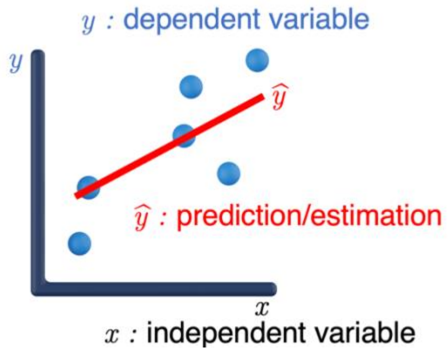
Review: Functions



Correlation and regression are methods that relate two sets of numerical data. Correlation gave us a sense of whether the variation in one quantity could be explained, or predicted, through the variation of the other; then we looked at regression as a way of finding a line of best fit. This line of best fit provides a model for the relationship between the two sets of data – a *function* that estimates the response $\hat{y}$ for a given predictor x. We assumed the function relating these variables was linear: that is, that it took the form $y = mx + c$. However, there are an infinite variety of possible relationships that could arise between the variables: in practice, there is a core set of common functions that we need to be familiar with, in a scientific context. Now we will turn our attention to what functions are, looking at a range of important functions that you should know, and some important skills that we need when working with functions.

Examples of functions

What are some situations where linear functions arise?

Any type of function that is not a linear function (not of the form $y = mx + c$) is called a *non-linear function.*

Important knowledge and skills when exploring relationships between data sets include

- familiarity with the more common nonlinear functions
  - what they look like when graphed, situations where they arise
- fitting nonlinear equations to data
- finding y for a given x, or vice versa
- finding important properties of the functions, such as maxima, minima, slopes, etc.

We will develop some of these skills in the coming weeks, learning about the relevant mathematical techniques and software.

What, exactly, *is* a function?

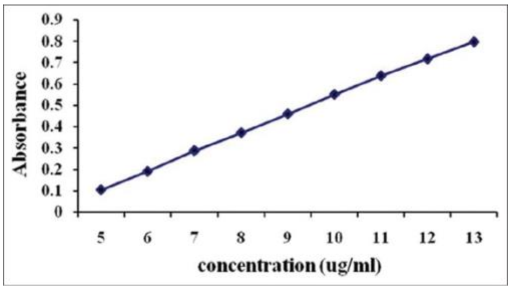
Functions express a relationship between two[1]* variables. They describe

- how a quantity evolves over time
- how a response depends on a predictor
- how an effect depends on a cause

[1] functions can exist between 3 or more variables, but we won't consider them in this course

Functions can be expressed in various ways

- As a graph



- In words
  - *Your final mark for this course is determine by …*
  - *To calculate the tax payable on your taxable income, …*
- As a formula

$$y = f(x) = sin(ax + b)$$

- As a procedure
  - *For natural number n, f(n) is the number of factors that divide into n.*

Function input and output

Function input (*abscissa*, *predictor*, *independent* variable):

- It's the **independent** variable, because we are free to choose its value.
- Generally called x if there is no preferred symbol
- Drawn on the **horizontal** axis
- The allowed input values are called the functions **domain** (where it lives)

Function output (*ordinate*, *response*, *dependent* variable):

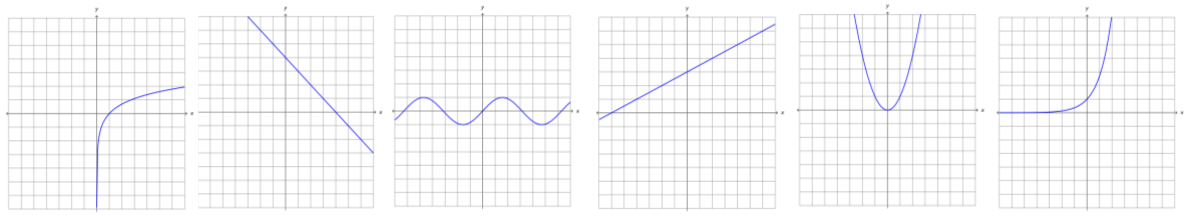
- It's the **dependent** variable: its value depends on the choice of x.
- Generally called y if there is no preferred symbol
- Drawn on the *vertical* axis
- The allowed output values are called the functions **range** (where it can reach)

If I have two sets of data, which do I choose which should be the independent variable x, and which should be the dependent variable y?

- Any variable you have control over: x
- Any timestamp variable[2]: x
- Any variable showing a response: y
- Any variable with repeated values: y

[2] Not a measurement of the time taken for something to occur, but a measurement of when (time and/or date) something has occurred, like the road toll data. We don't have control over time, but when a timestamp is one of the two data sets, we usually want to consider how the response is varying with time.

Drawing the graph of a function

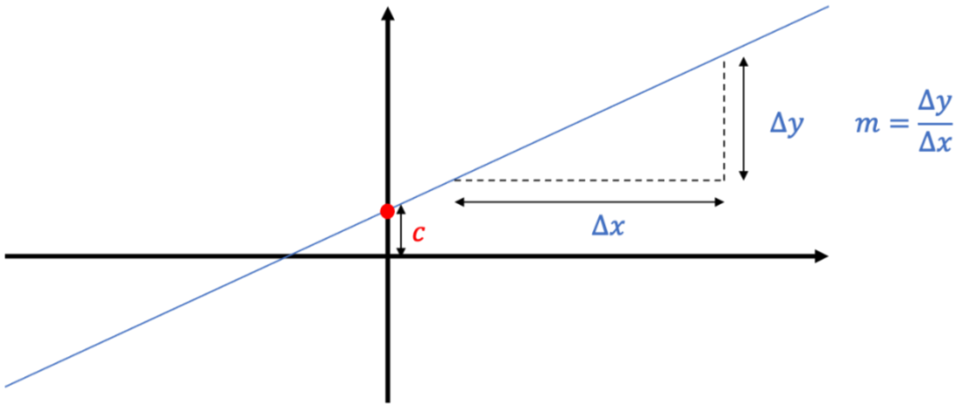


- The simplest approach is to remember the shapes of standard functions
- If you don't know/remember: use sample values of x to calculate f(x)

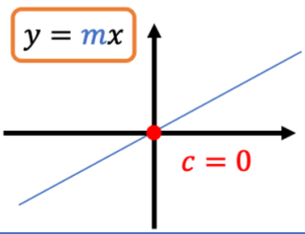
Linear functions

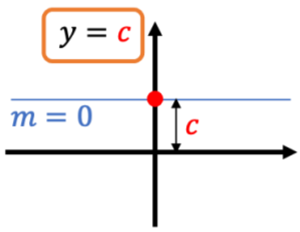
The straight line has the formula $y = mx + c$, where

- m is the slope
- c is the y-intercept



Special cases:

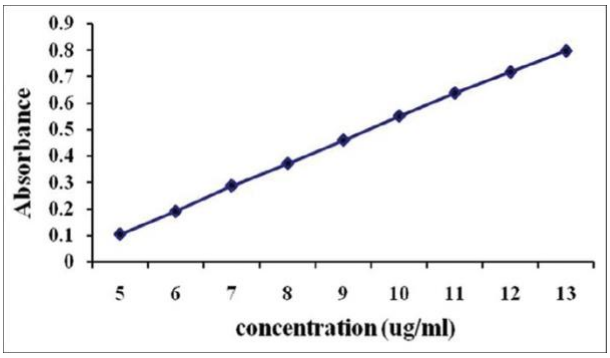




Example: using a concentration-absorbance regression curve of a drug administered to patients

- What concentration (W) is needed for an absorbance of 0.40?

$$A = 0.0875\,W - 0.3375$$



Non-linear functions

Any type of function that is not a linear function (i.e. not of the form $y = mx + c$ where the input variable x is raised to the power of 1) is called a *non-linear function*. We will be looking at a range of non-linear functions: polynomials, exponentials and logarithms, sinusoids, and a few other useful functions.

These types of functions had to be thought up to describe things we observe in real life – as important quantities in our world are rarely just simply a linear relationship!

Quadratics and other polynomials

A polynomial is a function of the form

$$y = a_0 + a_1x + a_2x^2 + a_3x^3 + a_4x^4 + \cdots$$

where the numbers $a_0, a_1, a_2, \ldots$ are all known fixed constants or *coefficients* of x to the various positive whole-number powers 0, 1, 2, 3, 4, … and so on.

The largest power of x that we see in any polynomial is called the *order* of the polynomial. Polynomials of first order are just linear functions, polynomials of second order are called quadratics, third-order polynomials are called cubics, fourth-order are quartics, etc.

What order are the following polynomials?

$y = 4x^2 - 2x + 1$	$y = 1 - x + x^2 - x^3$	$y = x^4 + 3x$
Order:	Order:	Order:

We can use also use tools like Excel to fit polynomials to data very easily, and because you have more coefficients to adjust as you increase the order, the fit usually gets closer. However, it is pretty uncommon for high-order polynomials to be the correct theoretical model for sets of data, which is why we usually stick to fitting low-order (usually linear or quadratic, sometimes cubic) polynomials unless there is a good reason to.

Solving polynomial equations is difficult for orders higher than the cubics – the best approach is to use software (such as WolframAlpha). It's suprisingly easy to do this, let's try solving the equations $x^3 - 2x + 1 = 0$ and $x^3 - 2x + 1 = e^x$ using WolframAlpha.

Go to WolframAlpha.com, type "Solve x^3-2x+1=0" or "Solve x^3-2x+1=e^x" into the input box that appears, and hit return to see the magic begin!

Write the output here of what solutions for x are for each case:

Quadratics

Quadratics is a name we give to second-order polynomials, so we can solve them using WolframAlpha as well as some algebra methods you may have seen in school or other study. However, they have some useful properties that are worth knowing, if you end up using them as a model for your data.

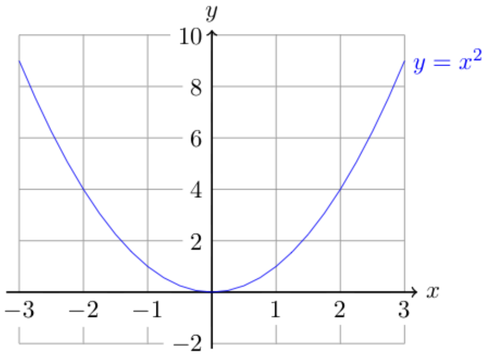
The general formula for a quadratic is $y = ax^2 + bx + c$, for coefficients a, b and c.

The coefficients can be used to understand a lot about the shape of a quadratic. Have a think about:

- The leading coefficient (coefficient of the highest power) a tells us …

- The constant c tell us …

- The graph of a quadratic is symmetric about its *extremum (the point where it takes its maximum or minimum value, depending on its shape)*. It is also symmetric about its *zeros (the points where y=0, which is the same as saying the points where it cuts the x-axis)*, if it cuts the *x*-axis.



- This axis of symmetry is located where $x = -b/2a$.

It is easy to get Excel to fit: we create a Chart of the data, and then add a Trendline via the Add Chart Element ribbon menu item. This trendline is a line of best fit, calculated to minimise the (least-squares) error in the same sort of approach we saw for linear regression. We can choose the form of the trendline from a range of functions, including polynomials, and Excel will show the equation and R^2 value.

Once we have the equation, we can use it to estimate the independent variable that produces the extremum (maximum or minimum value, depending on the shape of the quadratic).

Solving a quadratic

Apart from WolframAlpha, there are two standard algebraic approaches to solve quadratic equations[3] (this should be revision of some things from high school). One approach is to **factorise** the quadratic, the other is to use the **quadratic formula**

$$x = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}$$

where the quadratic we are solving is $ax^2 + bx + c = 0$. For a quadratic, we are looking for two solutions in general.

Factoring a quadratic expression means to express it as a product of two linear expressions multiplied together. As an example when *a*=1, we want to write a quadratic expression in the form:

$$x^2 + bx + c = (x + m)(x + n)$$

where *b*, and *c* are constants, *x* is the independent variable we want to solve, and *m*, *n*, are also number constants.

To factor a quadratic expression, we need to find two numbers that:

- multiply to give us the constant term *c*, and i.e. $m \times n = c$
- add (or subtract) to give us the coefficient of the *x*-term *b* i.e. $m + n = b$

Once we have these two numbers, we can use them to write the original qudaratic in the form $(x + m)(x + n)$

Here is an example to illustrate the process:

Factor and solve the quadratic expression: $x^2 + 5x + 6 = 0$

Step 1: Find two numbers that multiply to give us 6 and add up to give us 5.

The two numbers are 2 and 3, because 2×3 = 6 and 2+3 = 5.

Step 2: Use the two numbers to write the expression in factored form.

$$x^2 + 5x + 6 = (x + 2)(x + 3) = 0$$

Note that the two linear factors have the form (x+a) and (x+b), where a and b are the two numbers we found in step 1.

Step 3: Now we can see that either $x + 2 = 0$ or $x + 3 = 0$ for the above formula to work! That lets us now solve the original quadratic to say that $x = -2$ or $x = -3$

In general, factoring a quadratic can be more difficult, especially when the coefficients are not integers or when the constant term c has many factors. When things get harder we could use something like the quadratic formula to help us.

[3] This should be revision, but don't panic if you're rusty on this, because you'll always be able to use software to help you in this course (and most likely in real life, until you get to a point where it's quicker just to know how to do it reliably yourself!)

Solve $x^2 + 7x + 12 = 0$ using factorisation.

[Hint: what two numbers add to 7, and multiply to 12?]

Solve $x^2 + 7x + 12 = 0$ using the quadratic formula.

Solve $x^2 = 6 + x$ using either approach. *[Hint: first, re-arrange so it looks like a quadratic]*

A standard result to be aware of is called the *difference of two squares* - this is the result that

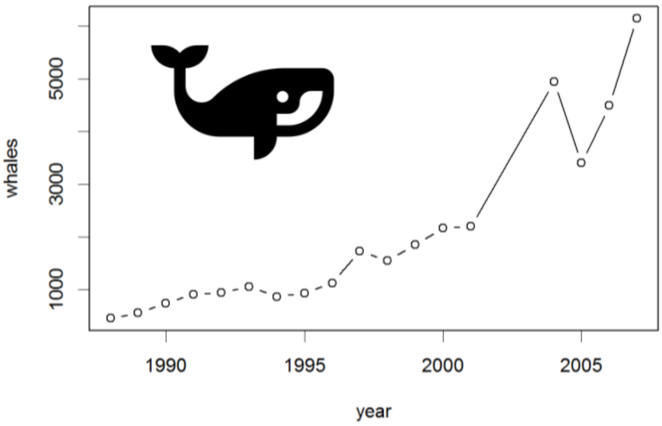
$$x^2 - a^2 = (x-a)(x+a)$$

which can give a quick and easy way to factorise the equation, and solution to an equation.

One final thing to note: From the quadratic formula, we can see that the axis of symmetry $x = -b/2a$ sits half-way between the two zeros

Introduction to exponential functions

In this plot of the evolution of the whale population in Hervey Bay from the late 1980s to the early 2000s, we can see that the population can be nicely approximately as linear until about 2000, but this this model breaks down beyond that point. How else might we model the data?



The straight-line model is a good model when the rate of increase – the number of whales the population grows by per year – is constant. This is because the straight-line model has constant slope $m = \Delta y/\Delta x$, where Δy is the change in whales over the period of time Δx. So m in the linear model represents the constant rate of increase of the whale population.

Here, we can do better than assume a constant rate of change. Populations don't usually change at a steady rate, in terms of change in the number of individuals per year. More usually, they change at a rate that is *proportional to their size*. That is because the chances of a mother giving birth might remain the same, but if the population is twice as big, there will be twice as many mothers, and twice as many infants born. Likewise, each individual faces the same risks to their life, so if the population is twice as big, twice as many individuals are likely to die.

Changing at a rate *proportional* to your size is equivalent to doubling (or halving) your size at a fixed rate. What type of function can be used to describe growth (or decay) of this kind? Imagine a population that doubles every year, starting at 1000. The population $P(x)$ after x years would be given by[4]

$P(0) = 1000 = P_0$

$P(1) = 2000 = 2 \times 1000 = P_0 \times 2$

$P(2) = 4000 = 4 \times 1000 = P_0 \times 2^2$

$P(3) = 8000 = 8 \times 1000 = P_0 \times 2^3$

$P(4) = 16000 = 16 \times 1000 = P_0 \times 2^4$

So from this pattern we see that $P(x) = P_0 \times 2^x$.

[4] we often represent the initial population using the same letter that we use to represent the population function, but with a '0' subscript to indicate 'initial value'. Here the function is $P(x)$, so we use symbol P_0

For the whale data, it looks more like the population doubles every 7 years. In that case,

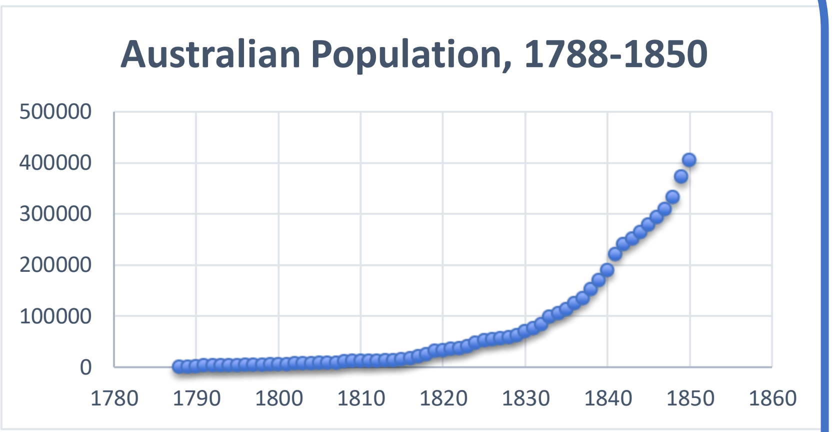
$P(0) = P_0$

$P(7) = P_0 \times 2 = P_0 \times 2^{7/7}$

$P(14) = P_0 \times 2^2 = P_0 \times 2^{14/7}$

So from this pattern we see that $P(x) = P_0 \times 2^{x/7} = P_0\left(2^{1/7}\right)^x$.

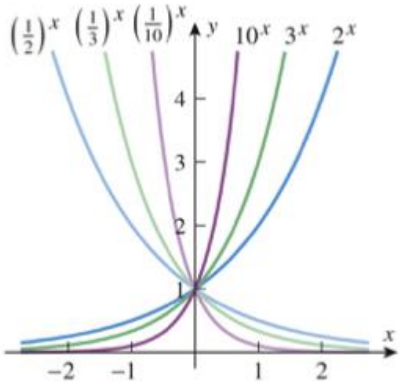
How might we describe the population of colonial Australia, over the period 1788-1850, using an exponential model?



The exponential function

Exponential functions are a family of functions of the form $y = b^x$, where b is called the base, and x is the exponent. We only use positive numbers for the base b.

Properties of the exponential functions



There are two bases for the exponential functino that we tend to prefer in maths:

- $b = 10$ is convenient because of our base-10 number system (and use of scientific notation)
- $b = e \approx 2.718\ldots$ is a common choice in mathematics. The number e plays a special role in the mathematics of exponentials, especially regarding their rates of change (as we'll see toward the end of the course). We can derive the number e by considering how interest compounds over shorter and shorter calculation times.

What is the return on investment if I invest at r % per annum for a year, if the interest is calculated daily?

How can I convert an exponential in base b to an exponential in base 10 or e? If I can write $b = 10^k$ or $b = e^r$, then this is straightforward:

But how can I find k, when $b = 10^k$ (or r, when $b = e^r$)? Taking the exponential allows me to find b when I know k, but here I have the opposite problem -- I need to 'undo' the process of finding the exponential, and find k when I know b.

In mathematics, we call this finding the *inverse* of the exponential (function). Finding the inverse of the exponential is so important that has its own function name – it is called the *logarithm*.

Logarithms

- If $y = b^x$, then $x = \log_b y$.
- If $y = 10^x$, then $x = \log_{10} y$.
- If $y = e^x$, then $x = \log_e y = \ln y$.

The logarithm with base e has its own name because of its fundamental importance in modern mathematics – it is called the natural logarithm, and has function name[5] ln.

It is commonly assumed (particularly in high-school maths curricula) that log by itself, without a given base, should be base 10, but that is by no means standard. It probably arises because many calculators and computer programs use log to mean $\log_{10}$, but you need to be check what base might be expected. In mathematics texts, log by itself could mean base 10, base e, or an arbitrary base.

Now we can work out how to convert to base 10 or base e from a different base:

Convert 2^t to base 10

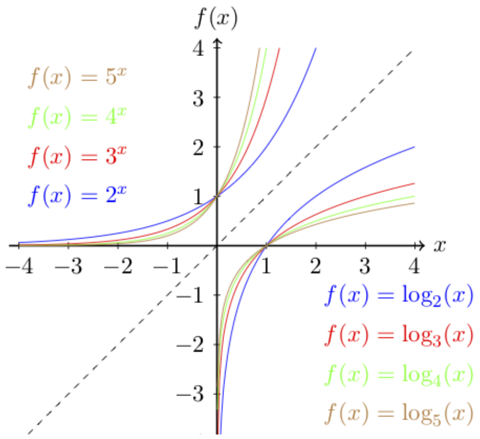
Convert $2^{t/7}$ to base e

[5] From the French *logarithme naturel*

This approach relies on the fact that calculators typically have buttons to calculate the logarithms in base 10 and base e, but not other bases. But what if I want to convert an exponential to a different base, such as base 2? To do this, I need to be able to calculate logarithms for arbitrary bases – how can I calculate $\log_b x$ for some number x, if my calculator can only find $\log_{10} x$ or $\ln x$? It turns out that[6]

$$\text{If } y = \log_b x, \text{ then } y = \frac{\ln x}{\ln b} = \frac{\log_{10} x}{\log_{10} b}$$

Properties of the logarithm functions



[6] To prove this: if $y = \log_b x$, then $b^y = x$. But, from the definition of the logarithm $b = e^{\ln b}$. Therefore

$$x = b^y = \left(e^{\ln b}\right)^y = e^{y \ln b}$$

Taking the natural logarithm of both sides, we get

$$\ln x = y \ln b, \qquad so\ y = \frac{\ln x}{\ln b}$$

We can repeat this, replacing e with 10 and ln with $\log_{10}$

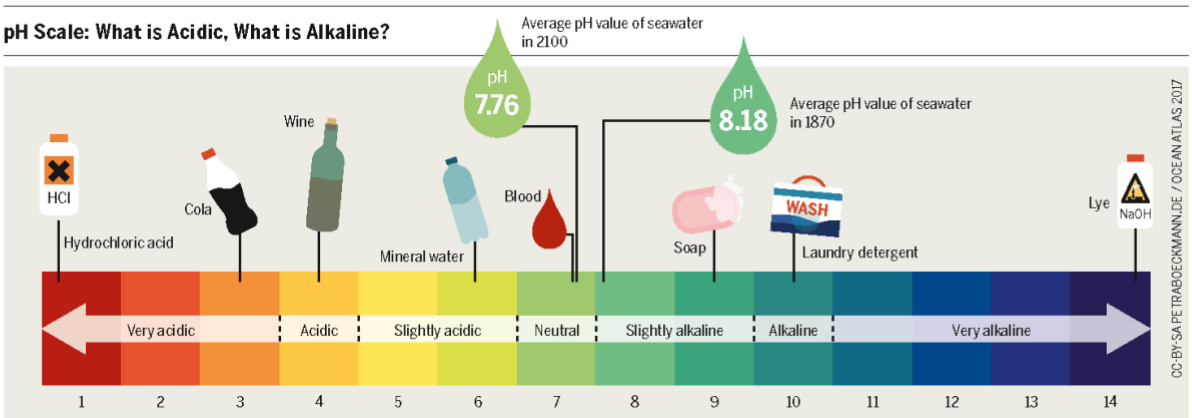
We can quickly estimate $\log_{10}$ from scientific notation. If the number is $r \times 10^m$ for $1 < r < 10$, it follows that

$$\log_{10}(r \times 10^m) = \log_{10} r + \log_{10} 10^m = m + \log_{10} r$$

and $0 < \log_{10} r < 1$, so we know that the $\log_{10}$ of the number sits between m and $m + 1$.

Estimate the log base 10 of Avogadro's number 6.022×10^{23}

Applications of exponentials and logarithms



The difference may seem small, but the decline in the pH value from 1870 to 2100 would mean a 170 percent increase in acidity. Much smaller changes already pose problems for many sea creatures.

It often isn't explained that the pH has a mathematical basis: the pH of a solution is the value of

$-\log_{10}[H^+]$, ie the negative log-base-10 of the molarity of hydrogen ions.

So a quite acidic solution, say pH=1, means that

$$-\log_{10}[\mathrm{H}^+] = 1, \text{ so } \log_{10}[\mathrm{H}^+] = -1, \text{ so } [\mathrm{H}^+] = 10^{-1} = 0.1 \text{ mol/litre}$$

What is the concentration of hydrogen ions in a swimming pool with pH=7.4?

In chemistry, there is also a quantity pOH, which is the negative log-base-10 of the molarity of hydroxyl (OH^-) ions, $-\log_{10}[OH^-]$.

It turns out that, at 25°C aqueous solutions satisfy $[H^+][OH^-] = 10^{-14}$, which means that

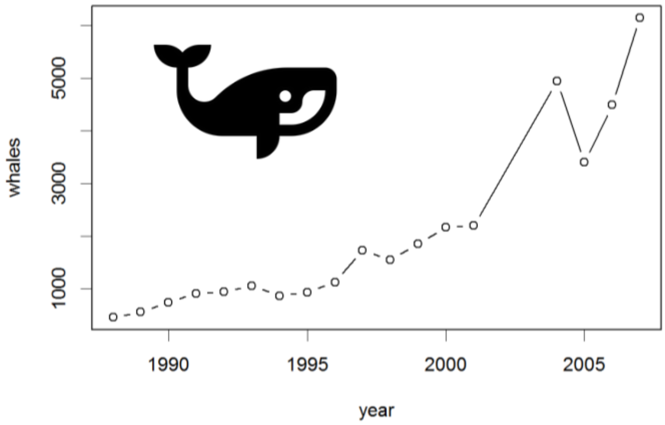
$$\begin{aligned} \text{pH} + \text{pOH} &= -\log_{10}[\text{H}^+] - \log_{10}(10^{-14}/[\text{H}^+]) \\ &= -\log_{10}[\text{H}^+] - \log_{10}[10^{-14}] + \log_{10}[\text{H}^+] = 14 \end{aligned}$$

What is the acidity of an aqueous 1mmol solution of NaOH?

If $P = 1000\, e^{-0.1t}$,

- is the population growing or shrinking?
- after how long does it double or halve?
- when would it reach a population of 100?

Express the population of humpback whales in Hervey Bay, using an exponential with base e



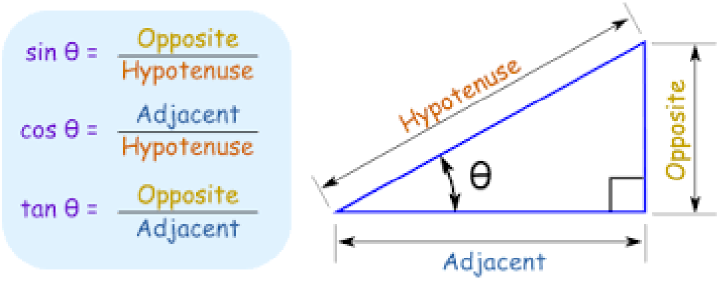
Periodic functions

In real life and science, we often deal with quantities that vary periodically. On a global scale, we have a daily 24-hour cycle due to the earth's rotation on its axis. But somethings undergo small daily changes, because they are influenced by other periodic processes – such as the phases of the moon, which have a period of around 29 days, or the time from sunrise to sunset, which varies over an annual time period due to the earth's revolution around the sun.

Describing quantities that vary periodically is a broad area of mathematics, and our modern approach begins with the functions sine and cosine.

We usually first encounter these functions when we explore the properties of right-angled triangles:

$sin(\theta)$ and $cos(\theta)$ are defined in terms of the ratios of sides of the right-angled triangle that has one angle equal to θ, as in the following picture.

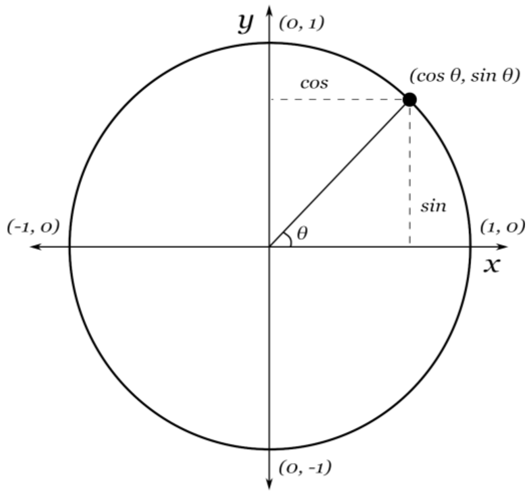


There are four more functions in this family, called the *sinusoids*. They are each defined as ratios of sides of a right-angled triangle – the remaining four sinusoidal functions can also be defined in terms of $sin(\theta)$ and $cos(\theta)$:

$$tan(\theta) = \frac{sin(\theta)}{cos(\theta)}, \quad sec(\theta) = \frac{1}{cos(\theta)}, \quad cosec(\theta) = \frac{1}{sin(\theta)}, \quad cotan(\theta) = \frac{1}{tan(\theta)}$$

We will only really use $sin(\theta)$ and $cos(\theta)$ here in this course.

The approach above allows us to define these functions for angles between 0° and 90°. We extend these definitions to arbitrary angles by drawing a circle of radius 1 at the origin (0,0) of a set of axes (this special circle is called the *unit circle*). We then overlap the above right-angled triangle onto that picture, putting the angle θ at the origin. The hypotenuse will end on the unit circle is we set its length to 1 – doing so also sets $sin(\theta) =$ Opposite, which is the y-coordinate of that point, and $cos(\theta) =$ Adjacent, which is the x-coordinate of that point.



If we rotate the point further around the circle, so that $\theta > 90°$, we can use the x- and y-coordinates of the point to extend our definition of $cos(\theta)$ and $sin(\theta)$, even though we can't have a right-angled triangle with one angle larger than 90°.

There are a few other relationships involving $sin(\theta)$ and $cos(\theta)$ that are good to be aware of.

Pythagoras' theorem tells us that $\text{Adjacent}^2 + \text{Opposite}^2 = \text{Hypothenuse}^2$, which gives us the relationship

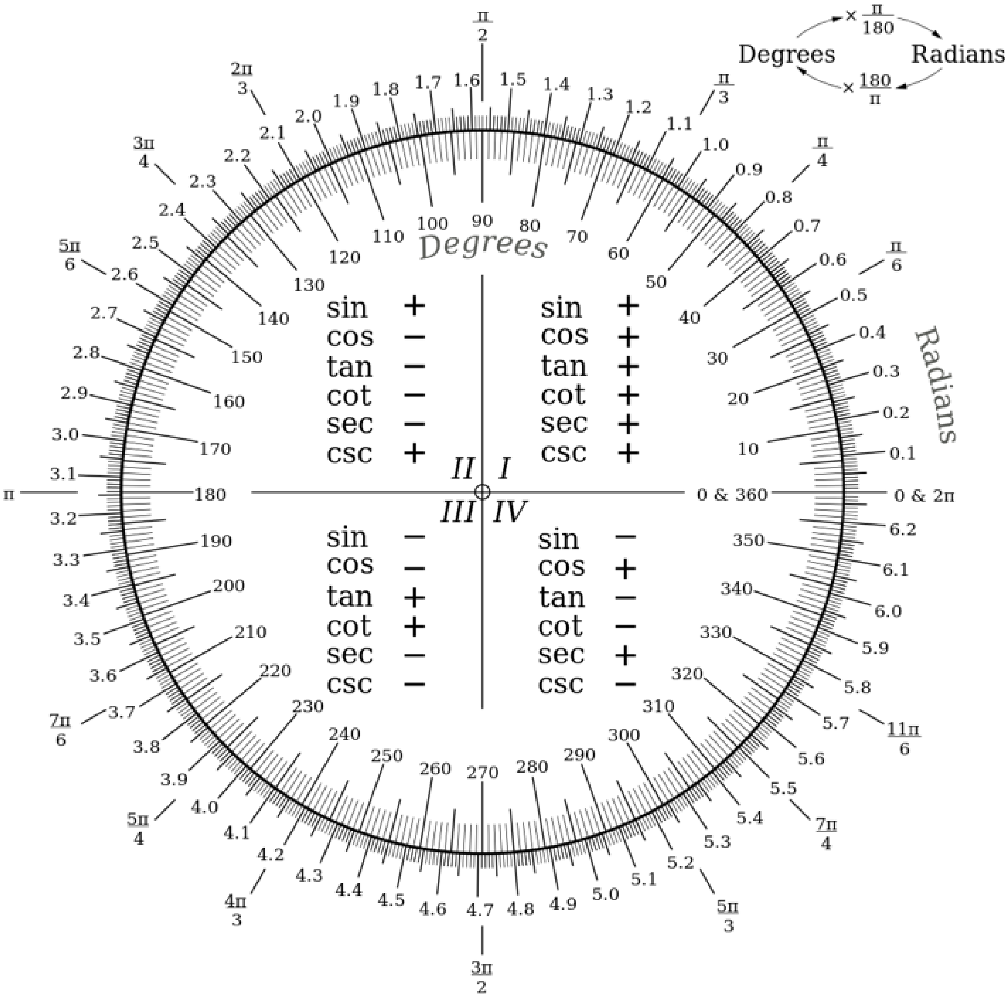
$$sin^2(\theta) + cos^2(\theta) = 1$$

Notice that the power of two is written between sin (or cos) and the angle – this is a slightly unusual notation, special to the sinusoidal functions. We do this so we don't get confused about whether we raising just the angle to a power – we are raising the whole thing to the power.

For arbitrary (non-right-angled) triangles, there is an extension of Pythagoras' theorem, known as the *law of cosines*, and there is also a *law of sines* that relates the size of angles of a triangle to the length of the opposite sites. We won't be using either of these laws in this course. These can be useful if you need to do more geometry problems in future!

Degrees vs Radians

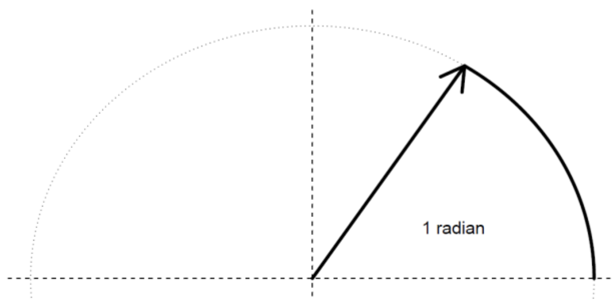
When you first learn about angles, you typically measure them in *degrees*. This is the common everyday measure – longitude and latitude are measured in degrees; builders and manufacturers will give angle specification in degrees. These are important everyday units to be able to understand and use. The degree is defined by assigning 360° to a complete rotation – this number is traced back to the ancient Babylonians, who understood that there were close to 360 days in a year, but also that 360 is a useful number because so many numbers divide into it: dividing a circle into 2, 3, 4, 5, 6, 8, 9, 10, 12, 15, 18 or 20 pieces gives whole-numbered angles for each piece (180°, 120°, 90°, 72°, 60°, 45°, 40°, 36°, 30°, 24°, 20° and 18°, respectively).



However, there is a second set of units for angles that mathematicians also use. While 360° is a useful number of degrees in a full circle, it is entirely arbitrary!

Instead, we may use the more fundamental approach. Draw the sector of a circle with angle θ, and define the size of the angle θ, in **radians**, as the length of arc traced out by the sector, divided by the circle's radius. This ratio is independent of the size of the circle, or the units used to measure lengths. Since a circle of radius r has circumference $2\pi r$, the angle of a full circle must be:

$$\theta = 2\pi r/r = 2\pi.$$



This gives us the conversion between degrees and radians:

$$360\ degrees = 2\pi\ radians$$

We can use this to convert between units of angles, using our usual approach of introducing a conversion factor, and canceling out the units we want to get away from:

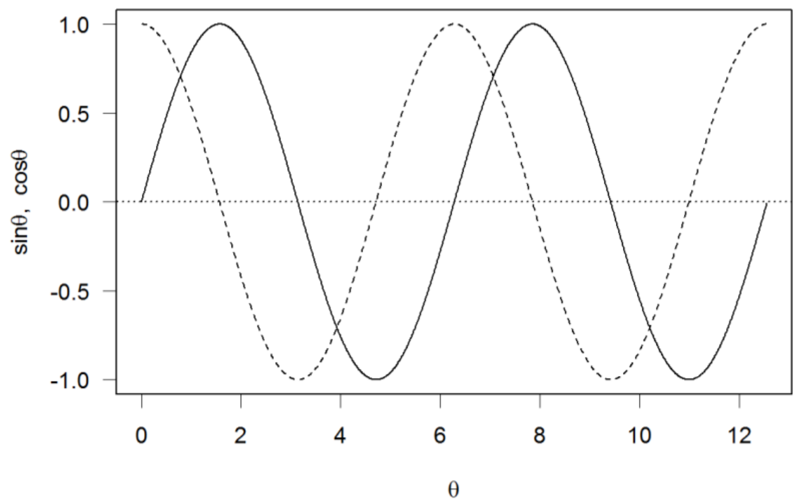
$$315° \cdot \frac{\pi}{180} = \frac{315\pi}{180} = \frac{7\pi}{4}$$

$$120° \cdot \frac{\pi}{180} = \frac{120\pi}{180} = \frac{2\pi}{3}$$

What is 90 degrees, in radians?

There are important mathematical reasons why we use radians, particularly related to the rates of change of $sin(\theta)$ and $cos(\theta)$ as the angle θ changes. For this reason, it is important to be aware of radians, and how to change between radians and degrees.

Plotting sin and cos



The graph here shows the plots of $sin(\theta)$ and $cos(\theta)$, as a function of the angle θ. The ***solid line*** is $sin(\theta)$ and the ***dashed line*** is $cos(\theta)$.

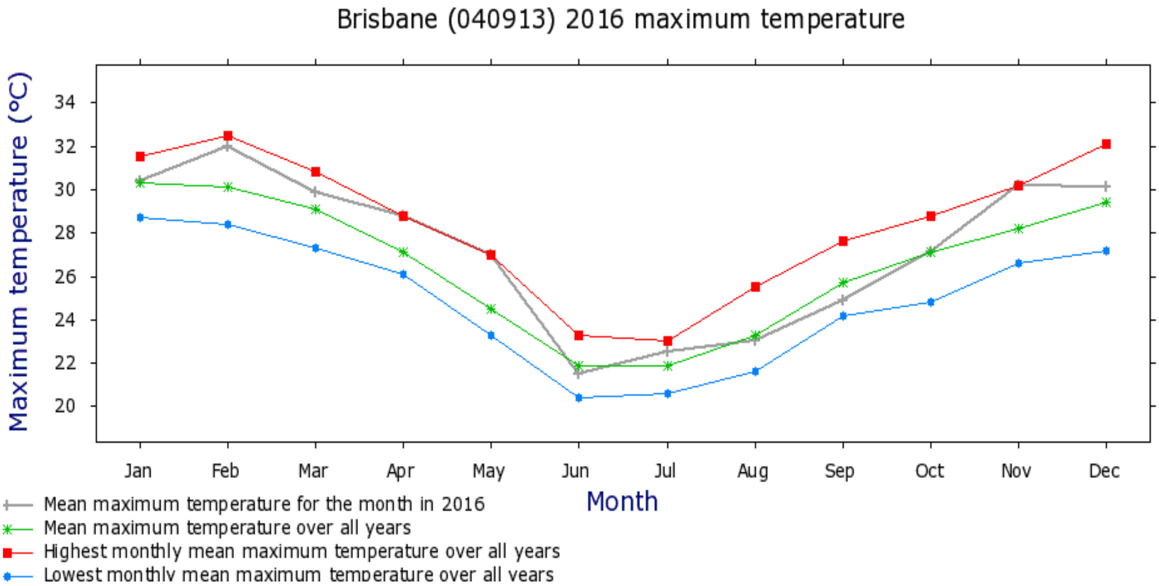
Notice that these functions have the same shape, but they are shifted relative to each other along the horizontal axis.

This is known as a *phase shift*, because the only change corresponds to adding a constant amount to the angle (also called the *phase*).

Also notice that these functions are periodic: they repeat every $2\pi \approx 6.28$, so that $sin(\theta + 2\pi) = sin(\theta)$ and $cos(\theta + 2\pi) = cos(\theta)$.

This means if you know the value of $sin(\theta)$ for the value of θ you have, if you add or subtract another 360 degrees or 2pi radians to θ, you will end up with the exact same value of $sin(\theta + 2\pi)$. This is a powerful result, and we can use it to describe so many natural phenomena, and even things in the business world for example.

This is the sort of variation that we commonly see in seasonal variation, such as the maximum temperature in Brisbane from year to year



However, if we defined t as the day number in the year ($t = 1$ for Jan 1, $t = 365$ for Dec 31), the maximum temperature is not just $sin(t)$ or $cos(t)$. We need to *change the scale* in the x- and y-directions, and *shift the function* in the x- and y-directions, so that the functions roughly overlap.

How can we do this?

Shifting and scaling functions

For any function of the form $y = f(x)$, we have a set of rules that allow us to rescale the function in the x- and y-directions, and to shift the function in the x- and y-directions:

<u>Shifting functions</u>: For a function $y = f(x)$, and positive constant $c > 0$:

New Function	Effect on Original Function
$y = f(x + c)$	moves the graph $y = f(x)$ **to the left** a distance of c
$y = f(x - c)$	moves the graph $y = f(x)$ **to the right** a distance of c
$y = f(x) + c$	moves the graph $y = f(x)$ **up a distance of** c
$y = f(x) - c$	moves the graph $y = f(x)$ **down a distance of** c

<u>Scaling functions</u>: For a function $y = f(x)$, and positive constant $c > 0$:

New Function	Range of c	Effect on Original Function
$y = cf(x)$	$c > 1$	stretches the graph vertically by a factor of c
$y = cf(x)$	$0 < c < 1$	compresses the graph vertically
$y = f(cx)$	$c > 1$	compresses the graph horizontally
$y = f(cx)$	$0 < c < 1$	stretches the graph horizontally

Examples of shifting and scaling functions:

It is worth noting that the changes needed to change the x-direction often seem counter-intuitive[7], whereas the changes needed to change the y-direction often seem quite natural.

[7] the opposite of what you'd expect

Putting these all together gives us a general formula for matching a sinusoidal variation:

$$y = A\sin[b(t+c)] + D$$

We can give specific meaning to some of these quantities:

- A is the *amplitude* of the sine wave: half the separation between the maximum and minimum values
- D is the *mean* of the sine wave: the average value of data (evenly spaced in time)
- b is the *(angular) frequency* of the function, telling us how many oscillations there are per unit time. If the function we are modelling has period T, then

$$b = \frac{2\pi}{T}$$

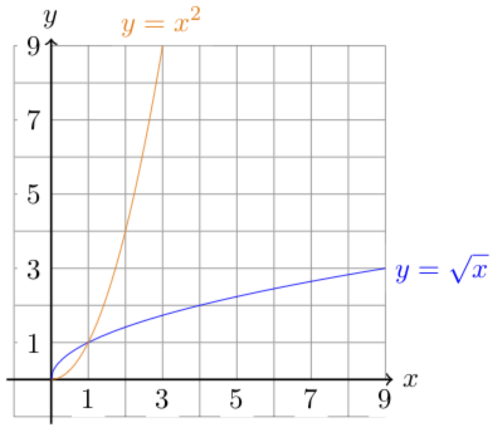
- c is the phase shift, measured in units of time.

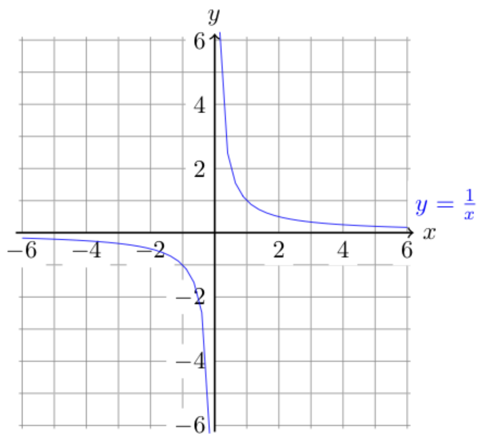
What values of A, b, c and D match the Brisbane temperature data?

Miscellaneous Functions (OPTIONAL)

It's impossible to give you all a comprehensive list of the functions you will need to know, because they depend so much on the applications of the mathematics you use. We have focused on three important classes of function: polynomials (and quadratics in particular), exponentials/logarithms, and sinusoids. To finish off, we look briefly at some other functions that you may encounter, so that you are aware of their existence and the context in which you might encounter them.

x to negative or fractional powers. The functions $y = x^{1/2} = \sqrt{x}$ and $y = x^{-1} = 1/x$ are plotted below. They look quite different to the graphs of polynomials.





The square-root function $y = \sqrt{x}$ looks like the quadratic $y = x^2$, except that the x and y axes have been switched. This is no accident. The square-root function is the *inverse* (the "opposite") of the square function, because it undoes the process of squaring a number: if you square a number, and then take the square-root, you end up where you started. The functions $x^{1/2}, x^{1/3}, x^{1/4}$, etc. are useful to model data that initially grows very quickly, but whose growth slows down as time increases (as an alternative to the logarithm function).

The function $y = 1/x$ has quite different behaviour to these examples. It blows up (diverges to infinity) at a finite value of x, so it allows us to model functions that have a similar divergent behaviour. By shifting along the x-axis, we can move the divergent x-value (the x-value where the function blows up). The power in the denominator controls how quickly the function diverges.

Hyperbolic sinuisoids. Despite their intimidating name, these functions are just combinations of the exponentials e^x and e^{-x}. We define

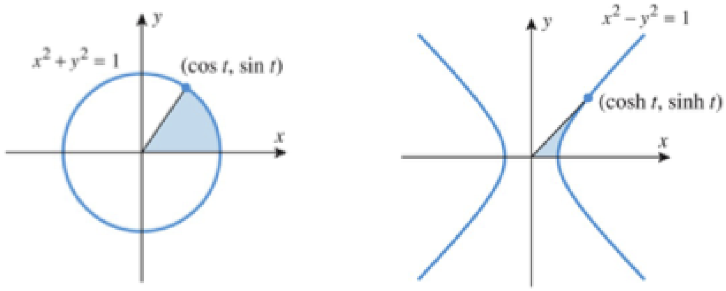
$$\cosh x = \frac{e^x + e^{-x}}{2}, \quad \sinh x = \frac{e^x - e^{-x}}{2}$$

And we define the other hyperbolic functions in analogy with their 'non-hyperbolic' counterparts:

$$\tanh x = \frac{\sinh x}{\cosh x}, \quad \operatorname{sech} x = \frac{1}{\cosh x}, \quad \operatorname{csch} x = \frac{1}{\sinh x}, \quad \coth x = \frac{1}{\tanh x}$$

The names of these functions are the same as the regular sinusoids, but with 'h' at the end. However, they are pronounced a little oddly. "Cosh" is pronounced as you'd expect, but "Sinh" is pronounced "shine", as if the 'h' were the second letter. Similarly, "Tanh" is pronounced "than" (with a long 'a' like in "tan", but starting with a "th").

They share some similarities with the sinusoids, but where we used the unit circle to define $sin(\theta)$ and $cos(\theta)$, the hyperbolic sinusoids can be defined graphically taking a similar approach on a hyperbola:



Since these functions are just fancy ways of combining exponentials, you might (quite reasonably) wonder why we both with them. Mainly, it is because there are a lot of similarities and mathematical connections between the properties of the sinusoids and the hyperbolic sinusoids, so in some situations it is more convenient to work with these hyperbolic sinusoidal functions. But in this course, we will mainly consider exponentials in the context of exponential growth or decay, in which case the hyperbolic sinusoids don't help us very much.

<u>The bell-shaped curve</u>. The Gaussian distribution plays a central role in statistics, so it is useful to be aware of its functional form. The Gaussian distribution with mean 0 and standard deviation 1 is given by:

$$f(x) = \sqrt{\frac{1}{2\pi}}\, e^{-x^2/2}$$

The square-root out the front guarantees that the area under the function is 1. We can use our shift and rescaling results to work out the distribution that has mean μ and standard deviation σ. This involves shifting the mean from $x = 0$ to $x = \mu$, and rescaling in the x-direction by a factor σ:

$$\sqrt{\frac{1}{2\pi}}\, e^{-x^2/2} \longrightarrow \sqrt{\frac{1}{2\pi}}\, e^{-\left(\frac{x-\mu}{\sigma}\right)^2/2} = \sqrt{\frac{1}{2\pi}}\, e^{-(x-\mu)^2/2\sigma^2}$$

However, in stretching the function, we increase its area, so we need to divide the whole thing by σ, to keep the area under the graph equal to 1.

So the general formula for a Gaussian of mean μ and standard deviation σ is

$$f(x) = \sqrt{\frac{1}{2\pi}\frac{1}{\sigma}}\, e^{-(x-\mu)^2/2\sigma^2} = \sqrt{\frac{1}{2\pi\sigma^2}}\, e^{-(x-\mu)^2/2\sigma^2}$$