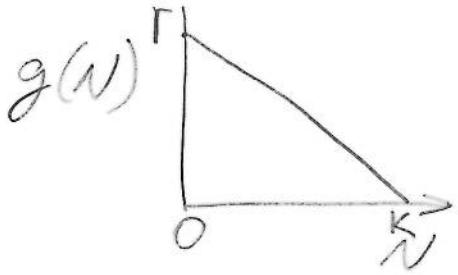


Photo 10.3: Kiss me, red. (Source: PA.)

- One way to model constrained growth is to assume that the growth rate varies with the population size N, rather than remaining constant.

***Question* 10.4.2**

Assume that a population has an unconstrained growth rate of r. As the population N increases, the effective growth rate g reduces linearly from r, until the population reaches the carrying capacity K at which point the effective growth rate is 0. Derive an expression for g as a function of N.



$y = a + bx$

$g(N) = r - \frac{r}{K}N$

$$g(N) = r\left(1 - \frac{N}{K}\right)$$

What is $\frac{N}{K}$? Fraction of carrying capacity full

$\therefore \left(1 - \frac{N}{K}\right)$ is fraction of " available or empty.

Logistic DE

Any function $N(t)$ that changes at a rate proportional to its value (with unconstrained growth rate r), **and also** in reverse proportion to how close the value is to a carrying capacity K, is modelled by the logistic DE

$$N' = r\left(1 - \frac{N}{K}\right) N.$$

- In the logistic model, the change in the population N' will:
  - **increase** as the population size gets larger and there are more individuals who can reproduce; and
  - **decrease** as the population size gets larger as individuals compete for scarce resources.
- The power of the logistic model is the interaction between two opposing factors, growth and competition.

Solution to the logistic DE

The logistic DE can be solved exactly or approximately using Euler's method. If N_0 is the value of N at time 0, the exact solution is

$$N(t) = \frac{K\,N_0}{N_0 + (K - N_0)e^{-rt}}$$

Example 10.4.3

A fish species with initial population $N_0 = 30$ and unconstrained growth rate of 10% per month lives in a reef with a carrying capacity of K=1000 fish. The function $N(t)$ gives the number of fish at time t months, and Figure 10.4 graphs $N(t)$ for 80 months, showing the typical "S"-shaped logistic curve.

$$N(t) = \frac{30000}{30 + 970e^{-0.1t}}$$

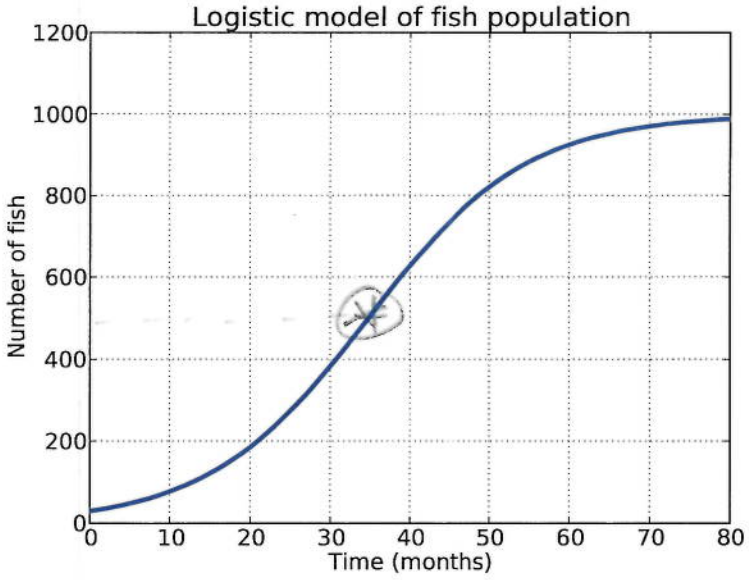


Figure 10.4: Logistic model with an initial population of 30 fish.

Question 10.4.4

With reference to Example 10.4.3, at what population size is the value of N' largest? Explain your answer briefly.

Slope = ROC = N'

Maximum ROC is @ (*) with $N = \frac{K}{2}$

Case Study 30: **Overfishing annoys an oyster**

Photo 10.4: Oysters. (Source: MG.)

- Chesapeake Bay is a large estuary on the east coast of the United States, near the states of Virginia and Maryland. The bay has a surface area of more than 11000 km^2, with a shoreline length of more than 18000 km.
- More than 150 watercourses enter the bay, including the Potomac River, which runs through Washington.
- In the past, the bay supported a diverse range of flora and fauna, including an abundant shellfish population, most notably oysters. However, it has experienced serious environmental degradation due to over-use, overfishing, and polluted runoff from agriculture, urban areas and industry.
- Substantial *marine dead zones*, which are areas of water so low in oxygen that they are unable to support life, are often found within the bay.
- Harvesting oysters is a long-term commercial industry in Chesapeake Bay, however, the size of the population (and hence the harvest) has drastically reduced, due to over harvesting and environmental damage; see Figure 10.5.
- The total current biomass of the oyster population is less than 0.5% of the previous, long-term, pre-harvesting total.
- Between 1982 and 2008, the value of the oyster harvest declined by 88%.
- Considerable research, money and education are being devoted to developing and implementing a sustainable, comprehensive management strategy.

- Figure 10.5 shows the historical annual official oyster harvest data for the Maryland part of the bay.

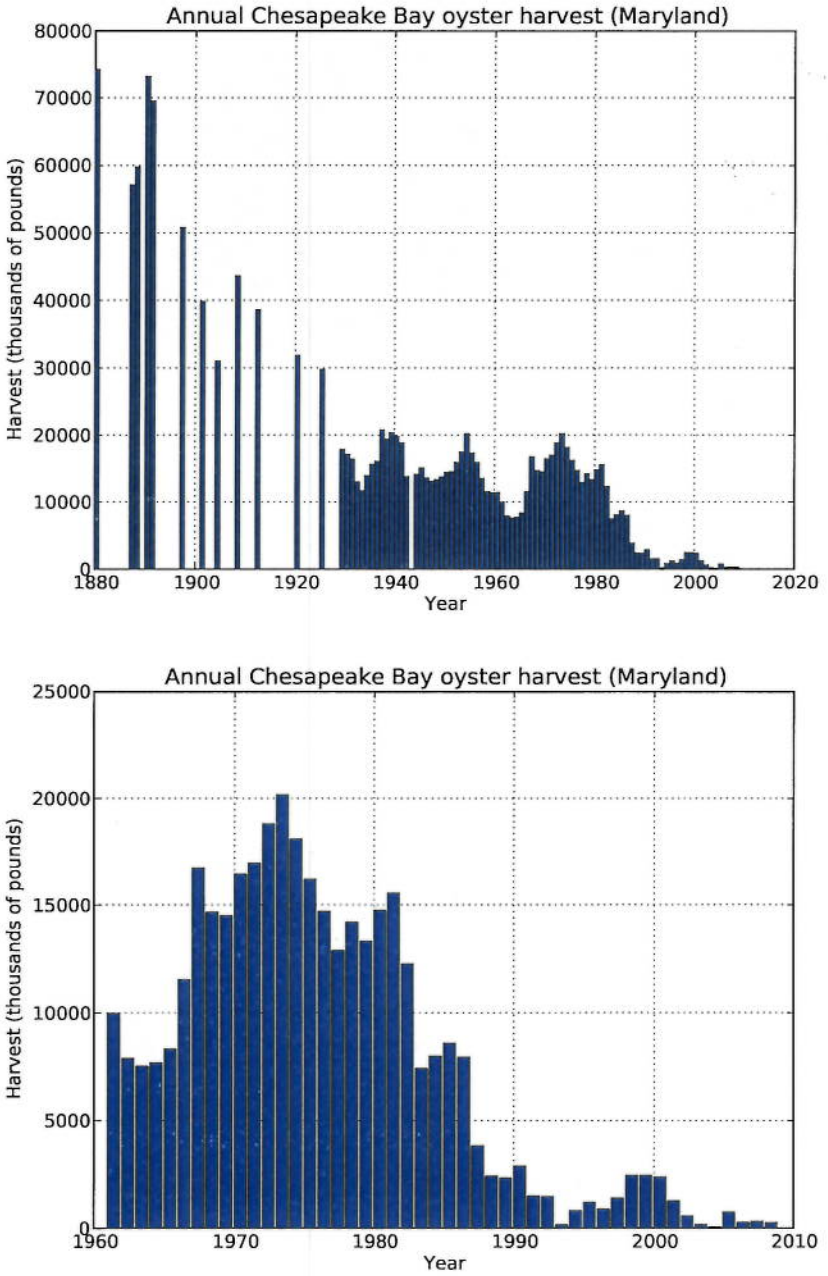


Figure 10.5: Annual Chesapeake Bay oyster harvests. Top: 1880 – 2008. Bottom: 1960 – 2008.

- The paper [53] studied the population of market-sized oysters in the Maryland part of the Chesapeake Bay.
- Using data from 1994 – 2007, researchers found that the effective unconstrained growth rate of market-sized oysters is around $r = 0.133$ per year.
- The estimated carrying capacity of the Maryland part of the bay is around 5×10^9 market-sized oysters.

Question 10.4.5

(a) Write a logistic DE for the population $N(t)$ of market-sized oysters, assuming no harvesting.

$$N' = rN\left(1-\frac{N}{K}\right) = 0.133N\left(1-\frac{N}{5\times10^9}\right)$$

(b) The current population of market-sized oysters in Chesapeake Bay is 81×10^6. Find the **current** annual increase in the population size.

When $N = 81\times10^6$, $N' = 0.133\times81\times10^6\left(1-\frac{81\times10^6}{5\times10^9}\right)$

$$\Rightarrow \boxed{N' = 10.6\times10^6}$$

(c) The 2007 harvest was 50×10^6 oysters. Is this sustainable? Why?

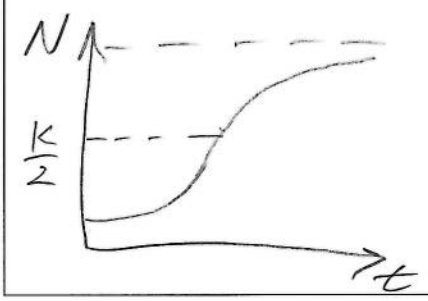
No. The amount it is growing is 10.6×10^6. Taking 5× too many.

(d) Find the largest number of oysters that could be harvested *from the current population* each year without reducing the total population.

10.6×10^6

(e) If there was no oyster harvesting conducted for a few years, how would your answer to Part (d) change? Why?

It depends on N relative to $\frac{K}{2}$.

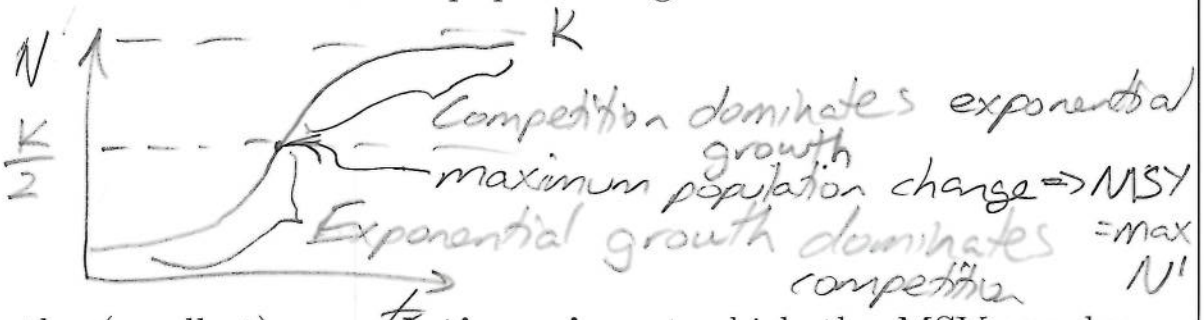


Max. amount harvestable = ROC
& for $N < \frac{K}{2}$ this ↑ as pop^n grows but ↓ for $N > \frac{K}{2}$

(continued over)

Question 10.4.5 *(continued)*

(f) In resource management, especially fisheries management, the *Maximum Sustainable Yield* (MSY) of a population is defined to be the *largest possible harvest size that could be maintained indefinitely.* Explain how the MSY relates to population growth rates.



(g) What is the (smallest) **population size** at which the MSY can be harvested? Why? When $N = \frac{K}{2} = \frac{5\times10^9}{2} = 2.5\times10^9$

OR $N' = rN(1-\frac{N}{K}) = rN - \frac{rN^2}{K}$. Maximum when $N''=0$

$0 = r - \frac{2rN}{K} \Rightarrow N = \frac{K}{2}$.

(h) Find the MSY of oysters in Chesapeake Bay.

Substitute $N=\frac{K}{2}$ into $N' = rN(1-\frac{N}{K})$

$\Rightarrow N' = \frac{rK}{2}(1-\frac{\frac{K}{2}}{K}) = \frac{rK}{2} - \frac{rK}{4} = \frac{rK}{4}$

So $\max N' = MSY = \frac{0.183\times 2.5\times10^9}{4}$

$\boxed{MSY = 0.166\times10^9}$

(i) Noting that there are around 5 oysters per pound, comment on the historical harvest rates in Figure 10.5.

$MSY = 166\times10^6$ oysters

$\Rightarrow \frac{166\times10^6 \text{ oysters}}{5 \text{ oysters/pound}} = 33\times10^6$ pounds.

Historical rates (before 1920) were much higher!

(continued over)

***Question** 10.4.5 (continued)*

(j) A logistic growth model of a population with unconstrained growth rate $r = 0.133$ per year, current population $N_0 = 81 \times 10^6$, carrying capacity $K = 5 \times 10^9$ and no harvesting predicts the following population:

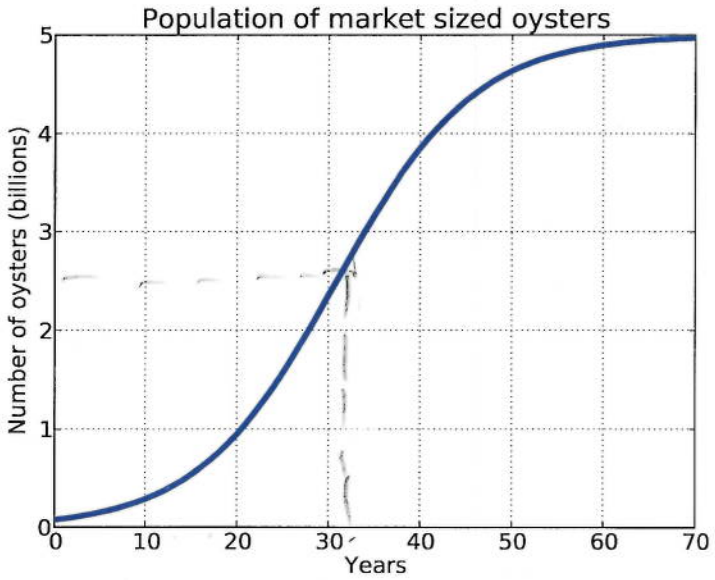


Figure 10.6: Logistic growth of the Chesapeake Bay oyster population with no harvesting.

Make some brief recommendations to assist the government with long-term oyster stock management in Chesapeake Bay.

Current sustainable harvest ≈ 10 million oysters
Could have 166 million oysters if you wait
Have to wait >30 yrs though - livelihoods?
It will grow if you take < growth.

Photo 10.5: Over-exploited? (Source: MG.)

End of Case Study 30: Overfishing annoys an oyster.

Case Study 31: **Overpopulation annoys us all**

- The human population of Earth is rising very rapidly. In 1950 the global population was about 2.5 billion, and in 2012 it first exceeded 7 billion.
- The doubling time has reduced in recent centuries: it took about 300 years for the population to double to one billion, then 120 years to double again, then 47 years to double again.

Question 10.4.6

The following graphs show the human population of Earth over 1000 years (top) and transformed using ln (bottom).

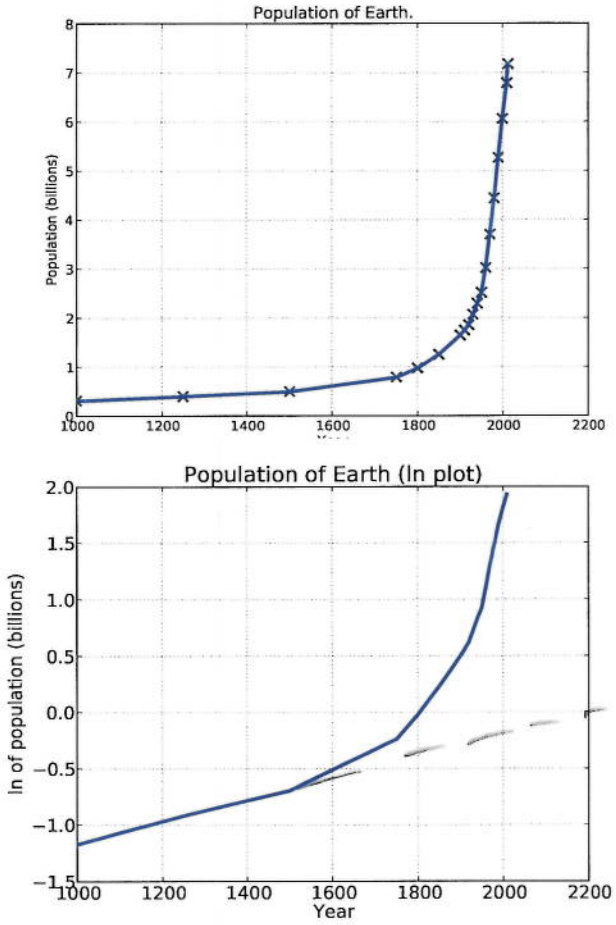


Figure 10.7: Human population of Earth over 1000 years (top), and transformed using ln (bottom).

(a) Is the population shown in Figure 10.7 growing exponentially?

$N = N_0 e^{rt}$ ← linear

$\ln N = \ln N_0 + rt$

∴ human population growing faster than exponential

(continued over)

***Question** 10.4.6 (continued)*

(b) If the population were growing exponentially, describe the shape of the graph showing the annual change in the global population.

$N = N_0 e^{rt} \Rightarrow N' = rN_0 e^{rt}$

∴ should still grow exponentially

(c) Figure 10.8 shows the annual change in global population since 1950. Interpret this graph. What does it mean for global population since 1950, and what might it mean into the future?

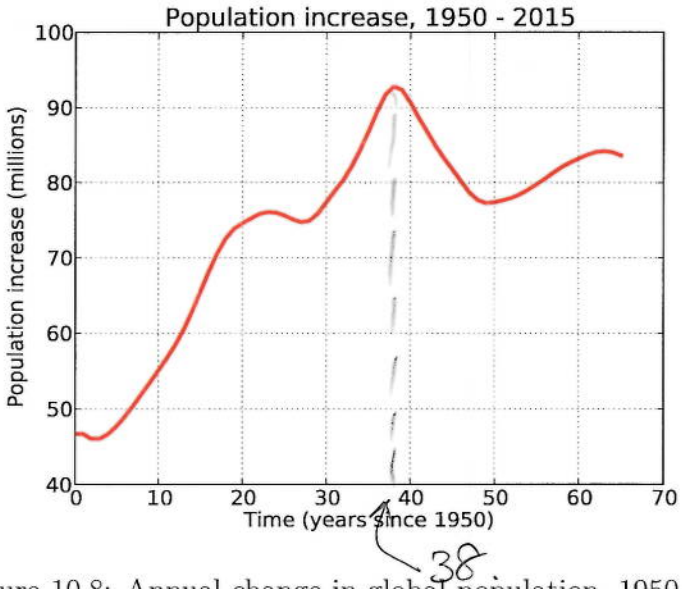


Figure 10.8: Annual change in global population, 1950–2015.

Annual population ↑ peaked in 1988. Population growth is declining or levelling off. Approaching K?

- The United Nations has estimated the populations of all countries, and also globally, for each year until 2100; see [54]. Their models take into account predicted shifts in demographic patterns in each country.
- Figure 10.9 shows the projected global population using four scenarios: maintaining current growth rate; and 'high', 'medium' and 'low' fertility.

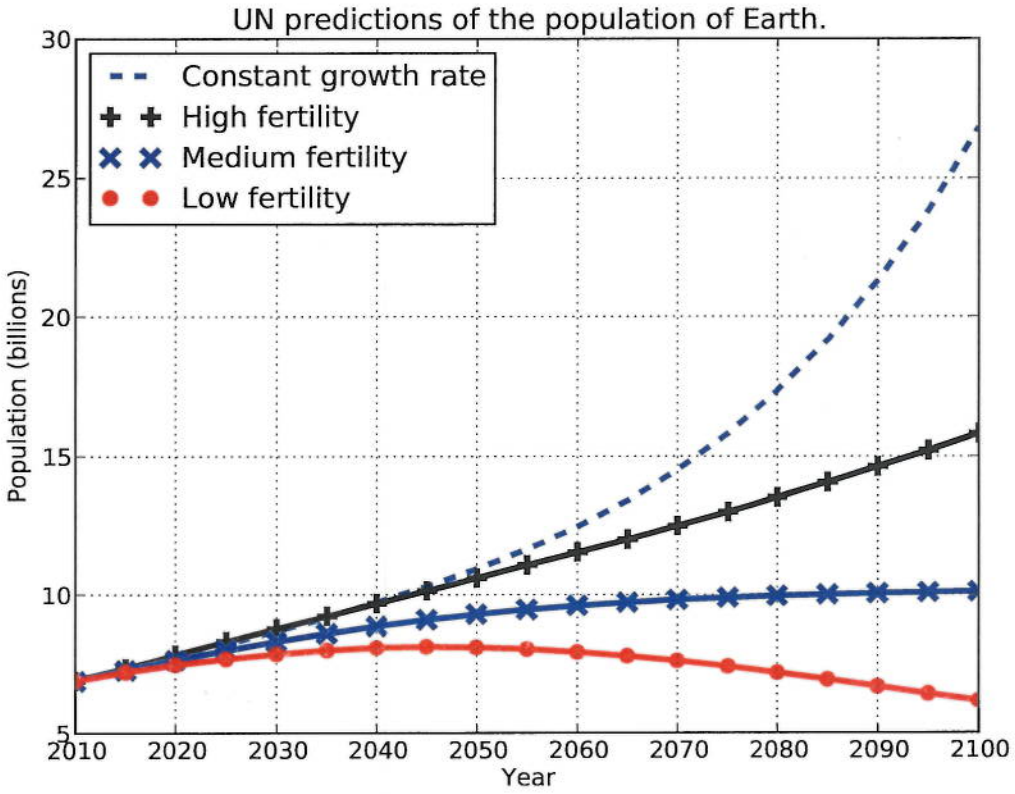


Figure 10.9: UN predictions of global population, 2010–2100.

***Question* 10.4.7**

(a) Interpret each of the graphs.

Current – Constant growth rate ⇒ 27B by 2100

High – Increasing ~linearly ⇒ 16B by 2100

Med – Plateaus ⇒ 10B by 2100

Low – Peaks in 2045 & then declines ⇒ 6B by 2100

(b) Discuss the uncertainty in the UN projections.

What will happen to fecundity?

As standard of living ↑, fecundity ↓

♀ emancipation/equality ↑, fecundity ↓

Question 10.4.8

Figure 10.10 shows the UN projected global population $P(t)$ at five year intervals from 2010 ($t = 0$) until 2100 ($t = 90$), assuming 'medium fertility'.

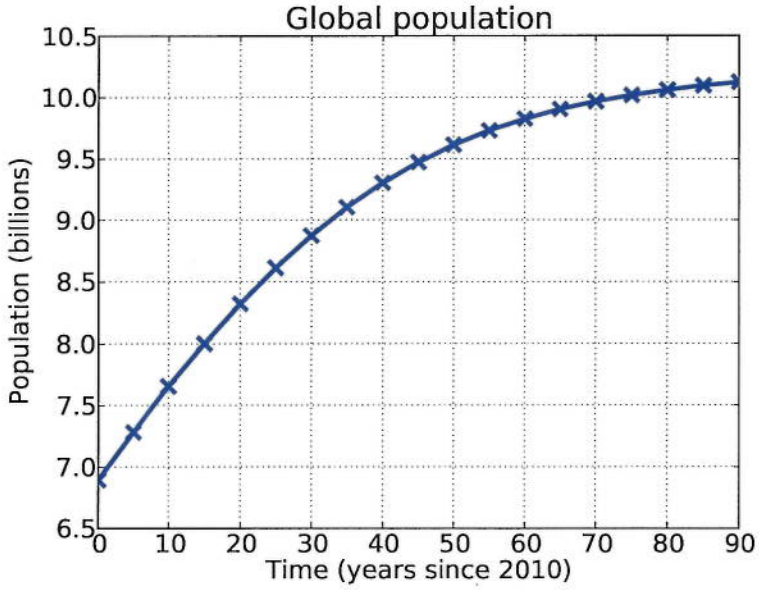


Figure 10.10: Predicted global population.

(a) Is it reasonable to model $P(t)$ using a logistic DE, $P' = rP\left(1 - \dfrac{P}{K}\right)$?

Not sure if sigmoidal (can't see before 2010)
Shows asymptote approach to K
Probably reasonable assuming logistic

(b) Estimate the value of K in the model.

$K \approx 10.2$ billion

(c) Discuss the statement: 'K is the carrying capacity of Earth for humans'.

- K is based on resource limitation
- The Earth has finite resources - only a fixed amount of arable land and fresh H_2O
- Competition for these resources will ↑

BUT social factors (# kids/♀)
medical " " (longevity, contraception)

Question 10.4.9

Figure 10.11 shows atmospheric CO_2 concentrations (that is, the Keeling curve) from 1958 to 2013, along with the global population from 1958 to 2013, and the UN projected global population over the next 100 years, assuming 'medium fertility' (units for CO_2 concentrations and numbers of people do not match; the graphs have been drawn this way for broad comparison). If the UN projections are correct, how would you expect atmospheric CO_2 concentrations to change over that time period? (In other words, extrapolate the Keeling curve for 100 years.) Justify your answer.

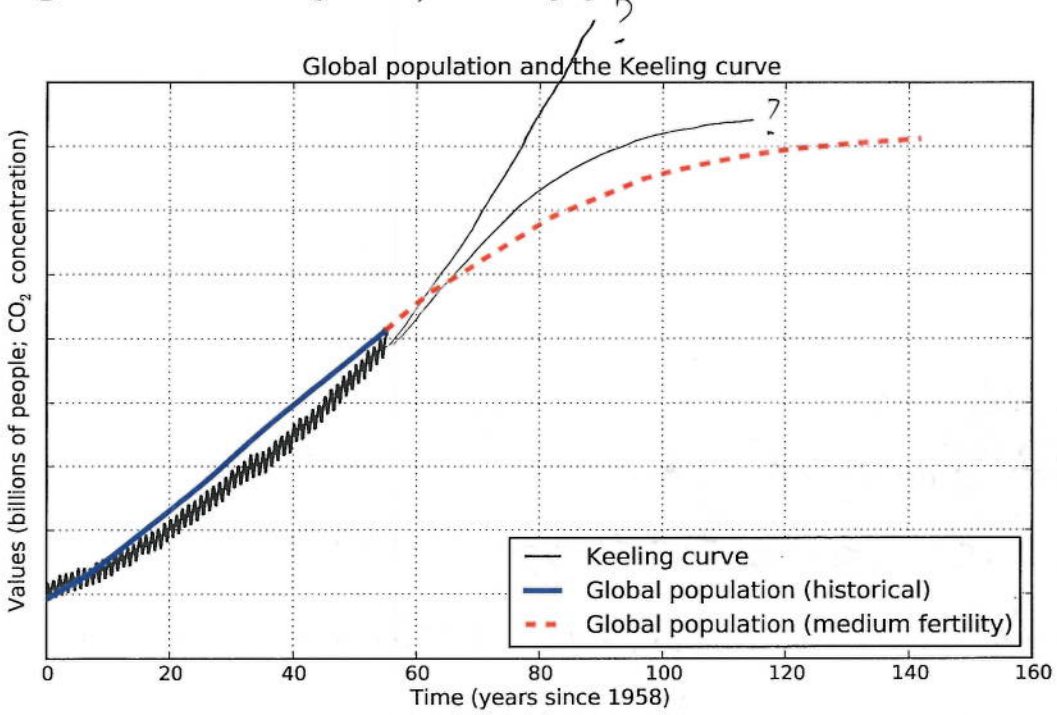


Figure 10.11: Historical and projected global population and the Keeling curve.

We don't know whether $[CO_2]$ will follow human population

If yes, then $[CO_2]$ increase will peak, slowing $[CO_2]$ growth

If no, because developing nations develop, it could continue exponentially

End of Case Study 31: Overpopulation annoys us all.

Chapter 11: **Systems of DEs**

On the farm, every Friday
On the farm, it's rabbit pie day.
So, every Friday that ever comes along,
I get up early and sing this little song
Run rabbit - run rabbit - Run! Run! Run!
Don't give the farmer his fun! Fun! Fun!
He'll get by
Without his rabbit pie
So run rabbit - run rabbit - Run! Run! Run!

Artist: *Flanagan and Allen.* Song: *Run rabbit run.*

Image 11.1: *The wild hunt: Åsgårdsreien* (1872), Peter Nicolai Arbo (1831 – 1892), Nasjonalgalleriet, Oslo. (Source: en.wikipedia.org)

11.1 Introduction to systems of differential equations

- The DE models we have studied so far have all modelled a single, distinct phenomenon.
- Often, multiple factors interact, requiring more sophisticated models.
- For example:
  - in predator-prey relationships, changes in population sizes of *two* species are interrelated;
  - in species with multiple distinct life stages, changes in the population sizes within each stage depend on the numbers in other stages; and
  - the rates at which epidemics spread through populations are influenced by the number of infected individuals **and also** by the number of susceptible individuals.
- Typically, models for these more complex situations use a *system* of DEs (that is, more than one DE).
- Just as with single DEs, analytical solutions exist for some systems of DEs, but other systems require approximate solutions.
- Euler's method can be used to solve a system of DEs approximately, by applying a single iteration to each equation in turn, and then repeating.

Photo 11.1: Predators: Siberian tigers, *Panthera tigris altaica*. (Source: PA.)

11.2 Going through a difficult stage

- We previously modelled populations using exponential and logistic DEs. In each case we assumed that populations were *homogeneous*; that is, every individual in the population had an identical impact on population growth.
- Many organisms have different life stages, each with substantial differences in typical survival rates and reproduction rates.
- For example, in many species, small juveniles have a low survival rate and do not reproduce, whereas mature individuals have a high survival rate and typically do reproduce.
- Hence, simple models based on single DEs are inaccurate for more advanced organisms, particularly those with long life spans. In such cases, systems of DEs give rise to better models.
- In one type of model, populations are classified into groups based on their *life stages*, such as *juvenile* or *breeding adult*.
- Rather than applying a constant growth rate to every individual in the population, a system of DEs includes:
  - the *distribution* of the population within the distinct groups;
  - differing rates of *reproduction* and *death* within groups; and
  - the *transitions* of individuals between groups.
- *Life-cycle diagrams* are useful aids to writing the equations in a system of DEs. These diagrams show the rates of *transition* between stages.

Life-cycle diagram

Life-cycle diagrams represent all possible transitions between stages in the life-cycle of an organism. Each stage is represented as a circle in the diagram, with a directed arrow joining Stage A to Stage B whenever it is possible for an individual to transition from Stage A to Stage B. Each arrow has an associated number, which is the *rate* of transition.

- The general form of a stage in a life-cycle diagram is shown in Figure 11.1. Not all stages will have all of these arrows, as some particular transitions may not be possible.

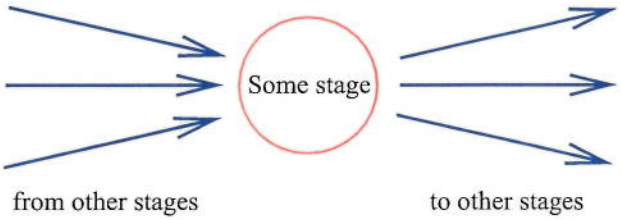


Figure 11.1: A stage in a life-cycle diagram, showing some transitions.

- In order to draw the life-cycle diagram for an organism, we need to know:
  - the number of stages;
  - all possible transitions to and from each stage, including:
    - reproduction;
    - transitions due to the passage of time, or other factors; and
    - deaths.
  - the number or probability associated with each possible transition.
- Once we have drawn a life-cycle diagram, it is usually easy to write a system of DEs for the number of individuals in each stage.

***Question* 11.2.1**

Consider an idealised fish species with two distinct life stages: juvenile and adult. Each month, on average:

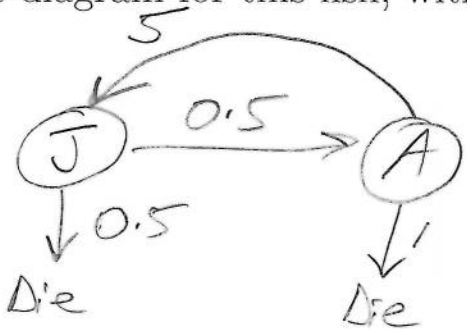
- Juveniles do not breed, have a 50% probability of surviving to adulthood, and a 50% probability of dying.
- Adults produce 5 offspring, and then die.

(continued over)

Question 11.2.1 *(continued)*

Photo 11.2: Bighead Gurnard Perch (*Neosebastes pandus*). (Source: DM.)

(a) Draw a life-cycle diagram for this fish, with juvenile and adult stages.



(b) Let the populations of juveniles and adults at any time be $J(t)$ and $A(t)$ respectively. Write a system of DEs for these populations.

	J	A
$J' =$	$-0.5J - 0.5J = -J$	$+5A$
$A' =$	$= 0.5J$	$-A$

So $J' = -J + 5A$

$A' = 0.5J - A$

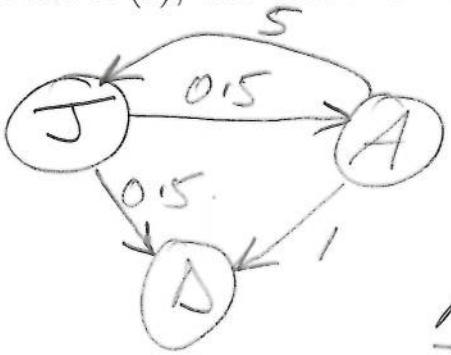
(continued over)

***Question** 11.2.1 (continued)*

(c) Assume that a specific population comprises 20 juveniles and 2 adults at time $t = 0$ months. Use Euler's method and a step size of 1 to estimate the number of fish in each stage at time $t = 2$ months.

t	J_{old}	A_{old}	$J' = -J + 5A$	$A' = 0.5J - A$	$J_{new} = J_{old} + J'h$	$A_{new} = A_{old} + A'h$
0	20	2	−10	8	10 $= 20 + (-10 \times 1)$	10 $= 2 + (8 \times 1)$
1	10	10	40	−5	50 $= 10 + (40 \times 1)$	5 $= 10 + (-5 \times 1)$
2	50	5				

(d) Sometimes it is convenient to include death as a stage in a life-cycle diagram. Draw a life-cycle diagram with three stages, including death, and write a DE for $D(t)$, the total number of dead fish at any time.



$D' = +0.5J + A$

Note: No arrow out of dead!

Case Study 32: **Total turtle turmoil**

Image 11.2: Loggerhead sea turtle. (Source: en.wikipedia.org.)

Photo 11.3: Sea turtle species. (Source: DM.)

- The loggerhead sea turtle (*Caretta caretta*) is a large marine turtle, reaching a length of around 1 m and a mass of more than 100 kg.
- The species is distributed throughout temperate, subtropical and tropical regions, and nests in a number of countries, including Australia.
- Individuals often live for more than 50 years.
- The species is listed as threatened, largely due to human activity, so is likely to become endangered within the foreseeable future.
- Ecologists have studied these turtles in detail, in order to better understand how populations change over time, to investigate possible management strategies and predict the impacts of further environmental change and human activity.
- Researchers in [4] and [7] found that these turtles move through seven distinct stages during their life cycle, and developed a population model based on these stages.
- (For interest, the researchers used a matrix model rather than a system of DEs. However, such models are equivalent to using a system of DEs and Euler's method with a step size of 1.)

- We will study a simplified version of their model, with the seven stages collapsed into three for ease of calculation.
- Figure 11.2 shows the life stages used for the simplified model, along with the estimated proportion of the total turtle population, and the global number of individuals, in each stage.

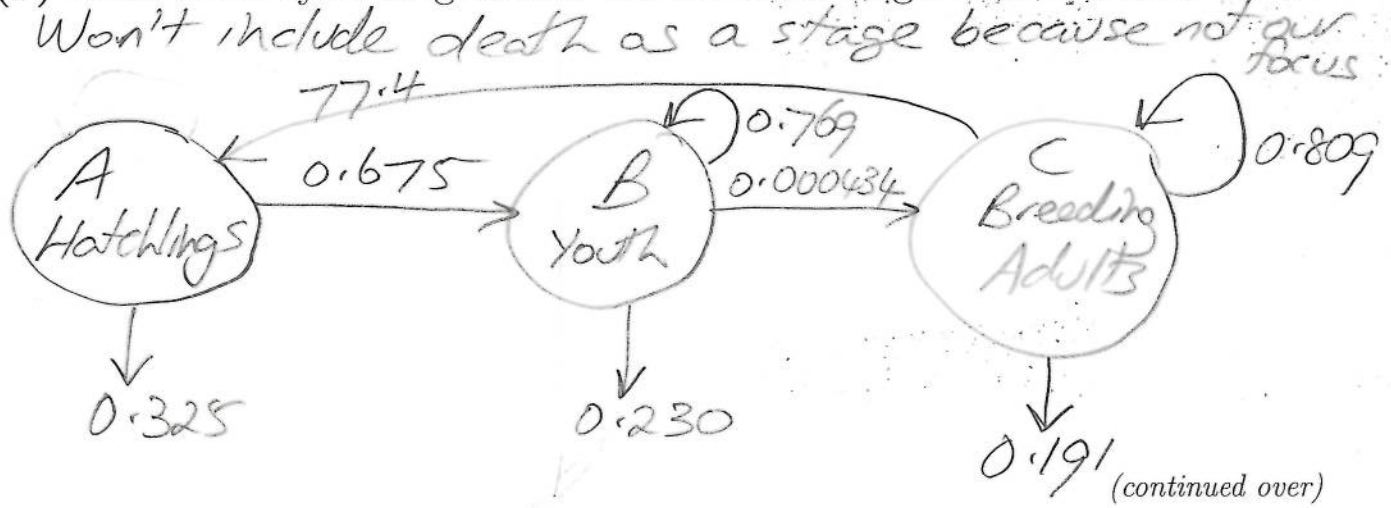
Stage	Description	Age (years)	Proportion	Global population
A	hatchlings	< 1	0.20651	1445570
B	youth	$1 - 23$	0.79097	5536790
C	breeding adult	$24 - 54$	0.00252	17640

Figure 11.2: Loggerhead sea turtles classified into three life stages.

- Each year, turtles transition with the following (rounded) probabilities:
  - Hatchlings become youths (probability $p = 0.675$) or die ($p = 0.325$).
  - Youths become breeding adults ($p = 0.000434$), remain in the youth stage (probability $p = 0.769$), or die ($p = 0.230$).
  - Breeding adults produce new hatchlings (77.4 per adult), and either remain as breeding adults ($p = 0.809$) or die ($p = 0.191$).
- The estimated global population across all life stages was 7 million.

Question 11.2.2

(a) Draw a life-cycle diagram for the three life stages of this turtle.



(continued over)

***Question** 11.2.2 (continued)*

(b) Write a system of DEs for the turtle population.

	A	B	C
A'	$-0.675A - 0.325A = -A$	$0B$	$+77.4C$
B'	$+0.675A$	$-0.230B - 0.000434B = -0.230434B$	$0C$
C'	$0A$	$+0.000434B$	$-0.191C$

(c) Use Euler's method with a step size of 1 year to estimate the number of hatchlings after one year.

$A_{new} = A_{old} + A'h \quad h=1 \quad A_{old} = 145570$

$A_{new} = 145570 + (-145570 + 77.4 \times 17640) \times 1$

$= 1365336.$

Homework: Calculate #s after 1 year for Youth & Breeding Adults

(d) Using Euler's method, after one year there are 5236685 turtles in the youth stage and 16673 breeding adults. Interpret these results.

$Y_{new} = Y_{old} + Y'h$

$A_0 = 1,445,570$ $\quad A_1 = 1365,336$

$B_0 = 5,536,790$ $\quad B_1 = 5,236,685.$

$C_0 = 17,640$ $\quad C_1 = 16,673$

∴ All stages declining!

This population can be modelled using a computer program.

Program specifications: Develop a Python program that uses Euler's

method with a step size of 1 to model the turtle population for 30 years.

Program 11.1: Turtles

```
# Uses Euler's method to model the turtle population.
from pylab import *

# Initialise variables.
maxt = 30
Apops = zeros(maxt+1)
Bpops = zeros(maxt+1)
Cpops = zeros(maxt+1)
Apops[0] = 1445570
Bpops[0] = 5536790
Cpops[0] = 17640
stepsize = 1

# Step through Euler's method for 30 years.
i = 1
while i < (maxt+1):
    dA = -Apops[i-1] + 77.4 * Cpops[i-1]
    dB = 0.675 * Apops[i-1] - 0.230434 * Bpops[i-1]
    dC = 0.000434 * Bpops[i-1]- 0.191 * Cpops[i-1]
    Apops[i] = Apops[i-1] + stepsize * dA
    Bpops[i] = Bpops[i-1] + stepsize * dB
    Cpops[i] = Cpops[i-1] + stepsize * dC
    i = i + 1

# Output the graph.
times = arange(0, maxt+1)
plot(times, Apops, "bx", markersize=8,mew=2,label='Stage A')
plot(times, Bpops, "r+", markersize=8,mew=2,label='Stage B')
plot(times, Cpops, "gs", markersize=6,mew=2,label='Stage C')
plot(times, Apops+Bpops+Cpops, "ko", markersize=6,mew=2,label='
   Total')
xlabel("Time (years)")
ylabel("Number of turtles")
title("Turtle population")
grid(True)
legend()
# savefig('turtlepop.pdf',format='pdf')
show()
```

Figure 11.3 shows the output from running the program.

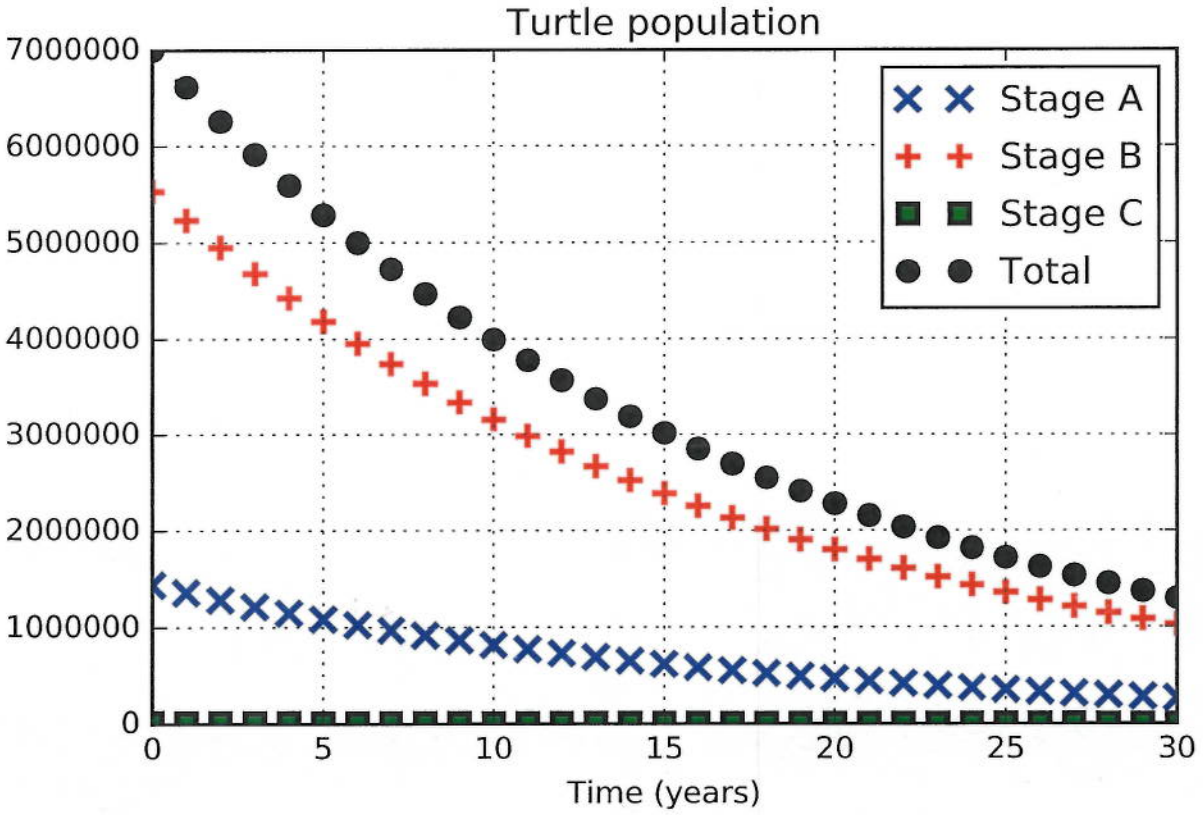


Figure 11.3: Turtle population modelled using Euler's method.

Question 11.2.3

Researchers and authorities have proposed various conservation strategies for the sea turtle; see [8]. Briefly discuss some possible strategies, and explain how the population model would change to reflect them.

Hatcheries ↑A (↑ effective feundity)

Turtle Exclusion Devices in nets ↑B ↑C (↓death)

Ban hunting " "

Remove foxes/dogs on beaches ↑A (↓death)

Remove lights on beaches ↑A (↓death)

End of Case Study 32: Total turtle turmoil.

11.3 Eat or be eaten

- In addition to modelling individual organisms with multiple life stages, systems of DEs can also model interactions between multiple species.
- For example, the classical *predator/prey* problem in ecology considers what happens to the populations of two species when one preys on the other.
- In laboratory situations there is control over these interactions. In nature, inter-species interactions are highly complex. We will first investigate a controlled example, then model a real interaction. The controlled example is very simple, but is not completely unreasonable.

Photo 11.4: Left: skeleton of *Tyrannosaurus rex*. Right: skeleton of *Triceratops horridus*. (Source: PA.)

Case Study 33: **It's just not cricket**

- One method of predicting what may happen in a real-world situation is to simulate it in a laboratory.
- Unpredictable phenomena complicate and impact predator/prey interactions in nature. However, controlled laboratory simulations can give valuable insight into real situations.
- Consider a controlled, time-compressed laboratory experiment simulating the effects of immigration, emigration, births and deaths on populations of frogs (predators) and crickets (prey).

- Initially the experiment comprises 60 frogs and 400 crickets. Each day:
  - 15 crickets are introduced into the experiment (modelling immigration and birth of crickets);
  - 25% of the frogs each eat a cricket (death of crickets);
  - 12 frogs are removed (modelling emigration and death of frogs); and
  - for each 25 crickets present, one new frog is introduced (modelling birth and immigration of frogs based on available food resources).

Photo 11.5: Left: Striped burrowing frog, *Litoria alboguttata*. Right: cricket. (Source: DM.)

Question 11.3.1

Let $F(t)$ and $C(t)$ be the populations of frogs and crickets at time t in days.

(a) Write DEs involving the rate of change of **each** of the populations.

$\frac{C}{25} \rightarrow (F) \rightarrow 12$ $\qquad$ $15 \rightarrow (C) \rightarrow \frac{F}{4}$

$$F' = \frac{C}{25} - 12$$

$$C' = 15 - \frac{F}{4}$$

(b) Using differentiation it can be shown that solutions to the DEs are:

$$F(t) = 40\sin 0.1t + 60 \qquad C(t) = 100\cos 0.1t + 300.$$

$\frac{d}{dt}(\sin t) = \cos t$ $\qquad$ $\frac{d}{dt}(\cos t) = -\sin t$

(continued over)

***Question 11.3.1** (continued)*

Figure 11.4 shows $C(t)$. On the graph, sketch $F(t) = 40 \sin 0.1t + 60$.

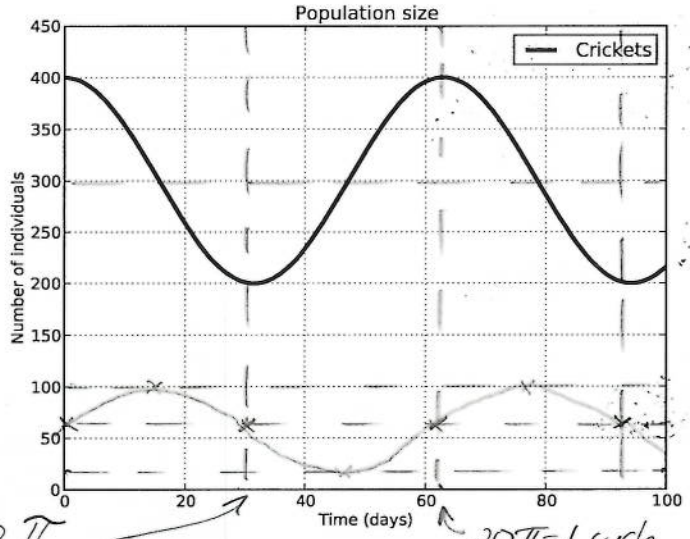


Figure 11.4: Population of crickets.

F = centre value
= 60
A = amplitude
= 40
$\frac{2\pi}{P} = 0.1$
$\Rightarrow P$ = period
$= \frac{2\pi}{0.1} = 20\pi \approx 62.$
10π = ½ cycle
20π = 1 cycle

(c) Interpret the population dynamics. One population is "leading" and the other "lagging". Identify which is which, and explain your answer.

C = Cricket = leading = prey
F = Frog = lagging = predator
C↑ followed by F↑
Then C↓ F↓

Photo 11.6: Left: genuine KTF. Right: fossilised cricket. (Source: PA.)

End of Case Study 33: It's just not cricket.

- Now we will develop a more realistic predator/prey model. In general, such models are based on the assumptions that:
  - the prey has no other predators, and the predator no other prey;
  - there is no significant change to the environment or species' genetics;
  - the prey species is not resource limited, so breeds rapidly and individuals **do not** compete with each other;
  - the predator species is resource limited, so breeds more slowly and individuals **do** compete with each other.

Question 11.3.2

Let W be a population of wolves (predators) and E be a population of elk (prey). How are the values of the *rates of change* of each of W and E influenced by the values of W and E, using the assumptions above? In each case, identify whether the impact is positive or negative.

(a) E' is: positively/negatively dependent on the value of E. Breeding

(b) E' is: positively/negatively dependent on the value of W. Predation

(c) W' is: positively/negatively dependent on the value of E. Food

(d) W' is: positively/negatively dependent on the value of W. Competition

Photo 11.7: Left: gray wolf (*Canis lupus*). Right: elk (*Cervus canadensis*). (Source: PA.)

- The best-known predator/prey model is the **Lotka-Volterra** model.

Lotka-Volterra model

Let $P(t)$ and $Q(t)$ be the population sizes of a predator (for example, wolf) and prey (for example, elk) species respectively, at any time t. The following system of DEs forms the *Lotka-Volterra model*:

$$\begin{aligned} Q' &= aQ - bPQ \\ P' &= -cP + dPQ \end{aligned}$$

where a, b, c and d are positive constants whose values depend on various characteristics of the species and their physical interactions.

Question 11.3.3

Carefully explain the meaning of each term in the Lotka-Volterra equations. In particular, explain the physical relevance of the terms involving PQ.

Prey $Q' = aQ - bPQ$

aQ: constant × # prey = Exponential growth

$-bPQ$: constant × interaction = Loss of prey thru predation = Interaction term

Predator $P' = -cP + dPQ$

$-cP$: constant × # predators = Exponential decay (competition)

dPQ: constant × interaction = Conversion of prey (after its capture) to predator numbers

Case Study 34: **Snowshoe hares and Canadian lynx**

Image 11.3: Canadian lynx chasing a snowshoe hare. (Source: www.animalspedia.com.)

- The Canadian lynx, *Lynx canadensis*, is a member of the feline family distributed predominantly in Canada and Alaska. Lynx are carnivorous, with individuals weighing 8 to 15 kg, and living for up to 15 years.
- The primary food source (up to 95%) of the Canadian lynx is the snowshoe hare, *Lepus americanus*. The hare has large hind feet (for moving on snow) and turns white in winter.
- People have hunted these lynx and hares for their fur for many years. Harvest records dating from the 1730s allow long-term population estimates.
- Figure 11.5 (from [31]) graphs these data over 90 years, and shows a series of reasonably regular fluctuations in the sizes of both populations. Note the similarity to the periodic population movements in the laboratory-controlled predator/prey relationship between frogs and crickets.

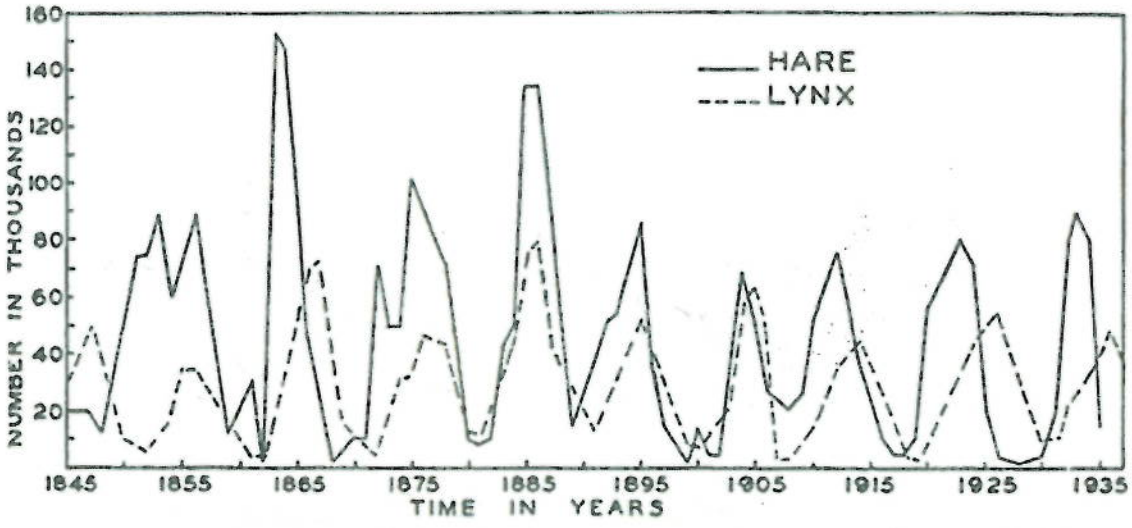


Figure 11.5: Numbers of Canadian lynx and snowshoe hares. (Source: [31].)

Question 11.3.4

Let $L(t)$ and $H(t)$ be the populations of lynx (predators) and hares (prey) respectively, **in thousands**. The Lotka-Volterra equations are:

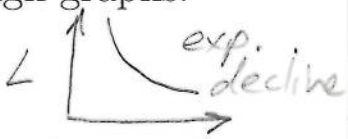
$$H' = aH - bHL \qquad L' = -cL + dHL$$

(a) How would "the lynx population has become extinct" be written mathematically?

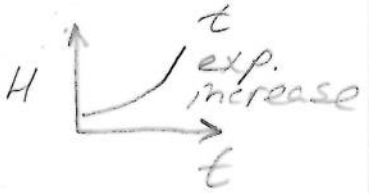
$L = 0$

(b) If either population suddenly became extinct, what does the model predict will happen to the other population? Sketch rough graphs.

if $H=0$ $\Rightarrow$ $H'=0$ & $L'=-cL$



if $L=0$ $\Rightarrow$ $H'=aH$ & $L'=0$



(c) What would you expect to happen in reality?

$H=0$: Lynx starve, but could switch to alternate prey

$L=0$: Hare ↑ but then limited by resources (eg. food) so could reach a carrying capacity

(d) If the units of H are hares and L are lynx and time is measured in years then find the units of a, b, c and d.

$H' = aH - bHL$

$\text{hares/yr} = \left(\frac{1}{\text{yr}}\right)\text{hares} - \left(\frac{1}{\text{yr.lynx}}\right)\text{hares.lynx}$

$L' = -cL + dHL$

$\text{lynx/yr} = \left(\frac{1}{\text{yr}}\right)\text{lynx} + \left(\frac{1}{\text{yr.hares}}\right)\text{hares.lynx}$

Example 11.3.5

Figures 11.6 and 11.7 show data from the Canadian Government and the Hudson's Bay Company, estimating the populations of hare and lynx in part of their range from 1900 to 1920. (All populations are in thousands.)

Year	Hares	Lynx	Year	Hares	Lynx	Year	Hares	Lynx	Year	Hares	Lynx
1900	30	4	1905	20.6	41.7	1910	27.1	7.4	1915	19.5	51.1
1901	47.2	6.1	1906	18.1	19	1911	40.3	8	1916	11.2	29.7
1902	70.2	9.8	1907	21.4	13	1912	57	12.3	1917	7.6	15.8
1903	77.4	35.2	1908	22	8.3	1913	76.6	19.5	1918	14.6	9.7
1904	36.3	59.4	1909	25.4	9.1	1914	52.3	45.7	1919	16.2	10.1
									1920	24.7	8.6

Figure 11.6: Populations of lynx and hares (in thousands).

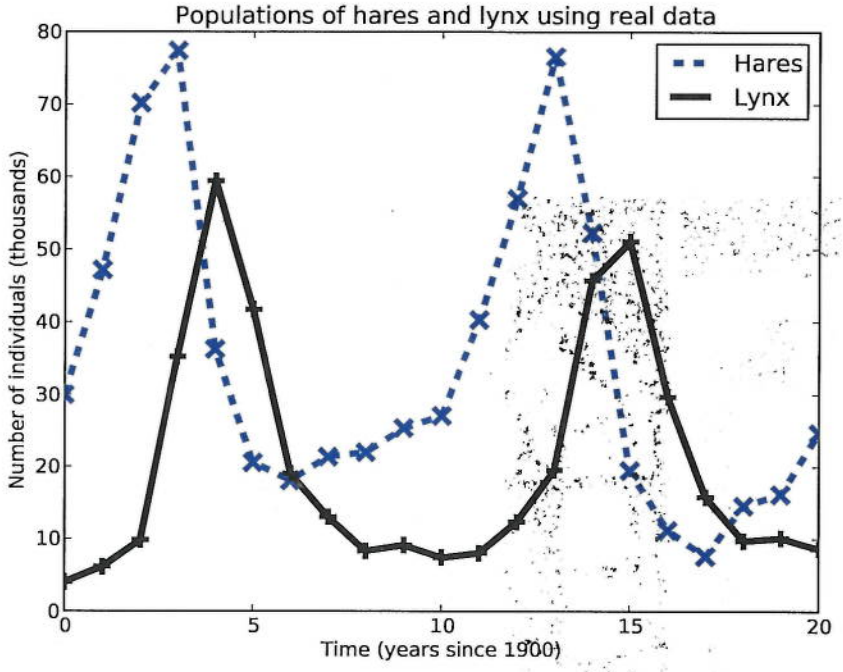


Figure 11.7: Graph of the populations of lynx and hares (in thousands).

- Experimentation and analysis show that for this time period, reasonable values for the constants a, b, c and d in the Lotka-Volterra equations are: $a = 0.484$, $b = 0.028$, $c = 1$ and $d = 0.032$ (in appropriate units).

Now we can use Euler's method to model the population sizes.

Program specifications: Develop a Python program that uses Euler's method with step size of 0.1 year to model the populations of lynx and hares.

Program 11.2: Lotka-Volterra model of hares and lynx.

```
# Uses Euler's method and Lotka-Volterra equations to model
# populations of lynx and hare from 1900 to 1920.
from pylab import *

# Initialise variables for Euler's method.
ss=0.1
time=arange(0,20.1,ss)
a = 0.484
b = 0.028
c = 1
d = 0.032
H = zeros(size(time))
L = zeros(size(time))
H[0] = 30.0
L[0] = 4.0
v=56
nn=size(time)

# Step through Euler's method with stepsize ss.
# Repeatedly calculate derivatives then update the 'next'
   values.
i = 0
while i < nn-1:
    dH =  a*H[i] - b*H[i]*L[i]
    dL = -c*L[i] + d*H[i]*L[i]

    H[i+1] = H[i] + ss*dH
    L[i+1] = L[i] + ss*dL
    i = i+1

# Output graphs.
text(4,7,"H(t)")
text(15.1,60,"L(t)")
xlabel("Time (years since 1900)")
ylabel("Number of individuals (thousands)")
title("Modelled populations of hares and lynx")
plot(time, H, "b-", linewidth=3)
plot(time, L, "k-", linewidth=3)
show()
```

Example 11.3.6

- At time $t = 0$ years (corresponding to year 1900), data show that there were 30 (thousand) hares and 4 (thousand) lynx in the monitored region.
- Figures 11.8 and 11.9 compare the modelled population sizes over 20 years with the real (measured) data for each population.

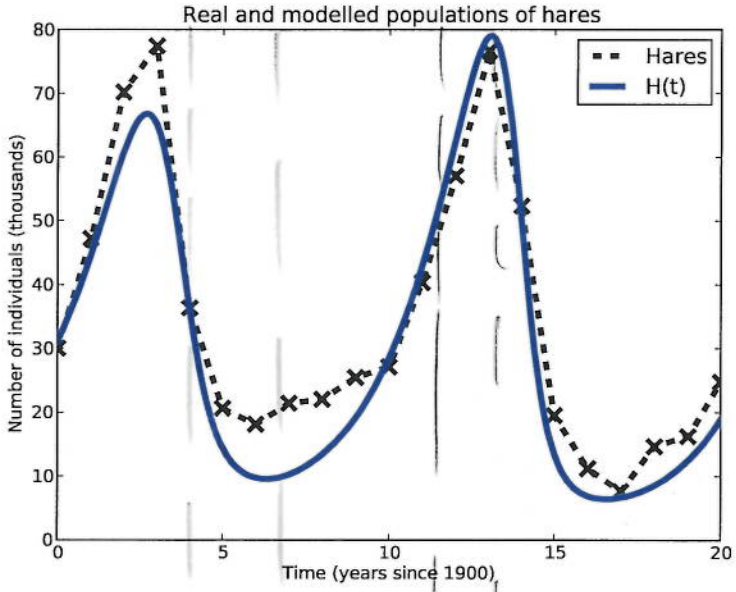


Figure 11.8: Real and modelled hare populations.

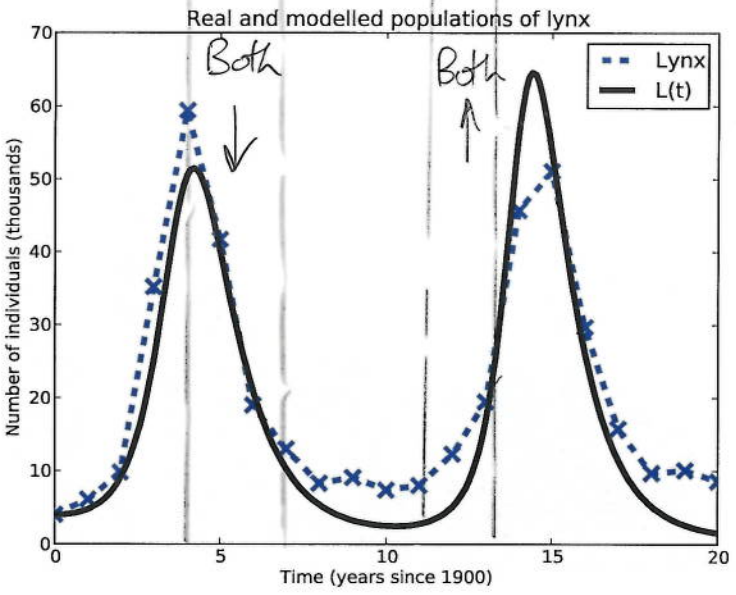


Figure 11.9: Real and modelled lynx populations.

Question 11.3.7

(a) Comment on the results in Example 11.3.6.

- Good model in time period investigated
- Captured leading H & lagging L
- Not sin/cos but can see why they are sometimes used for oscillations

(b) Critically evaluate the following possible media statement:

> *A survey has shown that the populations of lynx and snowshoe hares are both in decline. We need to act promptly or else one or both species will become extinct.*

- There are times when both ↓ & times when both ↑
- Does not make sense to talk about both declining when it's a cycle
- Be careful making sweeping statements from a short time series

Photo 11.8: Three top predators. Left: polar bear, *Ursus maritimus*. Centre: Komodo dragon, *Varanus komodoensis*. Right: Siberian tiger, *Panthera tigris altaica*. (Source: PA.)

End of Case Study 34: Snowshoe hares and Canadian lynx.

Chapter 12: **Fully sick**

From New Delhi to Darjeeling I have done my share of healing,
and I've never yet been beaten or outboxed.
I remember that with one jab of my needle in the Punjab
how I cleared up beriberi and the dreaded dysentery,
but your complaint has got me really foxed.
Oh doctor, touch my fingers.
Well, goodness gracious me.
You may be very clever but however, can't you see,
my heart beats much too much at a certain tender touch,
it goes boom boody-boom boody-boom boody-boom
boody-boom boody-boom boody-boom-boom-boom.

Artist: *Peter Sellers and Sophia Loren.* Song: *Goodness gracious me.*

Image 12.1: *The Triumph of Death* (1562), Pieter Bruegel the Elder (c. 1525 – 1569), Museo del Prado, Madrid. (Source: commons.wikimedia.org)

12.1 Epidemics and SIR models

- In this section, we will use systems of DEs to model the large-scale spread of communicable disease through a population over time.

Epidemic

A large-scale occurrence of disease in a human population is called an **epidemic** if new cases of the disease arise at a rate that "substantially exceeds what is expected" in a given time period. Localised occurrences are called **outbreaks**, and global occurrences are often called **pandemics**.

- Modelling diseases is important to understanding how they spread, and how their impact may be mitigated through approaches such as quarantine, vaccination and public health campaigns.

Photo 12.1: Images commemorating the bubonic plague in Eyam, the "Plague Village", UK. Left: stained glass 'Plague Window'. Centre: first page of the list of names of villagers who died from plague in 1665–6. Right: tombstone. (Source: PA.)

- Modelling disease spread often begins with estimates of the number of secondary infections that typically arise from an individual with the disease, and the rate at which individuals recover from the disease.

Basic reproduction number and infectious period

The **infectious period** of a disease is the average length of time during which an infective individual can infect a susceptible individual. Many diseases are infectious before symptoms become apparent.

The **basic reproduction number** R_0 of a disease is the average number of secondary infections caused by a single infective individual in a completely susceptible population, in the absence of any preventive interventions.

The value of R_0 is determined by factors including how infectious the disease is, how it is spread and the duration of the infectious period.

- Figure 12.1 gives information for some well-known communicable diseases.

Disease	Transmission method	R_0	Infectious period
Rubella	Airborne droplet	≈ 5	2 weeks
Measles	Airborne droplet	$12-18$	10 days
Whooping cough	Airborne droplet	$12-17$	3 weeks
Mumps	Airborne droplet	$4-7$	14 days
Swine flu	Airborne droplet	$1.4-1.6$	6 days
Seasonal influenza	Airborne droplet	$2-3$	6 days
Polio	Faecal/oral	$5-7$	$6-20$ days
HIV/AIDS	Sexual contact	$2-5$	unlimited
Syphilis	Sexual contact	≈ 1.5	up to 2 years
Human papillomavirus	Sexual contact	$1-3$	very variable
Pneumonic plague	Airborne droplet	≈ 1.3	2 days (100% death rate)

Figure 12.1: Transmission methods, infectious periods and values of R_0 for some communicable diseases.

Infection rate and recovery rate

The **infection rate** a is the rate at which secondary infections arise from a single infective individual, and is defined to equal the basic reproduction number divided by the infectious period (IP). Thus, $a = \dfrac{R_0}{IP}$.

The **recovery rate** b is the rate at which an infective individual recovers, and is defined to equal 1 divided by the infectious period. Thus, $b = \dfrac{1}{IP}$.

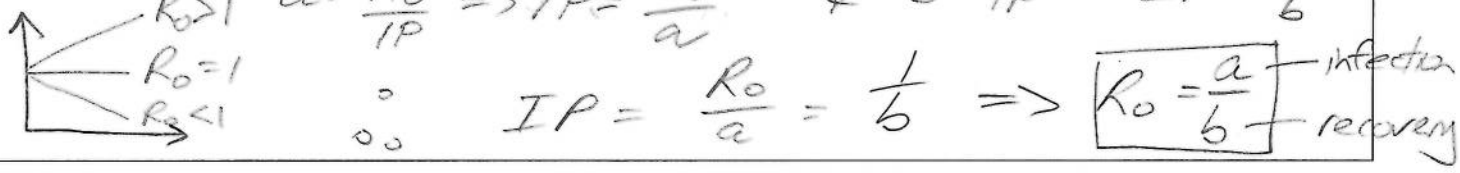
***Example* 12.1.1**

Figure 12.1 shows that for the disease rubella, $R_0 = 5$ and $IP = 2$ weeks.

Thus, $a = R_0/IP = 2.5$ per week and $b = 1/IP = 0.5$ per week. A person with rubella will typically infect 2.5 people per week in a fully susceptible population, and will 'half recover' each week.

***Question* 12.1.2**

Find an expression relating R_0, a and b.



- The SIR (**S**usceptible, **I**nfective, **R**emoved) epidemic model is used to model many diseases, including rubella, measles, cholera and bubonic plague.

SIR model of epidemics

The *SIR* epidemic model classifies a population into three distinct **compartments** or groups, and uses a system of DEs to predict the changes in the number of people in each group. At any time t:

(1) The *susceptible* compartment $S(t)$ is the group of people who are susceptible to the disease.

(2) The *infective* compartment $I(t)$ is the group of people who have the disease and can infect susceptible people.

(3) The *removed* compartment $R(t)$ is the group of people who cannot catch the disease, either because they have permanently recovered, are naturally immune, or have already died from the disease.

The only possible transitions in the SIR model are that: a susceptible person can become infective; and an infective person can become removed.

The model assumes that there are no births, no deaths from any other causes, and that the population mixes homogeneously.

Question 12.1.3

Draw a "life-cycle diagram" for the SIR model, with infection rate a and recovery rate b.

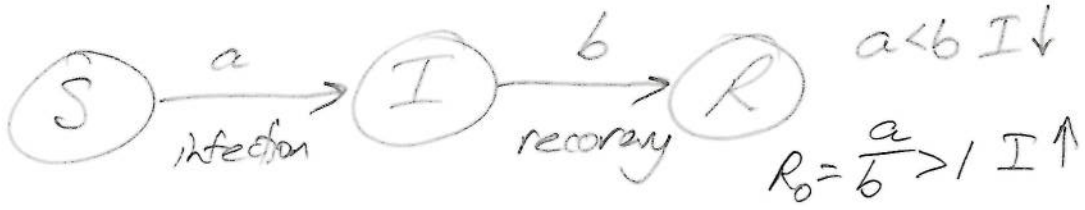


Photo 12.2: Nighttime at New Podgorski cemetery, Krakow, Poland, on All Saints' Day. (Source: PA.)

Effective reproduction number and infection rate

Often, not everyone in a population is susceptible to a disease. The **effective reproduction number**, R_e, is an estimate of the *average* number of secondary infections arising from an infective individual. If a population of size N contains S susceptible individuals then

$$R_e = R_0 \times \frac{S}{N}.$$

Similarly, if a is the infection rate in a fully susceptible population, then the **effective infection rate** in a population that is not fully susceptible is

$$a \times \frac{S}{N}.$$

The equations for the SIR model

If a population of N people at time t is divided into three compartments, susceptible $S(t)$, infective $I(t)$ and removed $R(t)$, then the SIR model is:

$$S' = -a \times \frac{S}{N} \times I$$

$$I' = a \times \frac{S}{N} \times I - bI$$

$$R' = b\,I$$

where a is the infection rate and b is the recovery rate. Note that the population size N remains constant because the numbers of people moving between compartments always balance.

Question 12.1.4

Explain what each of the terms in each of the SIR equations represents.

$S' = -a\frac{S}{N} \times I$ Note: S' always -ve so S decreases

$-a\frac{S}{N}$: Effective infection rate (considering proportion of population susceptible)

I: # infected

$I' = a\frac{S}{N} \times I - bI$

$a\frac{S}{N} \times I$: # infected

bI: Recovery rate × # infected = # people removed

$R' = bI$

bI: # people removed

Note: R' always +ve so R increases

Case Study 35: **Rubella**

- **Rubella** (or **German measles**) was (and in some countries, still is) a common disease, particularly in childhood.
- In most cases, symptoms are very mild, and may even pass unnoticed. However, if a woman is infected during the first 20 weeks of pregnancy then spontaneous abortion can occur (in about 20% of cases), or the child may be born with congenital rubella syndrome (CRS), which is a range of incurable conditions including deafness, blindness and mental retardation.
- The risk of developing CRS in an unborn child is as high as 90% if the mother is infected during the first 10 weeks of pregnancy.
- There was a rubella epidemic in the USA between 1962 and 1965. Data from [39] show that during 1964–65 there were:
  - 12.5 million rubella cases
  - 11000 abortions (spontaneous and surgical)
  - 20000 infants born with CRS (12000 deaf, 3580 blind, 1800 with mental retardation)
- During that epidemic, 1% of all children born in New York were affected.
- A vaccine was introduced in 1969 and is routinely administered in many countries. In Queensland, the Department of Health recommends all children have combined MMR (measles, mumps and rubella) vaccines at the ages of 12 months and 4 years.
- Vaccination campaigns have greatly reduced the incidence of rubella and the frequency of outbreaks. The Centers for Disease Control and Prevention announced that rubella was eliminated from the USA in 2004.
- In January 2008, at least four babies in Sydney became infected with rubella. All were less than 12 months old, so were under the age for vaccination with the MMR vaccine.

Example 12.1.5

Assume that a population of 10000 people contains 10 people infective with rubella, and that everyone else is susceptible. Using the values of a and b from above, the SIR equations for rubella are:

$R_0 = 5$, $IP = 2$

$a = \frac{R_0}{IP} = \frac{5}{2} = 2.5$ ✓

$b = \frac{1}{IP} = \frac{1}{2}$ ✓

$$S' = -2.5 \times \frac{S}{10000} \times I$$

$$I' = 2.5 \times \frac{S}{10000} \times I - 0.5I$$

$$R' = 0.5I$$

where $I(0) = 10$, $S(0) = 9990$ and $R(0) = 0$.

Question 12.1.6

Use Euler's method and a stepsize of one week to estimate the number of people in each category after one week.

$N = 10{,}000$; $S_{old} = 9990$; $I_{old} = 10$; $R_{old} = 0$

$Y_{new} = Y_{old} + Y'h$

$S_{new} = S_{old} + S'h = 9990 + 1 \times \left(-\frac{2.5}{10000} \times 9990 \times 10\right) = 9{,}965$

$I_{new} = I_{old} + I'h = 10 + 1 \times \left(\frac{2.5}{10000} \times 9990 \times 10 - 0.5 \times 10\right)$

$= 30$

$R_{new} = R_{old} + R'h = 0 + 1 \times 0.5 \times 10$

$= 5$

$10\,000$

Now we can develop a computer program to model a rubella epidemic.

Program specifications: Write a Python program that uses Euler's method and the SIR model to predict the spread of rubella with 10 infective individuals in a population of 10000 people.

Program 12.1: SIR model of rubella.

```
# This program uses Euler's method and the SIR equations to
# model the spread of rubella in a population with a proportion
# of between 0 and 1 of the population being vaccinated.
from pylab import *

# Input parameters for the model.
maxt = eval(input("Over how many weeks should the model run? "))

# Initialise variables for rubella; values of a and b are per week.
N = 10000
a = 2.5
b = 0.5

# Initialise variables for Euler's method. The stepsize is 0.1 week.
ss = 0.1
time=arange(0, maxt+0.1, ss)
nn=size(time)
SA = zeros(nn)
IA = zeros(nn)
RA = zeros(nn)

# Set the initial number of people in each category.
IA[0] = 10
SA[0] = N - IA[0]
RA[0] = N - SA[0] - IA[0]

# Step through Euler's method with stepsize ss.
i = 0
while i < nn-1:
    dS = -a * SA[i] * IA[i]/N
    dI = a * SA[i] * IA[i]/N - b * IA[i]
    dR = b * IA[i]
    SA[i+1] = SA[i] + ss*dS
    IA[i+1] = IA[i] + ss*dI
    RA[i+1] = RA[i] + ss*dR
    i = i+1
# Output
xlabel("Time (weeks)")
ylabel("Number of people")
title("SIR model of rubella")
plot(time, SA, "b—", linewidth=5,label="Susceptible")
plot(time, IA, "r-", linewidth=5,label="Infected",)
plot(time, RA, "k-.", linewidth=5,label="Removed")
legend(loc="center right")
grid(True)
savefig('rubella1.pdf',format='pdf')
show()
```

Example 12.1.7

Figure 12.2 shows the program output for a period of 12 weeks.

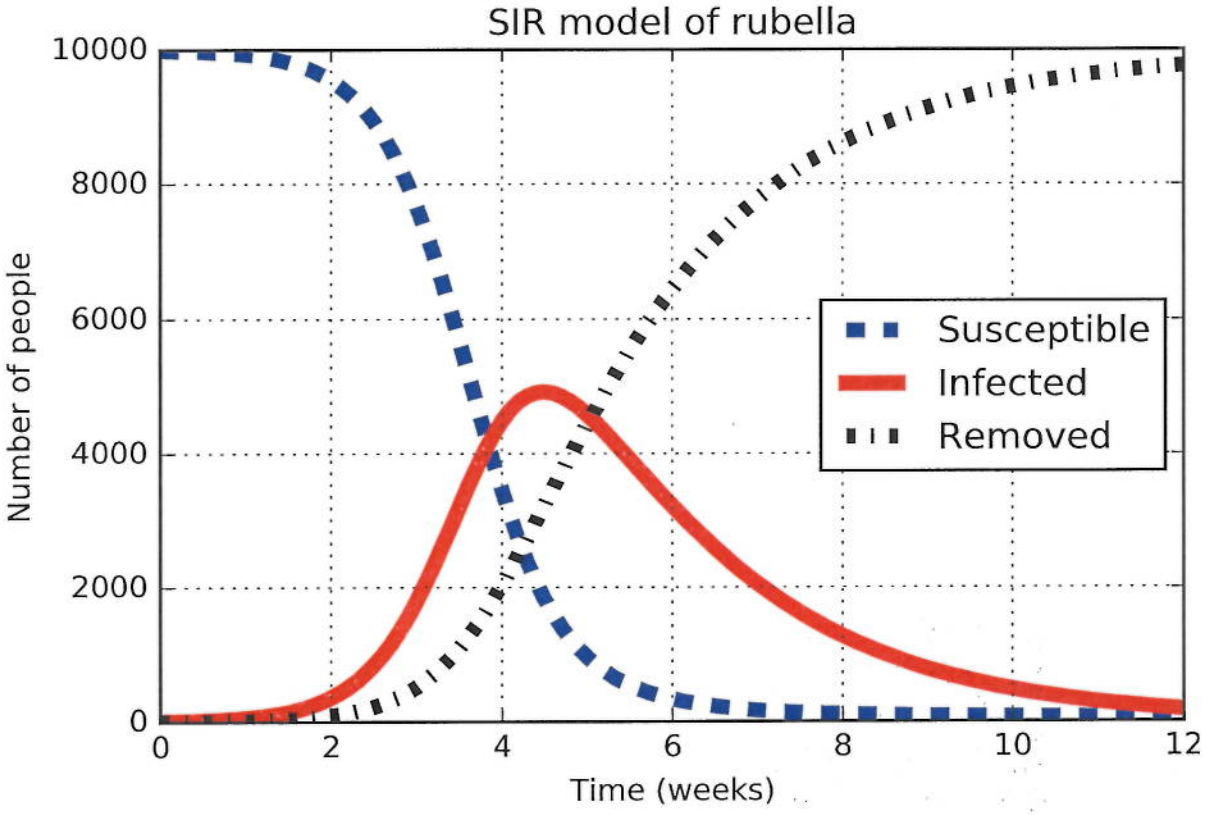


Figure 12.2: A rubella epidemic modelled using Euler's method, showing the numbers of people who are susceptible $S(t)$, infective $I(t)$ and removed $R(t)$.

Figure 12.2 shows that:

- An epidemic occurs. (This is expected, because the population mostly comprises susceptible individuals.)
- The epidemic lasts for about 12 weeks.
- The peak number of infectives at any time is 4925, which occurs about 4.4 weeks after infectives first entered the population.
- Almost everybody becomes infective at some time, although a small number never do.

End of Case Study 35: Rubella.

Case Study 36: **Vaccinations**

- An epidemic occurs if introducing a group of infected people into a population causes an increase in the number of infectives in the population.

Question 12.1.8

If a is the infection rate and b is the recovery rate then the DE for I' is:

$$I' = a \times \frac{S}{N} \times I - bI.$$

(a) Show that an epidemic occurs if the proportion of susceptibles in the population is more than $1/R_0$. (Hint: note that $R_0 = a/b$.)

Epidemic when $I' > 0$

$\therefore a\frac{S}{N}I > bI$

$\Rightarrow a\frac{S}{N} > b$

$\Rightarrow \boxed{\frac{S}{N} > \frac{b}{a} = \frac{1}{R_0}}$ Because $R_0 = \frac{a}{b}$

(b) Explain intuitively **why** an epidemic will occur if a fraction of more than $1/R_0$ of a population is susceptible. (Hint: consider the effective reproduction number, R_e.) $R_e = R_0 \times \frac{S}{N}$

R_0 = Basic reproduction number = average # secondary infections by a single infective. So when people are vaccinated this means they can't be infected. So need $R_e > 1$ for an epidemic (on average infect more than 1 person)

(c) Mass public vaccination aims to vaccinate a certain number of people. What level of coverage do authorities typically aim for? Why?

$\frac{S}{N} < \frac{1}{R_0}$ i.e.

$\Rightarrow R_e = R_0 \times \frac{S}{N} < 1$

$\therefore$ on average, an infected person infects <1 person

Question 12.1.9

Explain why the target vaccination rate for measles is (at least) 95%. What is the figure for rubella? Need $\frac{S}{N} < \frac{1}{R_0}$

Measles: $R_0 = 18$, so $\frac{S}{N} < \frac{1}{18} = 0.05$ ∴ need > 95% vaccination

Rubella: $R_0 = 5$, so $\frac{S}{N} < \frac{1}{5} = 0.2$ ∴ need > 80% vaccination

Question 12.1.10

Should vaccinations be compulsorily enforced? Why or why not?

Yes

- Public good
- Scientifically valid
- Lower death rates

No

- Personal choice
- Interference by Government
- Perceived side effects

Example 12.1.11

In 1998, a paper published in the Lancet (with lead author Dr Andrew Wakefield) claimed to identify a link between the MMR (Measles, Mumps and Rubella) vaccine and autism in children. The findings had a large impact on the public perception of the vaccination. As a result, more than 3 million young people in the UK were not fully vaccinated.

In recent years, the study linking MMR vaccines with autism has been completely discredited. Wakefield was found to have committed scientific fraud by falsifying data, to have acted dishonestly and irresponsibly, and to have a conflict of interest. A number of his research papers were retracted by the journals that had previously published them, and in 2010 he was struck off the UK medical register.

Question 12.1.12

In 1995/6, around 91% of 2 year old children in the UK had received the MMR vaccine. The figure then dropped steadily to 79.9% in 2003/4, and then rose to 92.3% in 2012/13. Figure 12.3 shows the number of laboratory confirmed cases of measles and rubella in England over recent years.

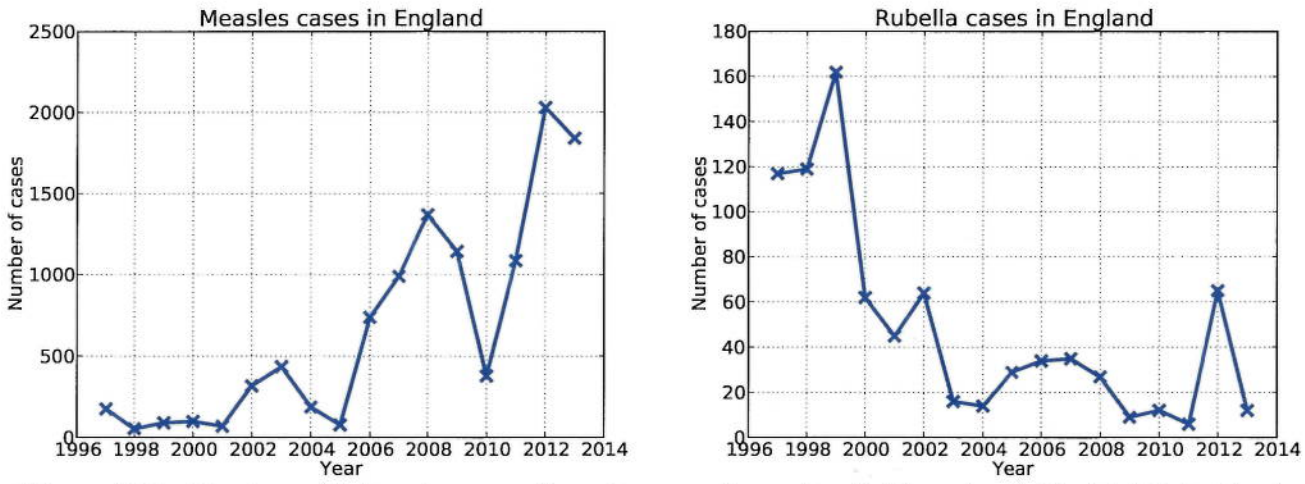


Figure 12.3: Number of laboratory confirmed cases of measles (left) and rubella (right) in England.

(a) Why might the measles graph in Figure 12.3 have that shape?

↑ because need vaccination rates >95% to prevent an epidemic, but it has been below this

(b) What shape do you think the measles graph will have over the next 10–15 years? Why?

Still ↑ but will start to drop off as vaccination rate approaches 95%

(c) Why might the rubella and measles graphs have different shapes?

Rubella is ↓ because vaccination rate generally above 80%. Rubella is less infectious than measles.

Question 12.1.13

How could the SIR model the effects of different vaccination rates?

where c = vaccination rate

Question 12.1.14

Earlier we modelled a rubella epidemic in a city with 9990 susceptible people and 10 infective people. Figure 12.4 shows the predicted numbers of infective people $I(t)$ under five scenarios, with vaccination rates of 0%, 20%, 40%, 60% and 80%. Interpret and explain the graphs. What are the benefits of increased vaccination rates?

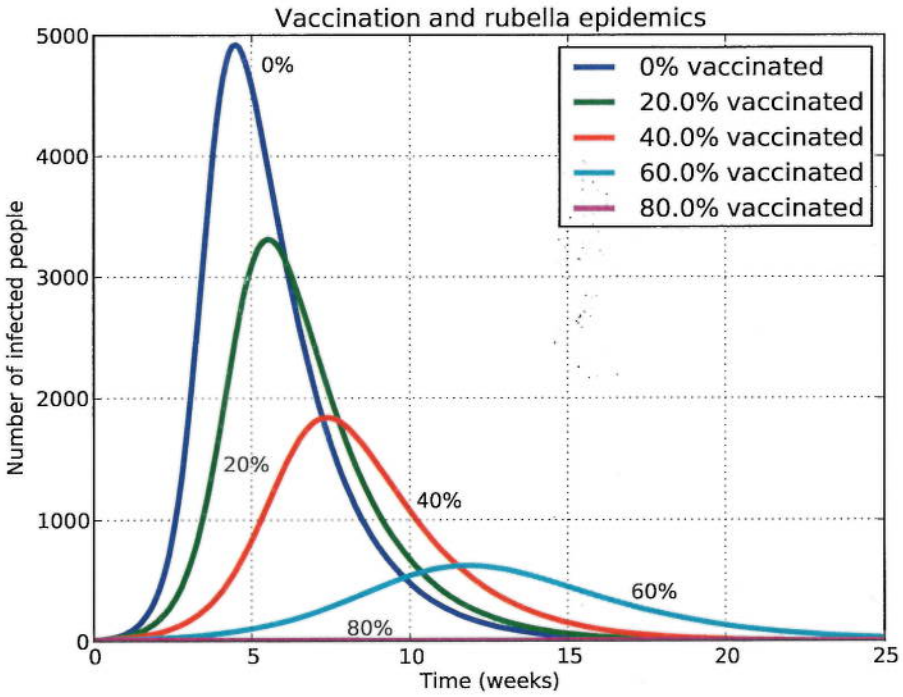


Figure 12.4: The effect of different vaccination rates on a rubella epidemic.

- Easier for medical infrastructure
- Save $
- Reduce epidemics
- Peak later - more time to prepare (quarantine & education & buy vaccines)

End of Case Study 36: Vaccinations.

12.2 Catastrophes

- Many governments conduct *catastrophe planning*, modelling the potential impacts of disastrous events, such as nuclear explosions, terrorist strikes, tsunamis, earthquakes and pandemics. Much of this work is highly secret, partly for security reasons, but also because some of the predicted outcomes are too frightening for public release.
- Historically, there have been many severe pandemics. In 2009, the Australian government spent $200 million responding to the (very mild) swine flu pandemic.

***Example** 12.2.1*

In terms of numbers of fatalities, three of the four worst catastrophes in (European) Australian history are diseases. These four events are:

- Spanish influenza in 1918–19, causing more than 12000 deaths;
- a polio epidemic in 1946–55, causing more than 1000 deaths;
- a naval battle in the Second World War, causing 727 deaths; and
- a bubonic plague epidemic in 1900–1910, causing 550 deaths.

In addition, thousands of indigenous Australians died from communicable diseases introduced by European settlement.

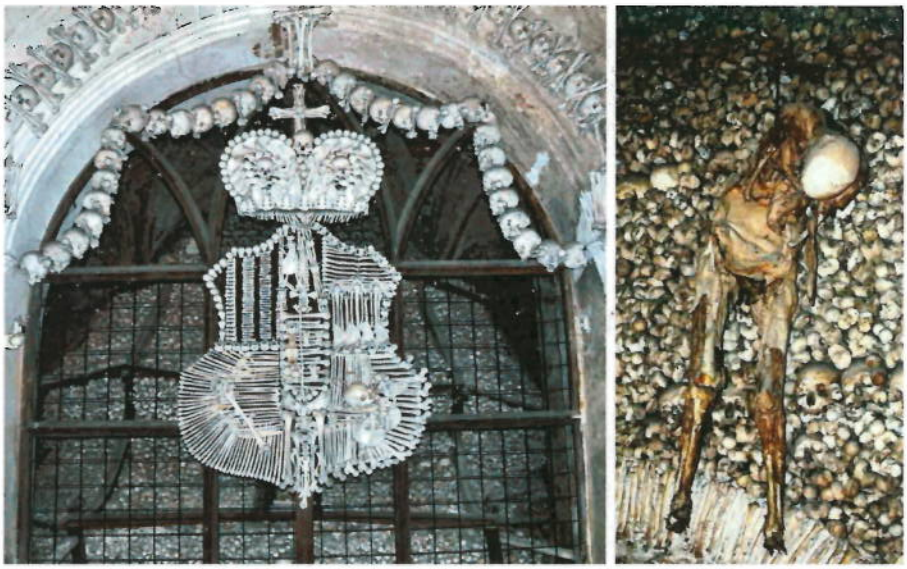


Photo 12.3: Bones of plague victims (and others). Left: Sedlec ossuary, Kutna Hora, Czech Republic. Right: Capela dos Ossos, Evora, Portugal. (Source: PA.)

Example 12.2.2

In the 1300s, the bubonic plague or *Black Death* killed around 20 million Europeans in six years, which was about one third of the population. In the worst affected urban areas, around half of the population died. The plague returned regularly, with around 100 epidemics occurring in the next 400 years. The resulting social, economic, humanitarian and psychological costs and disruption are incalculable, and unimaginable today.

Photo 12.4: Plague monuments. Left: Brno, Czech Republic. Centre: Vienna, Austria. Right: Olomouc, Czech Republic. (Source: PA.)

Since the year 2000, more than 80 probable or confirmed cases of plague have occurred in the USA, and nine people have died. The most recent fatality occurred in July 2015.

Example 12.2.3

A Spanish influenza pandemic occurred in 1918–1919. Within six months the global death toll was 25 million (more than the number who died from combat in the First World War). The flu was so virulent and deadly that it 'burnt itself out', disappearing completely within 18 months. The following was written (by Professor Grist) in an infected camp, 29/9/1918.

(continued over)

Example 12.2.3 (continued)

"These men start with what appears to be an ordinary attack of LaGrippe or Influenza, and when brought to the Hosp. they very rapidly develop the most viscous type of Pneumonia that has ever been seen. Two hours after admission they have the Mahogany spots over the cheek bones, and a few hours later you can begin to see the Cyanosis extending from their ears and spreading all over the face, until it is hard to distinguish the coloured men from the white. It is only a matter of a few hours then until death comes, and it is simply a struggle for air until they suffocate. It is horrible. One can stand it to see one, two or twenty men die, but to see these poor devils dropping like flies sort of gets on your nerves. We have been averaging about 100 deaths per day, and still keeping it up. There is no doubt in my mind that there is a new mixed infection here, but what I don't know...

...It takes special trains to carry away the dead. For several days there were no coffins and the bodies piled up something fierce, we used to go down to the morgue (which is just back of my ward) and look at the boys laid out in long rows. It beats any sight they ever had in France after a battle..."

Case Study 37: **Avian influenza**

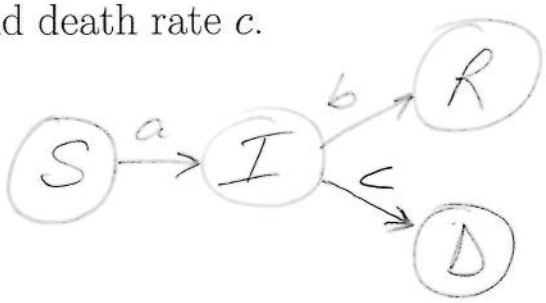
- The World Health Organisation (WHO) has warned that:
  - the risk of an influenza pandemic is high;
  - H5N1 (avian) influenza is endemic in many bird populations;
  - bird-to-human transmission has already caused fatalities; and
  - there is a serious risk that the virus could mutate and become human-to-human transmissible, leading to a "nightmare scenario".

Photo 12.5: Left: Marabou Stork, *Leptoptilos crumeniferus*. Right: Painted Desert, USA. (Source: PA.)

- We will use the **SIRD** model to investigate the potential impact of a catastrophe caused by human-transmissible avian influenza. The SIRD model divides the population into **four** distinct compartments: Susceptible, $S(t)$; Infective, $I(t)$; Recovered, $R(t)$; and Dead, $D(t)$.
- The only possible movements of people *between* compartments are: susceptible people can become infective; infective people can recover or die.

***Question* 12.2.4**

Draw a "life-cycle diagram" for the SIRD model, with infection rate a, recovery rate b and death rate c.



- Because there has never been a verified case of human-to-human transmission of avian influenza, the model is hypothetical. It is (obviously) important to choose reasonable values for all parameters.
- Researchers estimate the following values for the Spanish Flu pandemic in 1918–1919. We will use these values in our catastrophe model.

$$\begin{array}{lll} a = \text{infection rate} & b = \text{recovery rate} & c = \text{mortality rate} \\ = 1.9\ \text{week}^{-1}; & = 1.4\ \text{week}^{-1}; & = 0.065\ \text{week}^{-1}. \end{array}$$

Photo 12.6: Left: Spanish bird, nativity facade, Sagrada Famìlia, Barcelona, Spain. Right: Plague monument, Plzen, Czech Republic. (Source: PA.)

- Now we can use Euler's method and the SIRD model to investigate various scenarios in a city such as Brisbane with $N = 2 \times 10^6$.

Example 12.2.5

There is one infective person in a city in which $N = 2 \times 10^6$ and everyone else is susceptible. Figure 12.5 shows the results. The model predicts that the disease outbreak will last for about 45 weeks, around 870000 people will become ill, the largest number of infective people at any time will be about 59500, and that approximately 38500 people will die.

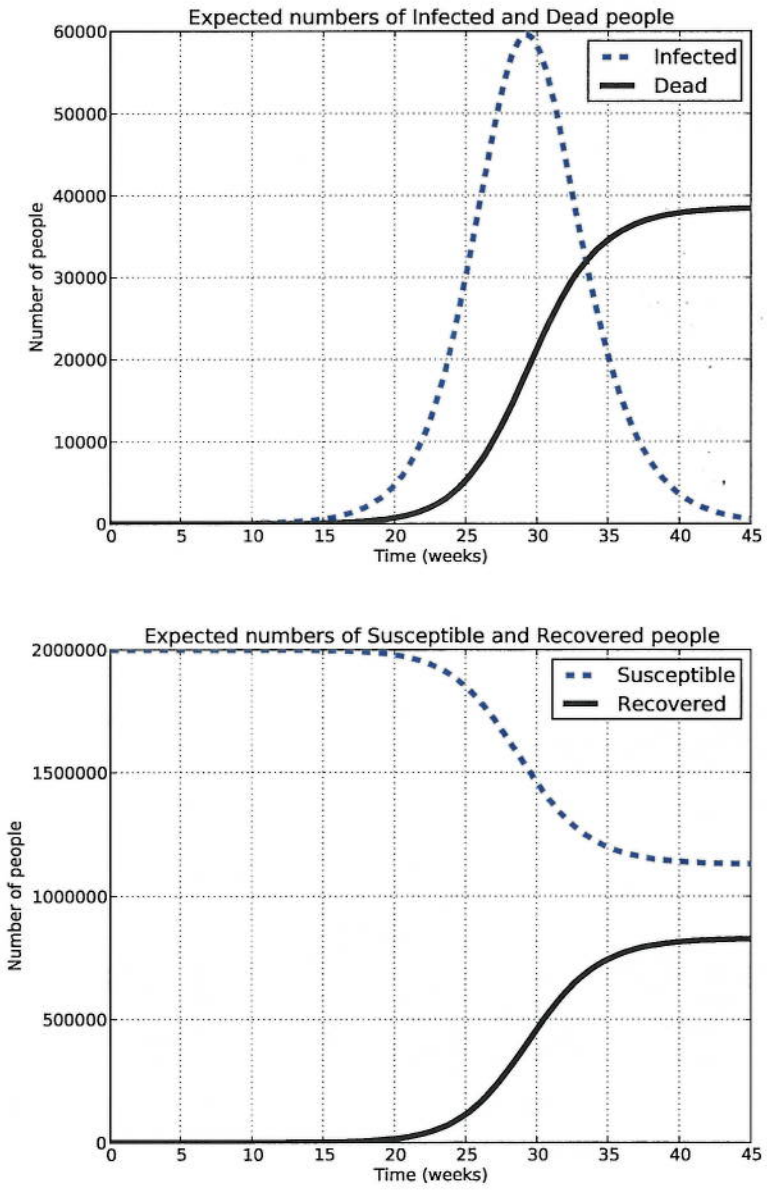


Figure 12.5: The impact of a possible human-transmissible avian influenza epidemic on a city of two million people, as modelled by Euler's method.

- Is this speculative catastrophe model realistic? For comparison, Figure 12.6 shows some mortality rates during the Spanish flu pandemic. The historical graphs have similar shapes to those in our catastrophe model.

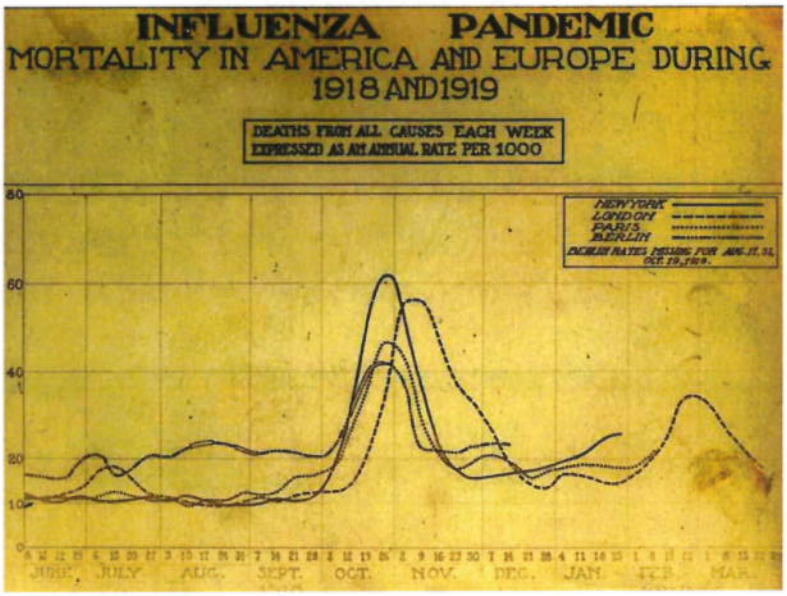


Figure 12.6: Spanish flu mortality rates in several cities, from 1918–1919. (Source: en.wikipedia.org)

- Our catastrophe model predicts an overall infection rate of 44% and a mortality rate of 4.4% of infective people. For the Spanish flu, infection rates reached around 50%, with mortality rates ranging from 2% to 5%.
- This suggests that our catastrophe model is (at least) plausible.
- Perhaps we should be very, very afraid? Or we should stop kissing poultry!

Photo 12.7: Peter really loves his poultry!

End of Case Study 37: Avian influenza.

Appendix A: **Python laboratory manual**

The coiling is fast
This time it's your last
Your soul asphyxiated
Final chance for escape terminated.
Enveloped in python
constriction complete
where dreams become nightmares
of total defeat.

Artist: *Torniquet.* Song: *Enveloped in python.*

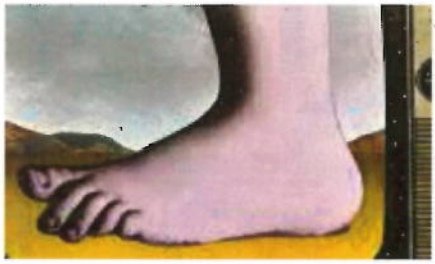


Image A.1: *Foot of Cupid*, Monty Python, BBC trademark. (Source: en.wikipedia.org)

Image A.2: *Venus, Cupid, Folly and Time*, (1540 – 45), Agnolo di Cosimo (usually known as Il Bronzino) (1503 – 1572), National Gallery, London. (Source: en.wikipedia.org)

A.1 Getting started

- Every computer program and computer model must be implemented in some computer *language*, which is a collection of commands that instruct the computer to perform associated operations and calculations.
- There are many different computer languages, each suited to various uses. In SCIE1000 we will use the language *Python*.[1]
- Commands vary between computer languages, however the following types of command are typical. (A brief Python example follows each command.)
  - **Comments** are ignored by the computer, but make programs easier to understand for people reading them.

    In Python programs, lines commencing with `#` are comments.
  - **Input commands** allow data to be entered into the program from the keyboard or a file.

    In Python, the command `input` reads data from the keyboard.
  - **Output commands** display data on the screen.

    In Python, the command `print` displays text and `plot` draws graphs.
  - **Variables** allow data values to be stored and manipulated.

    Python allows variables that store single data values, and variables called *arrays* that store multiple data values.
  - **Calculations** permit the computer to perform a range of mathematical calculations. Python supports all standard calculations.
  - **Booleans** evaluate expressions as being True or False.

    Python uses the values `True` and `False`, and words such as `and`, `or`.
  - **Conditional execution** allows the computer to execute certain commands if, and only if, a boolean expression is true.

    Python uses the conditional command `if-then-else`.
  - **Loops** execute commands multiple times, while a boolean expression is true. Python supports a number of loops, including `while` loops.

[1] Python was named after *Monty Python's Flying Circus*. We use Python because it is modern, freely available, fairly easy to learn, used in real science applications, and illustrates many important general computing concepts. Python users include Youtube, Google, Yahoo!, CERN and NASA.

A.2 Basic use of Python

We will write and run Python programs throughout semester. We start with a brief description of what a program is, and how to write, save and run programs. Then we cover some useful Python commands to use in your programs.

The usual approach to Python programming is to write a sequence of Python commands, save them to a file, and then run them. Saving programs in files means that

- programs can be run multiple times, on different computers.
- programs can be debugged more easily. For example, once a section of a program is thoroughly tested, it does not need any further testing.
- teams of people can design and write different parts of the program.
- problems of much greater complexity can be solved.

Creating and saving programs

To create and save a program:

- Use the *editor* window, which is similar to a word processor such as Microsoft Word, but simpler.
- Type in the Python commands for your program. You can enter multiple lines, and go back to fix any errors you might make.
- Do not include any spaces or tabs at the start of your lines – the reason will become clear later.
- When finished, choose **File** → **Save As** in the editor window. You may wish to create a special folder in which you save all of your Python programs.

To open an existing program, select **File** → **Open** and choose the file.

Naming your programs

Python programs are conventionally named `something.py` where `something` is a name that explains the purpose of the program. The text `.py` is the standard file extension for Python programs.

Running a program

To run the program that is in the editor window, select **Run** → **Run** from the menu bar. When you run your program, all output (and any requested input) will take place in a *console* window, separate from the editor window.

A.2.1 Comments in Python programs

Any lines starting with # in a Python program are *comments*, and are ignored by Python. Always use appropriate comments to explain what your program is doing, and to help anyone reading your program understand what is happening. Always try to use comments where appropriate.

A.2.2 Importing modules

Python provides access to mathematical functions and other useful operations, organised into *modules*. We will use the module called `pylab`. **All of your programs should have the following first line:**

```
from pylab import *
```

A.2.3 Printing to the screen

The *print* command outputs text and the results of calculations to the screen.

Printing text

To print text to the computer screen, use the `print` command in the following ways:

`print()`	prints a blank line.
`print("message")`	prints the message.
`print(expression)`	prints the **value** of the mathematical expression.

When printing multiple items (which may include a mixture of expressions and messages), items must be separated by commas.

Anything within quotation marks is printed *as it is*, whereas anything not inside quotation marks is evaluated as an expression and the *answer* is printed.

The following program demonstrates use of the `print` command. Note that line numbers are added for ease of reference and are not part of the program.

(If you wish to see how it works, type this program into a Python edit window, save it and then run it. Take care to type everything correctly.)

Program A.1: Printing things

```
1 from pylab import *
2
3 # Print some messages.
4 print("This is a message")
5 print("This is first", "and this is second")
6
7 # Print the results of some calculations.
8 print(3+4)
9 print("3*4 =",3*4)
```

Note that Lines 3 and 7 of the program contain comments. Also, you can use blank lines (such as Lines 2 and 6) to make your program more readable. Here is the output from running the program:

```
1 This is a message
2 This is first and this is second
3 7
4 3*4 = 12
```

A.2.4 Numerical calculations

As suggested by Lines 8 and 9 of the previous program, Python can use standard mathematical operations. The following table shows how to do this. (In each case, the letters a and b represent numbers.)

Mathematics	Python	Mathematics	Python
$a+b$	`a+b`	$a-b$	`a-b`
$a \times b$	`a*b`	$a \div b$	`a/b`
a^b	`a**b`	(...)	`(...)`

Important note

You may have seen `a^b` used to represent a^b on your calculator. In Python `a^b` means something completely different so be careful to use `a**b`.

The following program demonstrates some mathematical operations.

Program A.2: Simple calculations

```
from pylab import *

# Simple calculations:
print(3+2, 3-2, 3*2, 3/2)
print("3 squared =",3**2)

# Python correctly applies order of operations:
print(2+3*4," and ", (2+3)*4)
```

Here is the output from running the program:

```
5 1 6 1.5
3 squared = 9
14  and  20
```

Use spaces within expressions (almost) anywhere, to make the program easier to read and understand. When using spaces, remember that your main goal is *communication*; use your judgement about what works. The following example demonstrates one approach.

Program A.3: Spacing inside Python programs

```
from pylab import *

# Adding one space between numbers and symbols is reasonable.
print(6 + 4)

# You normally do not use space between brackets and numbers.
print((2 + 3) * (6 - 4))

# Sometimes spaces are used to show order of operations.
print(2 + 3*4 + 5*6)
```

Note that in Python, the symbol `e` represents scientific notation. For example, 6.02×10^{23} is displayed as `6.02e+023`, and 3×10^{-4} as `3e-04`.

A.3 Variables, functions and errors

A.3.1 Variables

A program can "remember" values by *assigning values* to *variables*. The programmer can create and use as many variables as needed in the program.

Each variable has a name, which is used to access it. Always choose *meaningful* names for your variables, to make the program easier to understand. Examples of names are `x`, `height`, `NumFish` and `x7`. Do not use spaces or other "special" characters in variable names. Also, note that variable names are case sensitive, so `numfish` and `NumFish` are different.

To assign a value to a variable called `name`, use the command

```
name = expression
```

where `expression` is either a value (such as `3` or `-2.25`), or an expression (such as `2 + 4`). Python calculates the result from the expression on the right hand side of the equals sign, and assigns that value to the variable.

Once you have assigned a value to a variable, you can use that variable name in subsequent calculations. The following program gives some examples.

Program A.4: Variables

```
from pylab import *

width = 20
height = 45
print("For a rectangle of size", width,"by", height)
print("the area is ",width * height)
perimeter = 2 * width + 2 * height
print("and the perimeter is", perimeter)
```

Here is the output from running the program:

```
For a rectangle of size 20 by 45
the area is  900
and the perimeter is 130
```

It is not possible to access a variable that does not yet have a value.

A.3.2 Python functions

Python includes many mathematical functions that act on one or more values and produce some output. To use a Python function, type the name of the function followed by the value to use, enclosed in brackets. Here is a list of some mathematical functions in Python.

`sqrt(value)`	square root
`sin(value)`	sine of the value (the value is in radians)
`exp(value)`	e raised to the given power
`log(value)`	ln
`log10(value)`	$\log_{10}$

Python also includes the constant `pi`, which (approximately) equals π. The following example demonstrates how to use mathematical functions.

Program A.5: Functions

```
from pylab import *

val = 9
print("The square root of 9 equals", sqrt(val))
print("e^1 = ", exp(1))
print("log to base 10 of 1000 equals",log10(1000))

# Evaluate sin of 90 degrees. First, convert to radians.
angleDeg = 90
angleRad = angleDeg * pi/180
print(angleDeg,"degrees =", angleRad,"radians.")
print(sin(angleRad))
```

Here is the output from running the program.

```
The square root of 9 equals 3.0
e^1 =  2.71828182846
log to base 10 of 1000 equals 3.0
90 degrees = 1.57079632679 radians.
1.0
```

A.3.3 The input function

It is often useful or necessary to ask the user to enter some input from the keyboard. The `eval(input(...))` function prints a message on the screen, and assigns the entered value to the specified variable.

```
variable = eval(input("message"))
```

The following example program shows how to use the `input` function.

Program A.6: Input

```
from pylab import *

# Input two values, then multiply and divide them.
a = eval(input("Tell me a number: "))
b = eval(input("Tell me another number: "))

prod = a * b
quot = a / b

print(a, "*", b, "=", prod)
print(a, "/", b, "=", quot)
```

Here is the output from running the program.

```
Tell me a number: 12
Tell me another number: 4
12 * 4 = 48
12 / 4 = 3.0
```

A.3.4 Software errors and bugs

All computer programmers have errors or *bugs* in programs they write. A key skill in programming is minimising the number of errors, and then identifying and fixing any that occur. There are many different types of error, including incomplete problem description, design faults in the software, unanticipated 'special cases', coding errors and logic errors.

In real life, the consequences arising from programming errors can be very serious: for example, they have caused plane crashes, rocket explosions and failure of entire transport systems. In SCIE1000 we will not give you control of aeroplanes, rockets or even small transport systems. The impact of any programming errors will be minor. You may feel a bit frustrated and need to ask for help, but no lasting damage will occur. Rather, you should learn from the process of finding and fixing your errors.

Avoiding errors

When writing programs, make sure that you:

- Understand the question **before** you start programming;
- Think about the best and most logical way to solve the problem;
- Consider planning your program on paper first;
- Put comments in your program so you know what you are trying to do;
- Test your programs on a range of data;
- Check some output carefully to make sure it is correct; and
- Pay attention to any error messages!

Fear not!

Do not be afraid of error messages! Never let the fear of error messages stop you from playing around with Python and trying different commands. Getting an error message **does not** mean that you will fail the course. If it helps you to figure out what you did wrong, then you have learned something! Equally important, do not ignore error messages. They give you useful advice about what is going wrong.

The following Python program illustrates a number of errors.

Program A.7: Multiple errors

```
1 from pylab import *
2
3 # Input two values, then multiply and divide them.
4 a = eval(input("Tell me a number: "))
5 b = eval(input("Tell me another number: "))
6
7 prod = a * b
8 quot = a / bb
9
10 print(a, "*", b, "=", prod)
11 print(a, "/", b, "=", prod )
```

Here is the output from running the program:

```
1 Tell me a number: 12
2 Tell me another number: 4
3
4 Traceback (most recent call last):
5   File "inputerror.py", line 8, in <module>
6     quot = a / bb
7 NameError: name 'bb' is not defined
```

To help you to identify the error:

1. The **last line** of the error message indicates **what** error occurred. In this case, it says `NameError: name 'bb' is not defined`
2. The third last line (Line 5) indicates **where** the error was detected. In this case, it says: `File "inputerror.py", line 8, in <module>`. This gives the name of the file and the line number where the error occurred.

The information in the error message allows the error to be located and then identified. In Line 8 of the program in the example above, the programmer has accidentally typed '`bb`' instead of '`b`'. Because the variable `bb` does not already have a value, the program cannot divide `a` by `bb`.

If a program contains multiple errors, Python will display the message for the first one it encounters. After you find and fix the error, Python may give a different error message. This is usually a good sign, indicating that the first error is fixed and you can move on to the next one.

Here is some output from running the previous program with `bb` changed to `b`.

```
1 Tell me a number: 12
2 Tell me another number: 4
3 12 * 4 = 48
4 12 / 4 = 48
```

Look carefully at the output – although there was no error message, there is still an error. The output says that `12 / 4 = 48`, which is incorrect. Hence you should **always check your output**, as it may be incorrect even if there is no error message. In this case, Line 11 should print the variable `quot`.

Finally, another type of error can arise. If you enter "0" for the second input number, then an error message will arise, saying that you cannot divide by zero! This is a type of "special case" error, which only arises for certain values. To avoid such errors, you need to test your program on a range of "special cases".

For reference, three common error messages are:

Error	Explanation and possible causes
`SyntaxError`	The command is not understood by Python. Perhaps: • You used incorrect bracket types (e.g. `( )` instead of `[ ]`) • You have forgotten a bracket • Your indentation is incorrect (wrong number of spaces at the start of a line)
`NameError`	There is no variable with the given name. Perhaps: • You have mistyped the name of a variable. • You have forgotten to set a starting value for a variable.
`IndexError`	You have used an invalid index to an array or sequence.

A.4 Conditionals

A.4.1 Introduction to conditionals

Programs often require the computer to do different things depending on whether a certain condition is true or false. For example, you might want to print various messages depending on the answer to the condition "are you aged over 18 and hence legally allowed to vote in Australia?" Python supports this by means of the *conditional command*, demonstrated below.

Program A.8: Conditionals, 1

```
from pylab import *

age = eval(input("What is your age? "))
if age >= 18:
    print("You can vote.")
print("Finished!")
```

Here is the output from running the program twice.

```
What is your age? 24
You can vote.
Finished!

What is your age? 17
Finished!
```

Here are some things to note about the conditional command.

- In the first line, the word `if` and the colon `:` are essential.
- The text between `if` and `:` must be a condition that is true or false.
- After the first line, any lines that are indented by four spaces will run if, and only if, the condition is true.
- Lines of the program that occur after any indented lines are outside the scope of the conditional command and therefore will run whether the condition is true or false.

A.4.2 Conditions

Python supports a number of conditions that can be true or false. For example:

Operation	Mathematics	Python
Greater than	$a > b$	`a > b`
Less than	$a < b$	`a < b`
Greater than or equal	$a \geq b$	`a >= b`
Less than or equal	$a \leq b$	`a <= b`
Equal to	$a = b$	`a == b`
Not equal to	$a \neq b$	`a != b`

Notice that the operator for checking whether two things are equal in Python is == and not just a single = sign.
Multiple conditions can be combined using the Python commands **and**, **or**, **not**, matching standard English usage of the words.

A.4.3 The `else` statement

In the previous example, the program prints a message if it is legal to vote, but gives no output if voting is not legal. When programming, we often have **two** possible situations – run some commands if a condition is true, and run other commands if the condition is false. This is done in the following way.

Program A.9: Conditionals, 2

```
from pylab import *

# Is it legal to vote?
age = eval(input("What is your age? "))
if age >= 18:
    print("You can vote.")
else:
    print("You cannot vote.")
print("Finished!")
```

Here is the output from running the program twice.

```
What is your age? 24
You can vote.
Finished!

What is your age? 17
You cannot vote.
Finished!
```

If there are more than two conditions to check then the command `elif` is used; it means "else if". Once again, we can use this to extend our example.

Program A.10: Conditionals, 3

```
from pylab import *

# Is it legal to vote?
age = eval(input("What is your age? "))
if age > 18:
    print("You can vote.")
elif age == 18:
    print("You can vote for the first time.")
else:
    print("You cannot vote.")
print("Finished!")
```

Here is the output from running the program three times.

```
What is your age? 24
You can vote.
Finished!

What is your age? 17
You cannot vote.
Finished!

What is your age? 18
You can vote for the first time.
Finished!
```

A.5 Loops

A.5.1 Introduction to loops

Programs often require some commands to run multiple times. For example, to model the growth of a population over 50 years, rather than writing 50 identical sections of code it is more convenient to write a single section, and run it 50 times. The programming concept which allows lines of code to execute multiple times is called a *loop*; here is some Python code demonstrating a loop.

Program A.11: Loops, 1

```
from pylab import *

# Print squares and cubes of numbers from 1 to 4.
i = 1
while i<5:
    print(i, i*i, i*i*i)
    i = i + 1
print("Finished!")
```

Here is the output from running the program.

```
1 1 1
2 4 8
3 9 27
4 16 64
Finished!
```

Here are some things to note about loops.

- In the first line, the word `while` and the colon `:` are essential, and the text between `while` and `:` must be a condition that is true or false.
- After the first line, any lines that are indented by four spaces will run while the condition is true.
- Lines of the program that occur after any indented lines will run once the condition is false.

Make sure you understand what is happening in the example loop above. A *loop control* variable, called `i`, is initially set to a value of 1. Each time the loop runs, the value of `i` is increased by 1, ensuring that the loop runs the required number of times, and stops when the condition `i<5` is false.

A.5.2 Loops and conditionals

There are some similarities between loops and conditionals. The *body* of each is indented by four spaces, and will only run if the initial condition is true. In a conditional, the body only runs once; in a loop, the body is run while the condition remains true. In both cases, the first line *after* the indentation will run after the conditional/loop has finished.

Multiple loops and conditionals can be *nested* within each other; indent by an extra four spaces each time.

A.5.3 Loop forever...

A *while* loop continues to run commands in the loop body until the condition is false. You must take care to choose a condition that will stop the loop at some stage. Consider the following loop:

Program A.12: Infinite loop

```
1 from pylab import *
2
3 i = 1
4 while i < 5:
5     print("forever...")
```

Notice that nothing within the body of the loop changes the value of `i`, so the condition `i < 5` is always true, and the loop will never terminate. This is called an *infinite loop*.

Stopping infinite loops

If you run a Python program and it seems to be taking a long time, it **may** contain an infinite loop. If you suspect that a running program contains an infinite loop, you can terminate it by pressing `Ctrl+C`.

A.6 Writing functions

A.6.1 Why write new functions?

Earlier we saw that Python includes mathematical functions such as `sqrt` and `sin`. Creating your own, new functions can be very useful. Once written, you can reuse them in multiple places or in different programs, without rewriting the commands each time.

Working with functions involves two related but distinct activities:

- *creating* (or *defining*) the function. You *define* a function by giving it a name and specifying the Python commands that actually do what the function requires.
- *using* (or *calling*) the function. You *use* the function by typing its name as one of your Python commands.

A.6.2 Writing a new function

In an earlier example, before using trigonometric `sin` we needed to convert the angle from degrees to radians. Here is some Python code demonstrating this.

Program A.13: Degrees to radians, 1

```
1 angleDeg = 90
2 angleRad = angleDeg * pi/180
3 print(sin(angleRad))
```

Converting from degrees to radians is a common calculation, so it may be useful to create a new function called (say) `toRadians` to do the conversion. Once the new function is written, the Python code could become

Program A.14: Degrees to radians, 2

```
1 angleDeg = 90
2 angleRad = toRadians(angleDeg)
3 print(sin(angleRad))
```

The second line has changed – now the new function performs the conversion to radians. Of course, before using the new function, you need to write it. Here is a Python program showing the new function.

Program A.15: Converting to radians

```
1 def toRadians(deg):
2 # This function converts degrees to radians.
3     rad = deg * pi/180
4     return (rad)
```

Pay careful attention to this example, as it demonstrates a number of important aspects of writing new functions. The first line and the last line are particularly important. Take the time to understand what is happening.

First line:

- The *name* of this function is `toRadians`. You should always choose **meaningful names** for your functions, to help you remember what the functions do.
- The word `def`, the brackets `(...)` and the colon `:` are essential!
- The variable inside the brackets will contain the value (in degrees) to be converted to radians. The common terminology is that this value is *passed into the function*. You can choose any name for this variable.

Last line:

- The last line 'returns' the value calculated by the function; in this case, it is the angle converted to radians.

Remaining lines:

- Except for the first line (or any comments), **every line** in the function **must** be indented; that is how Python can tell where the function ends. In SCIE1000, we will always use **four spaces** for indentation.
- These lines must perform the calculations required by the function, and create the value to be returned.

Here is an example showing how to use the new function.

Program A.16: Using a new function

```
from pylab import *

def toRadians(deg):
# This function converts degrees to radians.
    rad = deg * pi/180
    return (rad)

angleDeg = eval(input("Enter the angle in degrees: "))
angleRad = toRadians(angleDeg)
print(angleDeg,"degrees equals",angleRad,"radians.")
print("and sin() of this equals", sin(angleRad))
```

Here is the output from running the program:

```
Enter the angle in degrees: 90
90 degrees equals 1.57079632679 radians.
and sin() of this equals 1.0
```

A.6.3 Some notes on functions

Note that:

- Variables that are given values **inside** a function are **not accessible** outside the function, even if you use the same name in both places.
- You can call other functions from within a function.
- You can define multiple functions inside the same file. Remember that indenting shows where the body of each function starts and ends.
- The lines inside a function are **only** used when you call the function.
- A Python function is not restricted to only performing mathematical calculations. It can do anything that a program does.

A.7 Graphs

A.7.1 Plotting graphs

Drawing graphs is important in computer modelling. To draw a graph in Python, you need two *arrays* of the same size, one containing the x-coordinates of the points to plot, and the other containing the y-coordinates. (We will study arrays in detail in the next section. For the moment, note that an array allows you to store multiple values in a single variable name.) The following example plots three types of graph. Make sure you remember to use the `show()` function at the end of your program; if you forget, the graph will not be displayed.

Program A.17: Plotting graphs, 1

```
# This program demonstrates multiple plotti ng styles.
from pylab import *

A = arange(0,10)
# Plot y=A**2, using a solid line.
plot(A, A**2)
# Plot y=10A, using discrete points marked by 'x'.
plot(A, 10*A, 'x')
# Plot y=3A+20, using a wide solid line.
plot(A, 3*A+20, linewidth=3)
show()
```

Here is the output from running the program:

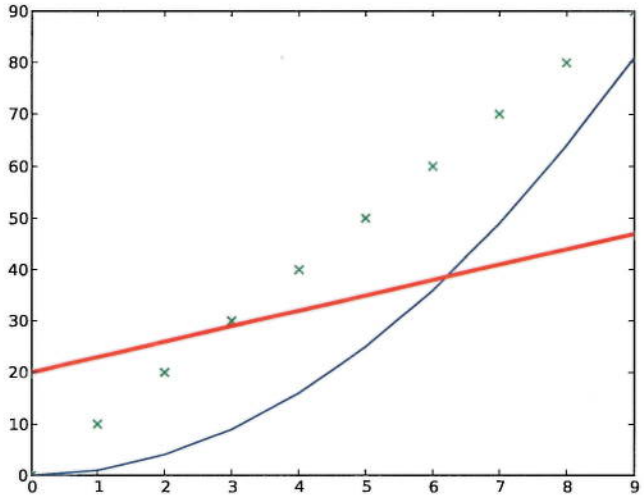


Figure A.1: Three graphs.

A.7.2 Graphing smooth functions

Computers cannot draw perfectly smooth curves: instead, they approximate smooth curves by drawing very short straight lines between points that are very close together. The more data points there are, the smoother the curve appears. The number of points needed to produce a smooth curve varies between different graphs. It is common to choose points whose x-coordinates are *equally spaced*, with a "small" spacing. The Python function `arange` creates an array of equally spaced points.

Equally spaced values in Python

The Python function

```
X = arange(a,b,s)
```

creates an array `X` of values starting at `a`, increasing by an equally spaced step of `s` each time, and stopping at the **last value less than** b.

Program A.18: Plotting graphs, 2

```
1  # This program shows how to use arange()
2  from pylab import *
3
4  # Plot sin(x) with x-coordinates separated by 0.5.
5  X1 = arange(0.0, 4.1, 0.5)
6  plot(X1, sin(X1), linewidth=3)
7
8  # Plot cos(x) with x-coordinates separated by 0.1.
9  X2 = arange(0.0, 4.1, 0.1)
10 plot(X2, cos(X2), linewidth=3)
11 show()
```

Note that in this program, the variable `X1` is an array which contains the values $0, 0.5, 1, \ldots, 4$. (Thus, `X1` contains nine values. We will see more about this in the next section on arrays.) Figure A.2 shows the output from running the program.

A.7.3 Customising your graphs.

Python provides a number of commands to customise your graphs. The commands in the following program should be self-explanatory.

Program A.19: Plotting graphs, 3

```
# This program shows how to customise graphs.
from pylab import *

# Create equally spaced points and plot sin and cos.
X = arange(0.0, 4.1, 0.1)
plot(X, sin(X), linewidth=3, label="sin(x)")
plot(X, cos(X), "--", linewidth=3, label="cos(x)")

# Create title and label axes.
title("Graphs of sin(x) and cos(x)")
xlabel("x")
ylabel("y")

# Draw a grid, create a legend and display the graph.
grid(True)
legend()
show()
```

Figure A.3 shows the output from running this program.

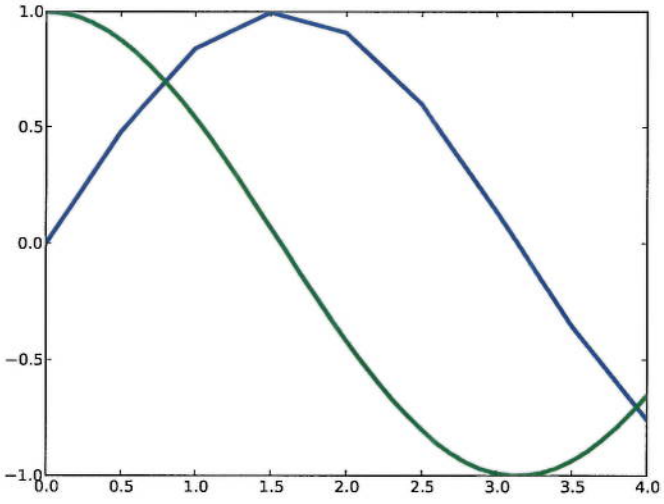


Figure A.2: Two graphs, with spacings of 0.5 and 0.1 between points.

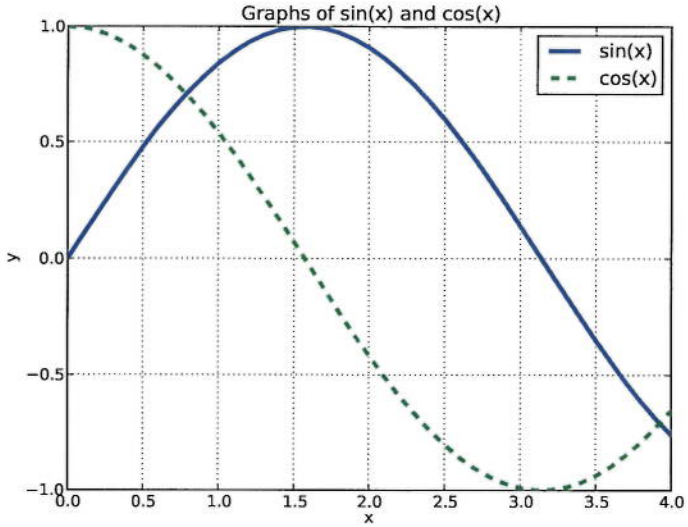


Figure A.3: Customising graphs.

A.8 Arrays

A.8.1 Introduction to arrays

Apart from when we were plotting graphs, we have only used Python to store individual data values in variables. A Python *array* is a new type of variable, which allows *multiple* items of data to be stored in the **same** variable. The following program creates and prints an array called `primes`.

Program A.20: Arrays, 1

```
1 from pylab import *
2
3 # Create an array containing the first 5 prime numbers.
4 primes = array([2, 3, 5, 7, 11])
5 print("Primes are:", primes)
```

Here is the output from running the program.

```
1 Primes are: [ 2  3  5  7 11]
```

Here are some things to note about arrays.

- Arrays are variables, so must have meaningful names.
- Python uses square brackets [and] to distinguish arrays from other variables. These are shown in Line 4 of the above program, and in the output.
- To create an array, use the `array(...)` function, with entries separated by commas; this is shown in Line 4 of the program. Alternately, you can use the `arange(...)` function we saw when plotting graphs.
- Line 5 uses the `print` command to display the entire contents of an array.
- The above array holds five values, so its `size` is 5.

A.8.2 Operations on arrays

Most Python commands we have already seen also act element-by-element on entire arrays at once, producing new arrays as the result. Pay particular attention to Lines 6 and 7 of the following example program.

Program A.21: Arrays, 2

```
from pylab import *

# Create arrays of primes and powers of 10
primes = array([2,3,5,7,11])
pows = array([0.01, 0.1, 1, 10, 100, 1000])
primeSq = primes * primes
pows = log10(pows)
print("Squares:   ", primeSq)
print("log(pows):", pows)
```

Here is the output from running the program.

```
Squares:    [  4   9  25  49 121]
log(pows): [-2. -1.  0.  1.  2.  3.]
```

A.8.3 Accessing individual array entries

In addition to dealing with an entire array, it is often useful to access individual entries in the array. The **index** of an entry refers to the *position* of that entry in the array (somewhat similar to the room numbers in the corridor of a building). To access individual entries, type the name of the array, immediately followed by the index surrounded by square brackets. For example, `A[i]` refers to the value at position `i` in the array `A`.

In Python, the **first** entry in an array has index `0`. This is important to remember! If `A` is an array of size n (so it contains n entries), then valid values of the index are from 0 to $n-1$ (inclusive). This is illustrated in Figure A.4.

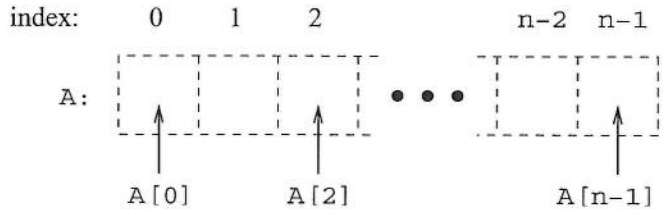


Figure A.4: The index of each entry in an array A with n entries

The following program demonstrates how to access individual array entries.

Program A.22: Arrays, 3

```
from pylab import *

# Create an array of primes and access various entries
primes = array([2,3,5,7,11])
print(primes)
print(primes[0], primes[4])
primes[0] = 13
primes[1] = 2 * primes[1] + 1
print(primes)
```

Here is the output from running the program.

```
[ 2  3  5  7 11]
2 11
[13  7  5  7 11]
```

A.8.4 Arrays and loops

Until now, we have used the `array` function to create arrays by listing their entries. The following program uses the function `zeros(n)` to create an array with `n` cells each equal to zero, then uses a loop to place values in the array.

Program A.23: Arrays, 4

```
from pylab import *

# Create an empty array then put values in it
X = zeros(5)
i=0
while i<5:
    X[i] = i*i
    i = i+1
print(X)
```

Here is the output from running the program.

```
[  0.   1.   4.   9.  16.]
```

A.9 Python summary

- All programs must commence with the line:

```
from pylab import *
```

- All lines commencing with '#' are comments.

```
# Use lots of comments to explain what your program does.
```

- The `print(...)` command displays text on the screen. Text inside quotation marks is displayed. Text outside quotation marks is treated as calculations or variables, and the values are printed.
- Values can be assigned to variables using '='.
- Use the operation `**` to find a power.

```
print("x cubed =", x**3)
```

- Useful inbuilt functions include `sqrt(value)`, `sin(value)`, `exp(value)`, `log(value)` and `log10(value)`.

```
print("The square root of x =", sqrt(x))
```

- The command `eval(input(message))` reads input from the keyboard.

```
numFish = eval(input("How many fish are there?"))
```

- The bodies of conditionals, loops and functions must be indented.
- The conditional commands `if`, `else` and `elif` control which commands are run, when a particular condition is true or false.

```
if x==y:
    print("x equals y")
elif x<y:
    print("x is less than y")
else:
    print("x is greater than y")
```

- The `while` loop allows commands to be run multiple times. Make sure that your loops will stop sometime!

```
i = 0
while i<5:
    print(i, "squared =",i**2)
    i = i+1
```

- To create a new function, first *define* it, then *call* it.

```
def toRadians(deg):
# Define a new function to convert degrees to radians.
    rad = deg * pi/180
    return (rad)

# Call the function:
print("Converted to radians, 90 degrees is", toRadians(90))
```

- Arrays allow multiple values to be stored in a "table", using a single variable name. The *index* of an element is its location in the array.

```
A = array([1,10,100,1000])
print("The last entry is",A[3])
B = log10(A)
C = zeros(10)
```

- Graphs are drawn using `plot()`, followed by `show()`. Commands such as `title()`, `xlabel()` and `ylabel()` allow you to customise your graphs.

```
x = arange(0,10.1,0.1)
y = x**2
plot(x, y)
show()
```

List of Figures

List of Images

List of Original Photographs

These notes contain a large number of original photographs. Qld Health images are copyright to the State of Queensland (Queensland Health) 2010 and are used with permission. In other cases, the photographer or photograph owner is identified, as follows:

AR: Anthony Richardson	DH: Dominic Hyde	DM: Delphia Manietta	WG: William Grey
MG: Melanie Giandzi	MGa: Marcus Gallagher	PA: Peter Adams	POD: Peter O'Donoghue

List of Programs

Bibliography

[1] Australian Institute of Health and Welfare (AIHW) and Australasian Association of Cancer Registries (AACR), *Cancer in Australia 2001.* AIHW cat. no. CAN 23. Canberra: AIHW (Cancer Series no. 28, 2004).

[2] Australian Medical Association, accessed 27 December 2010, <ama.com.au/node/2931>

[3] Benowitz *et al.*, *Nicotine intake and dose response when smoking reduced-nicotine content cigarettes*, Clinical Pharmacology and Therapeutics **80** (2006) 703–714.

[4] Caswell, *Matrix population models: Construction, analysis and interpretation*, 2nd Edition, 2001. Sinauer Associates, Sunderland, Massachusetts. ISBN 0-87893-096-5.

[5] Centers for Disease Control and Prevention, accessed 27 December 2010,
<www.cdc.gov/cancer/breast/pdf/cdc_mammogram_brochure.pdf>

[6] Cook *et al.*, *Pharmacokinetics of methamphetamine self-administered to human subjects by smoking S-(+)-methamphetamine hydrochloride*, Drug metabolism and disposition **21:4** (1993) 717–723.

[7] Crouse *et al.*, *A Stage-Based Population Model for Loggerhead Sea Turtles and Implications for Conservation*, Ecological Society of America **68** (1987) 1–13.

[8] Crowder *et al.*, *Predicting the impact of turtle excluder devices on loggerhead sea turtle populations*, Ecological Applications **4(3)** (1994) 437–445.

[9] Damon *et al.*, *Radiocarbon Dating of the Shroud of Turin*, Nature **337: 6208** (1989) 611–615.

[10] Dingle *et al.*, *What Is Migration?*, BioScience **57(2)** (2007) 113–121.

[11] Drugs.com, accessed 11 September 2010, <www.drugs.com>>

[12] Egevang *et al.*, *Tracking of Arctic terns* Sterna paradisaea *reveals longest animal migration*, PNAS **107** (2010) 2078–2081.

[13] Engber, *Vodka + movies = pseudoscience*, Chronicle of Higher Education, October 29, 2004.

[14] Engber, "Dodgy Boffins", accessed 17th August 2012,
<www.slate.com/articles/health_and_science/science/2012/05/science_in_the_telegraph_and_the_daily_mail_what_s_wrong_with_british_journalism_.single.html#pagebreak_anchor_2>

[15] Duijm *et al.*, *Sensitivity, specificity and predictive values of breast imaging in the detection of cancer*, British Journal of Cancer **76(3)** (1997) 377–381.

[16] Explorers Web 2004, accessed 6 September 2010,
<www.explorersweb.com/adventureweather/charts>

[17] Framingham Heart Study 2010, accessed 31 August 2010,
<www.framinghamheartstudy.org/participants/index.html>

[18] Framingham Heart Study 2010, accessed 31 August 2010,
<www.framinghamheartstudy.org/about/index.html>

[19] Framingham Heart Study 2010, accessed 31 August 2010,
<www.framinghamheartstudy.org/about/milestones.html>

[20] Galileo, *Dialogue Concerning the Two Chief World Systems*, trans. Stillman Drake, Berkeley: Univ. California Press (1953) 207.

[21] Gigerenzer *et al.*, *Helping Doctors and Patients Make Sense of Health Statistics*, Psych. Science in the Public Interest **8 (2)** (2007) 53–96.

[22] Goldacre, "Bad Science", accessed 19th September 2012, <www.badscience.net>

[23] Goldman *et al.*, *A re-examination of the age and growth of sand tiger sharks, Carcharias taurus, in the western North Atlantic: the importance of ageing protocols and use of multiple back-calculation techniques*, Environ Biol Fish **77:3-4** (2006) 241–252.

[24] Grau *et al.*, *Breast cancer in Rubens paintings*, Breast Cancer Research and Treatment **68** (2001), 89–93.

[25] Greenbaum *et al.*, *Impaired β-cell function, incretin effect, and glucagon suppression in patients with type 1 diabetes who have normal fasting glucose*, Diabetes **51** (2002) 943 – 950.

[26] Haub, "How many people have ever lived on Earth?", accessed 5th July 2013, <www.prb.org>

[27] Hoegh-Guldberg *et al.*, *Coral Reefs Under Rapid Climate Change and Ocean Acidification*, Science **318:5857** (2007) 1737 – 1742.

[28] International Organization for Standardization, ISO 2533:1975.

[29] Keeling *et al.*, *Exchanges of atmospheric CO_2 and $^{13}CO_2$ with the terrestrial biosphere and oceans from 1978 to 2000*, I. Global aspects, SIO Reference Series, No. 01-06, Scripps Institution of Oceanography, San Diego, (2001), 88 pages.

[30] Lundquist *et al.*, *Cyanide concentrations in blood after cigarette smoking, as determined by a sensitive fluorimetric method*, Clinical Chemistry, **33:7** (1987), 1228 – 1230.

[31] MacLulich, *Fluctuations in the numbers of the varying hare (Lepus americanus)*, University of Toronto Studies, Biological series **43** (1937).

[32] Maher, *Growth and development of the Yellow-bellied Sunbird* Nectarinia jugularis *in North Queensland*, Emu **91** (1991) 58-61.

[33] Moolenaar *et al.*, *Absorption rate and bioavailability of paracetamol from rectal aqueous suspensions*, Pharmacy World and Science, **1:1** (1979), 201–206.

[34] Stephenson *et al.*, *Mars Climate Orbiter Mishap Investigation Board, Phase I Report*,
<ftp.hq.nasa.gov/pub/pao/reports/1999/MCO_report.pdf>

[35] National Breast and Ovarian Cancer Centre, *Breast cancer risk factors: a review of the evidence*, National Breast and Ovarian Cancer Centre, Surry Hills, NSW, 2009.

[36] National Health Service, accessed 19 September 2012, <www.nhs.uk/News/Pages/NewsIndex.aspx>

[37] National Heart, Lung and Blood Institute 2008, accessed 31 August 2010,
<www.nhlbi.nih.gov/health/dci/Diseases/Hbc/HBC_WhatIs.html>

[38] Niyomnaitham *et al.*, *Bioequivalence study of 50 mg sertraline tablets in healthy Thai volunteers*, J Med Assoc Thailand **92:9** (2009), 1229–33.

[39] Pan American Health Organization, *EPI Newsletter Volume XX, Number 4*, August 1998.

[40] Popper, *The Logic of Scientific Discovery*, London: Hutchinson (1934).

[41] Ponce-Reyes *et al.*, *Vulnerability of cloud forest reserves in Mexico to climate change*, Nature Climate Change **2** (2012), 448–52.

[42] Posey and Mozayani, *The estimation of BAC, Widmark revisited*, Forensic Science, Medicine and Pathology, **3** (2007) 33–39.

[43] Rawlines *et al.*, *Pharmacokinetics of paracetamol (Acetaminophen) after intravenous and oral administration*, European Journal of Clinical Pharmacology **11** (1977), 283 – 286.

[44] Rodenhouse *et al.*, *Potential effects of climate change on birds of the northeast*, Mitigation and Adaptation Strategies for Global Change, **13** (2008) 487–516.

[45] Ronfeld *et al.*, *Pharmacokinetics of sertraline and its N-demethyl metabolite in elderly and young male and female volunteers*, Clinical Pharmacokinetics, **32** Suppl 1 (1997) 22–30.

[46] Salm *et al.*, *Blood flow in coronary artery bypass vein grafts: volume versus velocity at cardiovascular MR imaging*, Radiology, **232** 915–920.

[47] Schwartz *et al.*, *The role of numeracy in understanding the benefit of screening mammography*, Annals of Internal Medicine, **127** (1997) 966–972.

[48] Scripps Institution of Oceanography 2010, La Jolla, CA, accessed 24 September 2013,
<scrippsco2.ucsd.edu/data/atmospheric_co2.html>

[49] Seidl *et al.*, *The calculation of blood ethanol concentrations in males and females*, International Journal of Legal Medicine, **114** (2000) 71–77.

[50] Sheen and Hwang, *Mathematical modelling the cross-contamination of E. coli* O157:H7 *on the surface of ready-to-eat meal product while slicing*, Food Microbiology **27** (2010) 37–43.

[51] South Australian Government 2010, accessed 1 September 2010,
<www.dassa.sa.gov.au/site/page.cfm?u=122>

[52] Stenseth *et al.*, *Population regulation in snowshoe hare and Canadian lynx*, PNAS **94** (1997), 5147–5152.

[53] Wieland and Kasperski, *Estimating Net Present Value in the Northern Chesapeake Bay Oyster Fishery*, Prepared for: NOAA Chesapeake Bay Office, Non-native Oyster Research Program, Grant: NA05NMF4571231, (2008).

[54] United Nations, Department of Economic and Social Affairs, Population Division, *World Population Prospects: The 2010 Revision*, CD-ROM Edition (2011).

[55] Wilkinson *et al.*, *Pharmacokinetics of ethanol after oral administration in the fasting state*, Journal of Pharmacokinetics and Biopharmaceutics, **5:3** (1977) 207–224.

[56] Wolever *et al.*, *The glycemic index: methodology and clinical implications*, American Journal of Clinical Nutrition, **54** (1991) 846–854.

[57] World Health Organisation, *Global burden of disease*, accessed 10 October 2012, www.who.int/healthinfo/global_burden_disease/en

Index