# math612.1 612.1 Numbers, arithmetic and algebra

Gunnar Stefansson (editor) with contributions from very many students

3. september 2018

**Copyright** This work is licensed under the Creative Commons Attribution-ShareAlike License. To view a copy of this license, visit http://creativecommons.org/licenses/by-sa/1.0/ or send a letter to Creative Commons, 559 Nathan Abbott Way, Stanford, California 94305, USA.

**Acknowledgements**
MareFrame is a EC-funded RTD project which seeks to remove the barriers preventing more widespread use of the ecosystem-based approach to fisheries management.
`http://mareframe-fp7.org`
This project has received funding from the European Union's Seventh Framework Programme for research, technological development and demonstration under grant agreement no.613571.
`http://mareframe-fp7.org`
The University of Iceland uses the tutor-web in many courses and funds content-development as a part of this use.
The University of Iceland Research Fund has funded many of the studies developing algorithms uses in tutor-web.
`http://www.hi.is/`

# Efnisyfirlit

# 1 Numbers, arithmetic and basic algebra

## 1.1 Natural Numbers

The positive integers are called natural numbers.

These numbers can be added, multiplied together and so forth.

Notation: $\mathbb{N} = \{1, 2, 3, 4, ....\}$

Subtraction and division are not defined on these numbers.

An arbitrary element of $\mathbb{N}$ is most commonly denoted by $i,\ j,\ n$,or $m$, but any symbol can be used.

### 1.1.1 Details

**Definition 1.1.** The set of positive integers is usually denoted by $\mathbb{N}$, i.e. $\mathbb{N} = \{1, 2, 3, 4, ....\}$ and is called the set of **natural numbers**. In some cases the number zero is included as a natural number, but here we will use the symbol $\mathbb{N}_0$ to denote the integers 0, 1, 2 and up.

Within this set of numbers it is possible to add and multiply numbers together. Arithmetic operations are denoted by $+$ for addition and $\cdot$ (or $\times$) for multiplication. A natural number can also be raised to the power of a natural number, e.g. $3^5 = 3 \cdot 3 \cdot 3 \cdot 3 \cdot 3$ or in general $m^n = m \cdot m \cdot \ldots \cdot m$ ($n$ times).

When stating general properties of the natural numbers one needs to use symbols to indicate that the property holds for an arbitrary number. It is not enough to just write the property for a few numbers. For example, to declare that one can interchange numbers in a sum, it is not enough to say $4 + 3 = 3 + 4$ but one must explicitly state "the addition operator has the property that any two natural numbers, $n,\ m \in \mathbb{N}$ satisfy $n + m = m + n$".

An arbitrary element of $\mathbb{N}$ is most commonly denoted by $i,\ j,\ n$,or $m$, but any symbol, $a,\ b,\ c, \ldots$, can be used.

Several rules of arithmetic apply (some by definition, others can be derived) such as

$$\begin{aligned}
ab &= ba \\
a + b &= b + a \\
a + bc &= a + (bc) \\
a(b + c) &= ab + ac \\
(a + b) + c &= a + (b + c) \\
(ab)c &= a(bc)
\end{aligned}$$

Subtraction and division are not generally defined. In addition, we define one integer, $n$, to the power of another, $m$, to mean $n$ multiplied by itself $m$ times: $n^m = \underbrace{n \cdot n \cdot \ldots \cdot n}m$.

**Definition 1.2.** The power is an **operator** just like addition and multiplication, and is defined to have higher priority than the other two.

### 1.1.2 Examples

**Example 1.1.** If we have $x = 4$ and $y = 2$ and want to evaluate

$$x^y + y^x$$

then we replace the values of x and y in the expression, and evaluate it, taking care to observe the correct order of operations:

$$4^2 + 2^4 = 16 + 16 = 32.$$

## 1.2 Starting with R

Download R from the R website: http://www.r-project.org/

Look at on-line information on R, and take the tutor-web R tutorial: http://tutor-web.net/stats/stats240.1

Simple R commands:

- Assignment: $\texttt{x} < -2$
- Arithmetic: $2 * 5 + 4$

### 1.2.1 Details

To assign values to a variable in R one can use « -ϙr -"; however, these are **NOT** equivalent. Using the equals sign is confusing and therefore not recommended.

### 1.2.2 Examples

**Example 1.2.** Assigning values to a variable:

```
x<-2
y<-3
z<-x+y
```

**Example 1.3.** Viewing assigned values:
Type the name,i.e. "z", to view the assigned value.

```
z
[1] 5
```

## 1.3 The Integers

The set of positive and negative integers:
$\mathbb{Z} = \{..,..,-2,-1,0,1,2,......\}$

### 1.3.1 Details

**Definition 1.3.** The set of all integers is denoted by $\mathbb{Z}$, i.e.

$$\mathbb{Z} = \{..,..,-2,-1,0,1,2,......\}.$$

*Note 1.1.* Note that within this set it is possible to subtract as well as add and multiply. Within this set we cannot, however, in general, perform division.

When preforming multiple mathematical operations within the same equation, i.e. $79-8\cdot 3$, there is a conventional order for which the operations must be performed.

**Definition 1.4.** The conventional order of operations for equations with multiple mathematical operations is referred to as an **operator precedence**.

### 1.3.2 Examples

**Example 1.4.** To compute $79-8\cdot 3$ start by multiplying and then subtracting:
$79-8\cdot 3 = 79-24 = 55$

**Example 1.5.** To compute $15-(24+36)$ we first note that the parentheses (brackets) imply a precedence; anything inside brackets should be evaluated first.
Thus, we first add 36 to 24 and then we subtract that from 15.
15 - (24+36) = 15 - 60 = - 45

Note that the answer is a negative number.

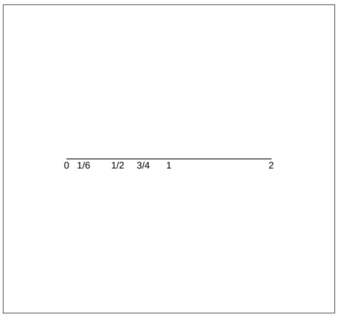
**Example 1.6.** Simple arithmetic in R is easily done at the command prompt.

```
79-8*3
[1] 55
15-(24+36)
[1] -45
```

## 1.4 Rational numbers

**Rational numbers** are fractions denoted p/q, where p and q are integers. We can simplify fractions if the numerator and denominator contain common terms.

### 1.4.1 Details




**Definition 1.5. Rational numbers** are fractions denoted $p/q$, where $p$ and $q$ are integers. The set of all rational numbers is usually denoted $\mathbb{Q}$.

*Note 1.2.* Note that every integer is a rational number (obtained by taking $q = 1$).

We can simplify fractions if the numerator and denominator contain common terms.

When the rationals are ordered on to a line there are points missing, i.e. there are "gaps", for example there is no rational number $p/q$ such that $(p/q)^2 = 2$.

### 1.4.2 Examples

**Example 1.7.** $\frac{2}{6}=\frac{2}{2\cdot3}=\frac{1}{3}$

The rational numbers can be put in order along a line as in the figure.

**Example 1.8.** As an elaborate example of a fraction, consider the evaluation of the quantity

$$\frac{\frac{2}{3}+\frac{2}{5}}{\frac{1}{3}+\frac{1}{2}}$$

**Example 1.9.** Evaluate

$$\frac{\frac{2}{3}+\frac{2}{5}}{\frac{1}{3}+\frac{1}{2}}$$

Solution: We can either start by calculating the numerator

$$\frac{2}{3}+\frac{2}{5}$$

or the denominator

$$\frac{1}{3}+\frac{1}{2}$$

.

Here we choose to start with the numerator. The first step is to make the two fractions in the numerator have a common denominator. We can either find the least common denominator or multiply the fractions with each others denominator. Here they are the same number, 15. So the first step is:

$$\frac{2}{3}\cdot 5+\frac{2}{5}\cdot 3=\frac{2\cdot 5}{3\cdot 5}+\frac{2\cdot 3}{5\cdot 3}=\frac{10}{15}+\frac{6}{15}$$

.

Now it is possible to add the two fractions which is the second step:

$$\frac{10+6}{15}=\frac{16}{15}$$

Now the same process has to be done on the denominator.
With the same method (LCM - least common multiple) we get:

$$\frac{1\cdot 2}{3\cdot 2}+\frac{1\cdot 3}{2\cdot 3}=\frac{2}{6}+\frac{3}{6}=\frac{5}{6}$$

Then the total answer is:

$$\frac{\frac{16}{15}}{\frac{5}{6}}=\frac{16}{15}\cdot\frac{6}{5}=\frac{96}{75}=\frac{96/3}{75/3}=\frac{32}{25}$$

We can see that in the last step of the equation, the factor has been simplified. To do this we use factoring. We break down the numbers into smaller factors or multiple prime numbers. Therefore we have:

$$\frac{96}{75}$$

=

$$\frac{3\cdot 32}{3\cdot 25}$$

We can now remove "3", or the multiplier, as it is on both sides of the fraction. So we have:

$$\frac{32}{25}$$

=

$$\frac{25}{25}+\frac{7}{25}=1\frac{7}{25}$$

In step 1 above we used Cross-Multiplication.

**Definition 1.6. Cross-Multiplication** is when we multiple the numerator by the reciprocal of the denominator.

So in this case we rewrite

$$\frac{\frac{16}{15}}{\frac{5}{6}}$$

or

$$\frac{16}{15}\div\frac{5}{6}$$

as

$$\frac{16}{15}\cdot\frac{6}{5}$$

As you can see all we are doing is turning

$$\frac{5}{6}$$

upside down: and multiplying it with

$$\frac{16}{15}$$

This gives:

$$\frac{96}{75}$$

In some cases it is possible to draw a **square root** of a fraction $s=\frac{p}{q}$, i.e. find a number $r\in\mathbb{Q}$ such that $r^2=s$. The square root is denoted $\sqrt{r}$.

**Example 1.10.** Consider the expression

$$(\sqrt{\frac{1}{9}}\times 2^4)+(\frac{1}{5}\times\sqrt{25})$$

.
To evaluate this expression, first consider separately the two parts on each side of the plus symbol.
The first part is

$$(\sqrt{\frac{1}{9}}\times 2^4)$$

and the second part is

$$(\frac{1}{5} \times \sqrt{25})$$

.
In addition, by definition of root,

$$\sqrt{\frac{1}{9}} = \frac{1}{3}$$

.
First part:

$$(\sqrt{\frac{1}{9}} \times 2^4) = \frac{1}{3} \times 16 = \frac{16}{3}$$

Second part:

$$(\frac{1}{5} \times \sqrt{25}) = \frac{1}{5} \times 5 = 1$$

Finally, add the first part and the second part:

$$\frac{16}{3} + 1 = \frac{19}{3}$$

**Example 1.11.** Consider the following fraction example, to be solved step by step:

$$\frac{\frac{4}{2} + (\frac{1}{4} \cdot \frac{5}{3})}{\frac{2}{6} \div \frac{1}{5}}$$

First we need to be aware of operator presedence, meaning that first we solve the brackets, then multiplication/division, then addition/subtraction and finally the main fraction.

$$(\frac{1}{4} \cdot \frac{5}{3}) = \frac{5}{12}$$

After solving the bracket we can proceed with adding

$$\frac{4}{2}$$

to

$$\frac{5}{12}$$

as there is no other action left for the nominator of the main fraction. So:

$$\frac{4}{2} + \frac{5}{12}$$

When adding fractions together we first have to find a common denominator, in this case 12 would work as

$$2 \cdot 6 = 12$$

So we multiply both the numerator and the denominator of that fraction by 6 and then add the two numerators of the fractions together, keeping the same denominator.

$$\frac{4}{2} + \frac{5}{12} = \frac{4 \cdot 6}{2 \cdot 6} + \frac{5}{12} = \frac{24}{12} + \frac{5}{12} = \frac{29}{12}$$

Now we have the top half of the fraction solved. We then proceed with dividing the two fractions of the bottom half. When dividing fractions we use the so called cross multiplication technique. This arithmetic trick is derived from the fact that if you divide a fraction by its duplicate you get 1. If you multiple a fraction by its reciprocal (it's reverse) you also get 1. Like so:

$$\frac{1}{2} \div \frac{1}{2} = 1$$

and

$$\frac{1}{2} \cdot \frac{2}{1} = 1$$

These functions always provide the same result and therefore we can turn the fraction we are dividing by upside down and multiply it to the other fraction as that is usually much easier.

We can therefore rewrite

$$\frac{2}{6} \div \frac{1}{5}$$

as

$$\frac{2}{6} \cdot \frac{5}{1} = \frac{10}{6}$$

We've now solved both halves of the original fraction and can therefore proceed to solve it, again with the cross multiplication technique as fractions are after all just divisions:

$$\frac{29}{12} \div \frac{10}{6} = \frac{29}{12} \cdot \frac{6}{10} = \frac{174}{120}$$

Now

$$\frac{174}{120}$$

is a pretty bad looking fraction and we'd preferably like to simplify it.

To do this we use factoring.

**Definition 1.7. Factoring** essentially means to break a number done into it's smallest factors or multipliable prime numbers.

In this case we get

$$\frac{2 \cdot 3 \cdot 29}{2 \cdot 3 \cdot 20}$$

These are the smallest prime numbers that can multiply together into 174 and 120 respectively.

A way of doing this in your head is by first dividing both numbers (174,120) by two. Which gives us:

$$\frac{2 \cdot 87}{2 \cdot 60}$$

and then dividing those numbers (87,60) by 3, since they can't be divided by 2. Dividing by 3 gives you

$$\frac{3 \cdot 29}{3 \cdot 20} = \frac{29}{20}$$

which is a lot nicer than

$$\frac{174}{120}$$

The reasoning behind this factoring simplification is that we can remove multipliers if they are on both sides of a fraction. This is because the result of a fraction where the numerator and the denominator are the same is always 1. Like so:

$$\frac{1}{1} = 1$$

or

$$\frac{2}{2} = 1$$

or

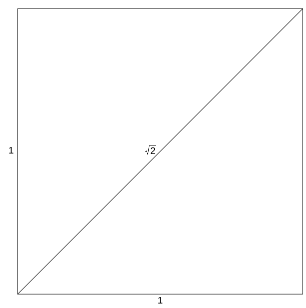
$$\frac{3}{3} = 1$$

The final answer therefore is

$$\frac{\frac{4}{2} + (\frac{1}{4} \cdot \frac{5}{3})}{\frac{2}{6} \div \frac{1}{5}} = \frac{29}{20}$$

## 1.5 The real line

Some obvious numbers are not fractions.
The set of numbers making up the real line is denoted by the symbol $\mathbb{R}$.




The diagonal of a rectangle with unit side lengths of $\sqrt{2}$, Note that $\sqrt{2}$ ia not a fraction.

### 1.5.1 Details

Some obvious numbers, which commonly occur, are not fractions. These are in between the rational numbers (fractions). Filling in the missing points to obtain a continuum results in the set of "real numbers".

Denoted by $\mathbb{R}$ the entire set of "real numbers"which corresponds to "filling in"the "missing piecesọf the line.

### 1.5.2 Examples

**Example 1.12.** If $C$ is the circumference of a circle and $D$ is the diameter and we define $\pi = \frac{C}{D}$ then $\pi$ is not a fraction.

**Example 1.13.** One example of a non fraction is the number e (Euler's number) which can be defined by

$$e = \sum_{n=0}^{\infty} \frac{1}{n!}$$

**Example 1.14.** If you have a right triangle with unit side length, what is the length of its hypotenuse and what class of numbers does it belong to?
An isosceles triangle is defined as having adjacent and opposite sides of same length, connected by a 90° angle. Unit side length of these, refers to a side length of

$$1$$

.
As we have a 90° angle, we can use Pythagoras' theorem:

$$a^2 + b^2 = c^2$$

With

$$a = adjacent$$

$$b = opposite$$

$$c = hypotenuse$$

So with

$$a, b = 1$$

:

$$c^2 = 1^2 + 1^2$$

$$c^2 = 1 + 1$$

$$c^2 = 2$$

We take the square root to get

$$c$$

$$c = \sqrt{2}$$

Now that we answered the first part of the question, it needs to be defined, which class of number

$$\sqrt{2}$$

belongs to.

$$\sqrt{2}$$

is an irrational number, and belongs thereby to the set of real numbers

$$\mathbb{R}$$

Real numbers can be imagined as points on an infinitely long line, which is also called the real line.

# 2 Data vectors

## 2.1 The plane

Pairs of numbers can be depicted as points on a plane.
The plane is normally denoted by $\mathbb{R}^2$.

### 2.1.1 Details

Pairs of numbers can be depicted as points on a plane.

**Definition 2.1.** A **plane** is a perfectly flat surface with no thickness and no end, it can extend forever in all directions. It has two-dimensions, length and width. We need two values to find a point on the plane.

Normally we talk about "the plane"as the collection of all pairs of numbers and denoted it by

$$\mathbb{R}^2 = \{(x,y) : x,y \in \mathbb{R}\}$$

, giving coordinates to each point.

### 2.1.2 Examples

**Example 2.1.** Plotting the point (2,4) in the x-y plane using R.

```
plot(2,4,xlim=c(0,6),ylim=c(0,6),xlab="x",ylab="y",cex=2)
text(2,4,"(2,4)",pos=4,cex=2)
```

Additional points can be added using the p*oints* function:

```
points(3,5, cex = 0.5) ## a point at (3,5)
```

If you have 2 sets of coordinates on a plane you can calculate the distance between the 2 points and graph the line connecting the points

**Example 2.2.** What is the distance between the 2 points (3,9) and (5,1)?
We will use the Pythagorean theorem:

$$d = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2}$$

We insert our values into the formula:

$$d = \sqrt{(5-3)^2 + (1-9)^2}$$

When we combine inside the parenthesis we get:

$$d = \sqrt{(2)^2 + (-8)^2}$$

Squaring both terms:

$$d = \sqrt{4+64}$$

Then we take the square root:

$$d = \sqrt{68}$$

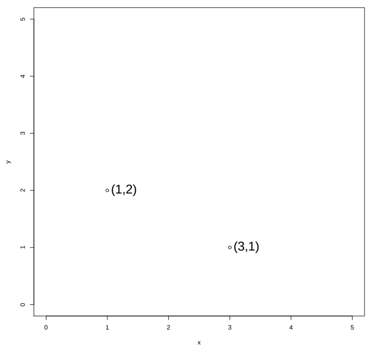
The result:

$$d = 8.2462$$

## 2.2 Simple plots in R

Graphing functions in R

- plot - plots a scatter plot (as a line plot)
- points - adds points to a plot
- text - adds text to a plot
- lines - adds lines to a plot




Points on a plane, drawn with R.

### 2.2.1 Examples

**Example 2.3.** `plot(2,3)`

gives a single plot and

`plot(2,3, xlim=c(0,5), ylim=c(0,5))`

gives a single plot but forces both axes to range from 0 to 5.

**Example 2.4.** The following R commands can be used to generate a plot with two points:

```
plot(1,2,xlim=c(0,5),ylim=c(0,5),xlab="x",ylab="y")
points(3,1)
text(1,2,"(1,2)",pos=4, cex=2)
text(3,1,"(3,1)",pos=4, cex=2)
```

**Example 2.5.** In this example, we plot 3 points. The first two points are by including vectors with a length of 2 as the x and y arguments of the plot function. The third plot was added with the points function. The second and third points were labeled using the text function and a line was drawn between them using the lines function.

*Note 2.1.* Note that if you are unsure of what format the arguments of an R function needs to be, you can call a help file by typing "?"before the function name (e.g. "?lines")

```
plot(c(2,3),c(3,4),xlim=c(2,6),ylim=c(1,5),xlab="x",ylab="y")
points(4,2)
text(3,4,"(3,4)",pos=4, cex=2)
text(4,2,"(4,2)",pos=4, cex=2)
lines(c(3,4), c(4,2))
```

## 2.3 Data

Data are usually a sequence of numbers, typically called a vector.

### 2.3.1 Details

When we collect data these are one or more sequences of numbers, collected into data vectors. We commonly think of these data vectors as columns in a table.

### 2.3.2 Examples

**Example 2.6.** In R, if the command

```
x <- c(4,5,3,7)
```

is given, then `x` contains a vector of numbers.

**Example 2.7.** Create a function in R, give it a name "Myfunction"which takes the sum of x,y.

```
Myfunction<- function(x,y) {
 sum(x,y)
}
```

If you input the vectors 1:3 and 4:7 into the function it will calculate the sum of `x<-(1+2+3)` and `y<-(4+5+6+7)` as follows

```
> Myfunction(1:3,4:7)
28
```

## 2.4 Indices for a data vector

If data are in a vector $x$, then we use indices to refer to individual elements.

### 2.4.1 Details

If $i$ is an integer then $x_i$ denotes the $i$'th element of $x$.

Note that although we do not distinguish (much) between row- and column vectors, usually a vector is thought of as a column. If we need to specify the type of vector, row or column, then for vector $x$, the column vector would be referred to as $x'$ and the row vector as $x^T$ (the **transpose** of the original).

### 2.4.2 Examples

**Example 2.8.** If $x = (4,5,3,7)$ then $x_1 = 4$ and $x_4 = 7$

**Example 2.9.** How to remove all indices below a certain value in R

```
x <- c(1,5,8,9,4,16,12,7,11)
x
[1] 1 5 8 9 4 16 12 7 11
y <- x[x>10]
y
[1] 16 12 11
```

**Example 2.10.** Consider a function that takes to vectors

$$a \in \mathbb{R}^n, b \in \mathbb{N}^m$$

as arguments with

$$n \geq m$$

and

$$1 \leq b_1, \ldots, b_m \leq n$$

. The function returns the sum

$$\sum_{i=1}^{m} a_{bi} \tag{1}$$

Long version:
fN <- function(a,b)
result <- sum(a[b])
return(result)

Short version:
| fN <- function(a,b) sum(a[b])|

## 2.5 Summation

We use the symbol Σ to denote sums.
In R, the sum function adds numbers.

### 2.5.1 Examples

**Example 2.11.** If $x = (4,5,3,7)$
then

$$\sum_{i=1}^{4} x_i = x_1 + x_2 + x_3 + x_4 = 4+5+3+7 = 19$$

and

$$\sum_{i=2}^{4} x_i = x_2 + x_3 + x_4 = 5+3+7 = 15.$$

Within R one can give the corresponding commands:

```
 x<-c(4,5,3,7)
 x
[1] 4 5 3 7
 sum(x)
[1] 19
sum(x[2:4])
[1] 15
```

# 3 More on algebra

## 3.1 Some Squares

If a and b are real numbers, then

$$(a+b)^2 = a^2 + 2ab + b^2$$

### 3.1.1 Details

If a, b are real numbers, then:
$(a+b)^2 = a^2 + 2ab + b^2$
This can be proven formally with the following argument:

$$\begin{aligned}(a+b)^2 &= (a+b)(a+b)\\ &= (a+b)a + (a+b)b\\ &= a^2 + ba + ba + b^2\\ &= a^2 + 2ab + b^2\end{aligned}$$

## 3.2 Pascal's Triangle

Pascal's triangle is a geometric arrangement of the binomial coefficients in a triangle

| | | | | |
|---|---|---|---|---|
| | | 1 | | |
| | 1 | | 1 | |
| 1 | | 2 | | 2 |

### 3.2.1 Details

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| $n = 0$: | | | | 1 | | | |
| $n = 1$: | | | 1 | | 1 | | |
| $n = 2$: | | 1 | | 2 | | 1 | |
| $n = 3$: | 1 | | 3 | | 3 | | 1 |

To build Pascal's triangle, start with "1"at the top, and then continue placing numbers below it in a triangular pattern. Each number is just the two numbers above it added together (except for the edges, which are all "1").

### 3.2.2 Examples

**Example 3.1.** The following function in R gives you the Pascal's triangle for $n = 0$ to $n = 10$.

```
fN <- function(n) formatC(n, width=2)
for (n in 0:10) {
    cat(fN(n),":", fN(choose(n, k = -2:max(3, n+2))))
    cat("\n")
}
```

```
 0 :  0  0  1  0  0  0
 1 :  0  0  1  1  0  0
 2 :  0  0  1  2  1  0  0
 3 :  0  0  1  3  3  1  0  0
 4 :  0  0  1  4  6  4  1  0  0
 5 :  0  0  1  5 10 10  5  1  0  0
```

```
 6 :  0  0  1  6 15 20 15  6  1  0  0
 7 :  0  0  1  7 21 35 35 21  7  1  0  0
 8 :  0  0  1  8 28 56 70 56 28  8  1  0  0
 9 :  0  0  1  9 36 84 126 126 84 36  9  1  0  0
10 :  0  0  1 10 45 120 210 252 210 120 45 10  1  0  0
```

Changing the numbers in the line `for(n in 0:10)` will give different portions of the triangle.

## 3.3 Factorials

We define the factorial of an integer n as
$n! = n \cdot (n-1) \cdot (n-2) \cdot \ldots \cdot 3 \cdot 2 \cdot 1$

### 3.3.1 Details

**Definition 3.1.** We define the factorial of an integer n as

$$n! = n \cdot (n-1) \cdot (n-2) \cdot \cdots \ldots \cdot 3 \cdot 2 \cdot 1.$$

### 3.3.2 Examples

**Example 3.2.** Suppose you have 6 apples, $\{a,b,c,d,e,f\}$ and you want to put each one into a different apple basket, $\{1,2,3,4,5,6\}$.

For the first basket you can choose from 6 apples $\{a,b,c,d,e,f\}$, and for the second basket you have then 5 apples to choose from and so it goes for the rest of the baskets, so for the last one you only have 1 apple to choose from.

The end result would then be: $6! = 6 \cdot 5 \cdot 4 \cdot 3 \cdot 2 \cdot 1 = 720$ possible allocations.

This could also be calculated in R with the factorial function:

```
factorial(6)
[1] 720
```

## 3.4 Combinations

The number of different ways one can choose a subset of size $x$ from a set of $n$ elements is determined using the following calculation:

$$\binom{n}{x} = \frac{n!}{x!\,(n-x)!}$$

### 3.4.1 Details

**Definition 3.2.** A **combination** is an un-ordered collection of distinct elements

Suppose we want to toss a coin $n$ times. In each toss we obtain head (H) or tail (T) resulting in a sequence of H,T,T,H, ... T.

How many of these possible sequences contain exactly $x$ tails? There are $n$ positions in the sequence, we can choose $x$ of these in $\binom{n}{x}$ ways and put our "Ts"in those positions. If the probability of landing tails then each one of these sequences with exactly $x$ tails has probability $p^x(1-p)^{n-x}$ so the total probability of landing exactly $x$ tails in $n$ independent tosses is

$$\binom{n}{x} = \frac{n!}{x!\,(n-x)!}.$$

For convenience we define 0! to be 1.

### 3.4.2 Examples

**Example 3.3.** Consider tossing a coin four times.

(a) How many times will this experiment result in exactly two tails?

There are a total of 16 possible sequences of heads and tails from four tosses. These can simply all be written down to answer a question like this.

We get two tails in 6 of these tosses. We can explicitly write the corresponding combinations of two tails as follows

```
HHTT
HTHT
HTTH
THTH
TTHH
THHT
```

(b) How many times you will end up with 1 tail? The answer is 4 times and the output can be written as;

```
HHHT
HTHH
THHH
HHTH
```

The case of a single tail is easy: The single tail can come up in any one of four positions.

## 3.5 The binomial theorem

$$(a+b)^n = \sum_{x=0}^{n} \binom{n}{x} a^x b^{n-x}$$

### 3.5.1 Details

If a and b are real numbers and n is an integer then the expression $(a+b)^n$ can be expanded as:
$(a+b)^n = a^n + \binom{n}{1}a^{n-1}b + \binom{n}{2}a^{n-2}b^{+} \ldots + \binom{n}{n-1}ab^{n-1} + b^n$
$(a+b)^n = \sum_{i=1}^{n}\binom{n}{x}a^x b^{n-x}$
This can be seen by looking at $(a+b)^n$ as a product of n parentheses and multiply these by picking one item (a or b) from each. If we picked $a$ from x parentheses and $b$ from $(n-x)$, then the product is $a^x b^{n-x}$. We can choose the $x$ $a$'s in a total of $\binom{n}{x}$ ways so the coefficient of $a^x b^{n-x}$ is $\binom{n}{x}$.

### 3.5.2 Examples

**Example 3.4.** Since

$$(a+b)^n = \sum_{x=0}^{n} \binom{n}{x} a^x b^{n-x},$$

it follows that

$$2^n = (1+1)^n = \sum_{x=0}^{n} \binom{n}{x}$$

i.e.

$$2^n = \binom{n}{0} + \binom{n}{1} + \binom{n}{2} \ldots + \binom{n}{n}$$

# 4 Discrete random variables and the binomial distribution

## 4.1 Simple probabilities

### 4.1.1 Details

Of all the possible 3-digit strings, $\binom{3}{x}$ of them have $x$ heads. So the probability of landing $x$ heads is $\binom{3}{x}p^x(1-p)^{3-x}$.

### 4.1.2 Examples

**Example 4.1.** Consider a biased coin which has probability $p$ of landing heads up. If we toss this coin 3 independent times the possible outcomes are:

| *sequence* | *probability* | *Numberofheads* |
|---|---|---|
| $HHH$ | $p \cdot p \cdot p = p^3$ | 3 |
| $HHT$ | $p^2(1-p)$ | 2 |
| $HTH$ | $p^2(1-p)$ | 2 |
| $HTT$ | $p(1-p)^2$ | 1 |
| $THH$ | $p^2(1-p)$ | 2 |
| $THT$ | $p(1-p)^2$ | 1 |
| $TTH$ | $p(1-p)^2$ | 1 |
| $TTT$ | $(1-p)^3$ | 0 |

**Example 4.2.** It is also possible to aggregate these values into a table and describe only the number of heads obtained:

| heads | probability $p(x)$ |
|---|---|
| 0 | $(1-p)^3$ |
| 1 | $3p(1-p)^2$ |
| 2 | $3p^2(1-p)$ |
| 3 | $p^3$ |

If we are only interested in the number of heads, then this table describes a **probability mass function** $p$, namely the probability $p(x)$ of every possible outcome $x$ of the experiment.

**Example 4.3.** Given that a year is 365 days and each day has the same probability of being someone's birthday. What's the probability of at least 2 people sharing a birthday in a group of 25 people?

Now, calculating each of the possible outcomes could become very tedious. That is calculating the odds that 2 people share a birthday, 3 people, 4 people, etc. So instead we try to find out the odds that no one in the group shares a birthday and subtract those odds from 1 (100%).

First, let's look at the odds of only two people having distinct birthdays.

$$\frac{365}{365} \cdot \frac{364}{365} = 0.9973$$

Person one can be born on any day and the odds of having a distinct birthday are therefore 1. The next person can be born on everyday but the 1 the other person was born on, so 364 days.

Now let's say we add the 3rd person and calculate his/her odds of having a distinct birthday.

$$\frac{365}{365} \cdot \frac{364}{365} \cdot \frac{363}{365} = 0.9918$$

This can also be rewritten as

$$\frac{365 \cdot 364 \cdot 363}{365^3}$$

And we can do this on and on for all the 25 people we are interested in. But that may also become a bit tedious. So we use factorials instead. So instead of doing

$$\frac{365 \cdot 364 \cdot 363 \ldots \cdot 341}{365^{25}}$$

we do

$$\frac{\frac{365!}{340!}}{365^{25}} = 0.4313$$

Essentially the division of factorials here removes all the values < 341, leaving 340, 339, 338 ... 1

Now remember this is the probability that no one shares a birthday. So when we subtract this from 1 we get

$$1 - 0.4313 = 0.5687$$

or roughly 57% odds of at least 2 people in a group of 25 sharing the same birthday.

## 4.2 Random variables

A random variable is a concept used to denote the outcome of an experiment before it is conducted.

### 4.2.1 Examples

**Example 4.4.** Let $X$ denote the number of heads in a coin tossing experiment. We can then talk about the probabilities of certain events such as obtaining two heads, i.e. $X = 2$. We write this as

$$P[X = 2] = \binom{n}{2} p^2 (1-p)^{n-2}$$

In general:

$$P[X = x] = \binom{n}{x} p^x (1-p)^{n-x}$$

where $x = 0, 1, \ldots.., n$

### 4.2.2 Handout

**Definition 4.1.** A **random variable**, $X$, is a function defined on a sample space, with outcomes in the set of real numbers.

It is simpler to think of a random variable as a symbol used to denote the outcome of an experiment before it is conducted.

*Note 4.1.* Note that it is **essential** to distinguish between upper case and lower case letters when writing these probabilities - it makes no sense to write $P[x = x]$.

*Note 4.2.* Random variables are generally denoted by upper case letters such as $X$, $Y$ and so on.

*Note 4.3.* To see how a random variable is a function, it is useful to consider the actual outcomes of two coin tosses. These outcomes can be denoted $\{HH, HT, TH, TT\}$. Now consider a random variable $X$ which describes the number of heads obtained. This random variable attributed 2 to the outcome $HH$ and 0 to $TT$, i.e. $X$ is a function with $X(HH) = 2$, $X(HT) = X(TH) = 1$ and $X(TT) = 0$.

## 4.3 Simple surveys with replacement

If we randomly draw individuals (with replacement) and ask a question with two possible answers (positive or negative), then the number of positive answers will come from a binomial distribution.

### 4.3.1 Examples

**Example 4.5.** Suppose we are participating in a lottery. We pick a number from a lottery bowl (a simple random sample). We can put the number aside, or we can put it back into the bowl. If we put the number back in the bowl, it may be selected more than once; if we put it aside, it can be selected only one time.

**Definition 4.2.** When an element can be selected more than one time, we are sampling **with replacement**.

**Definition 4.3.** When an element can be selected only one time, we are sampling **without replacement**.

## 4.4 The binomial distribution

If we toss a biased coin $n$ independent times, each with probability $p$ of landing heads up, then the probability of obtaining $x$ heads is

$$\binom{n}{x} p^x (1-p)^{n-x}$$

### 4.4.1 Examples

**Example 4.6.** Suppose we toss a coin, with probability $p$ of landing on heads $n$ times obtaining a sequence of Hs (when it lands heads) and Ts (when it lands tails). Any sequence,

$$HTH...HTHHH$$

which has $x$ heads ($H$) and $n-x$ tails ($T$), has the probability $p^x(1-p)^{n-x}$. There are exactly $\binom{n}{x}$ such sequences, so the total probability of landing $x$ heads in $n$ tosses is

$$\binom{n}{x} p^x (1-p)^{n-x}.$$

**Example 4.7.** Let the probability that a certain football club wins a match be equal to 0.4.If the total number of matches played in the season is 30, what is the probability that the football club wins the match 10% of the time?

We first calculate the number of times a match was played and won by multiplying the percentage of wins by the number of matches played.

10% of 30 times = 3 times

We can now proceed to calculate the probability that they will win the match given that their probability of a winning is 0.4 if they play 3 times in a season. This can be computed as follows:

$$\binom{30}{3} \times (0.4)^3 \times (1-0.4)^{30-3}$$
$$= 0.000265$$

This can be calculated in R using the code below:

```
dbinom(3,30,0.4)

[1] 0.0002659437
```

This is equal to the manual calculation using the binomial theorem.

**Example 4.8.** Suppose a youngster puts his shirt on by himself every day for five days. The probability that he puts it on the right way each time is $p = 0.2$. We let $X$ be a random variable that describes the number of times the youngster puts his shirt on the right way. The youngster can either put the shirt on the wrong or the right way so $X$ follows the binomial distribution with the parameters $p = 0.2$ (the probability of a successful trial) and $n = 5$ (number of trials). We can now calculate for example the probability that the youngster will put it on the right way for at least 4 days.

Putting the shirt on the right way for at least 4 days means that the youngster will either put it on the right way for either four or five days (at least four or more days of five days total). We thus have to calculate the probability that the youngster will put his shirt on the right way for 4 and 5 days separately and then we add it together. We can write this process as follows:

$$P(X \geq 4) = P(X = 4) + P(X = 5)$$

$$= \binom{5}{4} \times 0.2^4 \times (1-0.2)^{5-4} + \binom{5}{5} \times 0.2^5 \times (1-0.2)^{5-5}$$

$$= 5 \times 0.2^4 \times 0.8^1 + 1 \times 0.2^5 \times 0.8^0$$

$$= 5 \times 0.2^4 \times 0.8 + 0.2^5 \times 1$$

$$= 5 \times 0.8 \times 0.2^4 + 0.2^5$$

$$= 4 \times 0.2^4 + 0.2^5$$

$$= 4 \times 0.0016 + 0.00032$$

$$= 0.00672$$

The probability that the youngster will put his shirt on the right way for at least four out of five is thus 0,7%.

This is possible to calculate in R in a several ways, either using the command dbinom or pbinom. The command dbinom calculates

$$P(X = k)$$

and the command pbinom calculates

$$P(X \leq k)$$

where $k$ is the number of successful trials. If $n$ is the number of trials and $p$ is the probability of a successful trials then the commands are used by writing: *dbinom*($k$,$n$,$p$) and pbinom($k$,$n$,$p$).

To calculate the probability that the youngster will put his shirt on the right way for at least four days of five we thus write the command:

```
dbinom(4,5,0.2) + dbinom(5,5,0.2)
```

which gives 0.00672.

This is the same as writing:

```
dbinom(c(4,5),5,0.2)
```

or

```
dbinom(4:5,5,0.2)
```

which give two separate numbers: 0.00640 and 0.00032 which can be added together to get 0.00672.

There is also a command to add them together for us:

```
sum(dbinom(c(4,5),5,0.2))
```

or

```
sum(dbinom(4:5,5,0.2))
```

They give the answer 0.00672.

The fourth way of calculating this in R is to use pbinom. As said before pbinom calculates

$$P(X \leq k)$$

where $k$ is the number of successful trials. Here we want to calculate the probability that the youngster will put his shirt on the right way in 4 or 5 times (of 5 total) so the number of successful trials is 4 or greater. That means we want to calculate

$$P(X \geq 4)$$

which equals

$$1 - P(X \leq 3)$$

. We thus put $k$ as 3 and the R command will be:

```
1 - pbinom(3,5,0.2)
```

which also gives 0.00672.

**Example 4.9.** In a certain degree program, the chance of passing an examination is 20%. What is the chance of passing at most 2 exams if the student takes five exams?

Solution:
In this problem, we compute the chance of a student passing, 0.1 or 2 exams.This is given by,

$$\begin{aligned} p(X = 0 \text{ or } 1 \text{ or } 2) &= \binom{5}{0} 0.2^0 0.8^5 + \binom{5}{1} 0.2^1 0.8^4 + \binom{5}{2} 0.2^2 0.8^3 \\ &= 1 \times 0.2^0 0.8^5 + 5 \times 0.2^1 0.8^4 + 10 \times 0.2^2 0.8^3 \\ &= 0.32768 + 0.4096 + 0.2048 \\ &= 0.94208 \end{aligned}$$

In the R console, we can use the command, `sum(dbinom(c(0:2),5,0.2))`, which also gives

$$0.94208.$$

The same answer is obtained with

```
dbinom(0,5,0.2)+dbinorm(1,5,0.2)+dbinom(2,5,0.2)
```

and with

```
pbinom(2,5,0.2)
```

**Example 4.10.** Consider the probability of someone jumping off a cliff is 0.35. Suppose we randomly selected four individuals to participate in the cliff jumping activity. What is the chance that exactly one of them will jump off the cliff?

Consider a scenario where one person jumps:
P (A =jump , B = refuse, C = refuse, D = refuse)
= P (A =jump) P (B = refuse) P (C = refuse) P (D = refuse)
$= (0.35)(0.65)(0.65)(0.65) = (0.35)^1(0.65)^3 = 0.096$

But there are three other scenarios( B, C, or D) in which one only person decides to jump. In each of these cases, the probability is again 0.096. These four scenarios exhaust all the possible ways that exactly one of the four people jumps:
$4 \cdot (0.35)^1(0.65)^3 = 0.38.$

In the R console we can use the command: `dbinom(1,4,0.35)` which gives the answer as 0.384475.

## 4.5 General discrete probability distributions

A general discrete probability distribution can be described by a list of all possible outcomes and associated probabilities.

### 4.5.1 Details

A general discrete probability distribution is described by the possible outcomes

$$x_1, x_2, \ldots$$

and associated probabilities, denoted by $p_1, p_2, \ldots$ or $p(x_1), p(x_2), \ldots$

If a random variable $X$ has this distribution, then we can write

$$P[X = x_i] = p(x_i) = p_i$$

or in general

$$P[X = x] = p(x)$$

where it is understood that $p(x) = 0$ if $x$ is not one of these $x_i$.

### 4.5.2 Examples

**Example 4.11.** If $X$ is the number of heads ($H$) before obtaining the first tail ($T$) when tossing an unbiased coin 4 independent times, then the possible basic outcomes are:

| In binary | Toss 1 2 3 4 | #$H$ before $T$ |
|---|---|---|
| 0000 | H H H H | 4 |
| 0001 | H H H T | 3 |
| 0010 | H H T H | 2 |
| 0011 | H H T T | 2 |
| 0100 | H T H H | 1 |
| 0101 | H T H T | 1 |
| 0110 | H T T H | 1 |
| 0111 | H T T T | 1 |
| 1000 | T H H H | 0 |
| 1001 | T H H T | 0 |
| 1010 | T H T H | 0 |
| 1100 | T H T T | 0 |
| 1101 | T T H H | 0 |
| 1110 | T T T H | 0 |
| 1111 | T T T T | 0 |

Since the coin is unbiased, each of these has the same probability of occurring. We can now count sequences to find the number of possibilities of a particular number of heads, $H$, before a tail in 4 coin tosses and thus obtain the corresponding probabilities as:

| Number of tosses before a heads $x$ | Probability $p(x)$ |
|---|---|
| 0 | $\frac{8}{16} = \frac{1}{2}$ |
| 1 | $\frac{4}{16} = \frac{1}{4}$ |
| 2 | $\frac{2}{16} = \frac{1}{8}$ |
| 3 | $\frac{1}{16}$ |
| 4 | $\frac{1}{16}$ |

## 4.6 The expected value or population mean

The expected value is the sum of the possible outcomes, weighted with the respective probabilities (discrete variable). Think of this in terms of an urn full of marbles, each labelled with number.

### 4.6.1 Details

If the possible outcomes are $x_1, x_2 \ldots$ with probabilities $p_1, p_2 \ldots$ then the expected value is

$$\mu = x_1 \cdot p_1 + x_2 \cdot p_2 + \ldots.$$

The fact that this is the only sensible definition of an expected value follows from considering random draws from a finite population where there are $n_i$ possibilities of obtaining the value $x_i$. If we set $n = \sum x_i$ and $p_i = n_i/n$ then the expected value above is the simple average of all the numbers in the original population.
In the case of the **binomial distribution** with $n$ trials and success probability $p$ it turns out that

$$\mu = n \cdot p$$

If $X$ is the corresponding random variable, we denote this quantity by $E[X]$.

### 4.6.2 Examples

**Example 4.12.** If we toss a fair coin 10 independent times, we expect on average $np = 10 \cdot \frac{1}{2} = 5$ heads.

**Example 4.13.** Toss a fair die and pay \$60 if a six comes up and nothing otherwise. The expected outcome is

$$\frac{5}{6} \cdot \$0 + \frac{1}{6} \cdot \$60 = \$10.$$

**Example 4.14.** In Las Vegas, a particular sports bet has about a 30% chance of winning. If the bet wins, the bettor will win 15 dollars. If the bet loses, the bettor will lose 10 dollars. The expected return of placing one of these bets is -2.50 dollars.
Detailed calculation:

$$\$15 \cdot 0.3 - \$10 \cdot 0.7 = -\$2.5$$

**Example 4.15.** Class starts at 8:00 and the last bus that will get you to class on time leaves at 7:30. The teacher has a policy that if you are late to class 6 of the 30 classes, then she drops your final grade by 1/10 points. You know that if you set your alarm for 7:15, you miss the 7:30 bus approximately every fourth time, but if you set it for 7:10, you'll only miss the bus approximately every eighth time. If you set it for 7:00, you'll only miss the bus every one hundredth time.

Part A: Assuming you try to go to class every time, can you expect to have your grade dropped in the following scenarios?
1 - You set your alarm for 7:15 throughout the duration of the class.
2 - You set your alarm for 7:15 until you reach 5 missed classes, then switch to 7:10.
3 - You set your alarm for 7:15 until you reach 5 missed classes, then switch to 7:00.

Part B: What is your expected grade in the course, assuming you would have had a 7/10 without the late penalty, and:
1 - You would never choose the first alarm-clock strategy and you would most likely choose scenario 2 (let's say 9/10 times), but there's a small chance you might choose the 3rd strategy (let's say 1/10 times).
2 - You would never choose the first alarm-clock strategy and you would most likely choose scenario 3 (let's say 9/10 times), but there's a small chance you might choose the

2nd strategy (let's say 1/10 times).

Answers:
A1 - Let's call X our random variable, which we want to be the number of times we make it to class on-time. With the alarm set to 7:15 we expect to make it to class on-time:

$$E[X] = 30 \times (1 - \frac{1}{4}) = 22\frac{1}{2}$$

You're grade would most likely be dropped.

A2 - First we need to see how many classes we go to before we reach the 5-late-classes threshhold:

$$E[X] = n \times (1 - \frac{1}{4}) = n - 5$$

$$E[X] = n((1 - \frac{1}{4}) - 1) = -5$$

$$E[X] = n = \frac{-5}{-\frac{1}{4}}$$

$$E[X] = n = \frac{20}{1} = 20$$

So, the night before our 21st class, you get worried and change alarm-clock strategies. If you set it at 7:15 for the rest of the course (10 classes), you will be on time:

$$E[X] = 15 + 10 \times (1 - \frac{1}{8}) = 23\frac{3}{4}$$

You're grade would most likely be dropped.

A3: If you instead start setting the alarm clock for 7:00 for the rest of the course, you will be on time:

$$E[X] = 15 + 10 \times (1 - \frac{1}{100}) = 24\frac{1}{9}$$

You're grade would most likely NOT be dropped.

Part B: **This seems to contain errors** In Part A, we calculated the mean of several binomial distributions that described the expected number of days that you will arrive on-time to class. Each distribution corresponded to a different alarm-setting scenario. In this part, we are describing a different binomial distribution. It describes your expected grade. Therefore, the grade is the outcome n, weighted by the probability of you choosing the particular alarm-clock setting procedure:

$$1 - E[X] = 0 \times 6 + 0.9 \times 6 + 0.1 \times 7 = 6.1$$

$$1 - E[X] = 0 \times 6 + 0.1 \times 6 + 0.9 \times 7 = 6.9$$

Note that the probabilities of these three choices (0 + 0.9 + 0.1) must equal 1, since these are the only three choices defined.

## 4.7 The population variance

The (population) variance, for a discrete distribution, is

$$\sigma^2 = E\left[(X-\mu)^2\right] = (x_1-\mu)^2 p_1 + (x_2-\mu)^2 p_2 + \ldots$$

where it is understood that the random variable $X$ has this distribution and $\mu$ is the expected value.

In the case of the binomial distribution, it turns out that:
$\sigma^2 = np(1-p)$

### 4.7.1 Details

**Definition 4.4.** If $\mu$ is the expected value, then the **variance of a discrete distribution** is defined as

$$\sigma^2 = (x_1-\mu)^2 p_1 + (x_2-\mu)^2 p_2 + \ldots.$$

If a random variable $X$ has associated probabilities, $p_i = P[X = x_i]$, then one can equivalently write

$$\sigma^2 = V[X] = E\left[(X-\mu)^2\right].$$

### 4.7.2 Examples

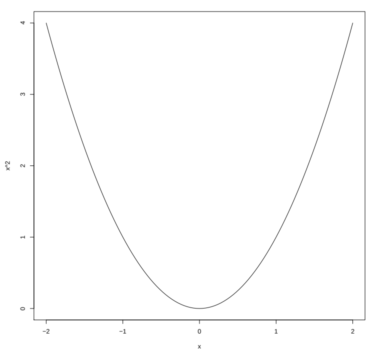
**Example 4.16.** In the case of the binomial distribution, it turns out that:

$$\sigma^2 = np(1-p).$$

# 5 Functions

## 5.1 Functions of a single variable

A function describes the relationship between variables.
Examples:
$f(x) = x^2$
$y = 2 + 3 \cdot x^4$



### 5.1.1 Details

Functions are commonly used in statistical applications, to describe relationships.

**Definition 5.1.** A **function** describes the relationship between variables. A variable $y$ is described as a function of a variable $x$ by completely specifying how $y$ can be computed for any given value of $x$.

An example could be the relationship between a dose level and the response to the dose.

The relationship is commonly expressed by writing either $f(x) = x^2$ or $y = x^2$.

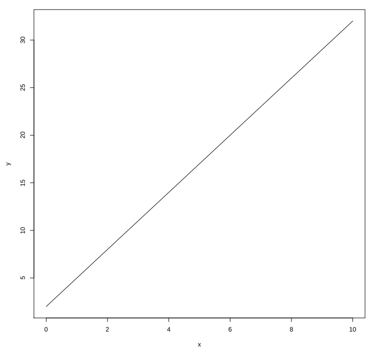
Usually names are given to functions, i.e. to the relationship itself. For example, $f$ might be the function and $f(x)$ could be its value for a given number $x$. Typically $f(x)$ is a number but $f$ is the function, but the sloppy phrase "the function $f(x) = 2x+4$"is also common.

### 5.1.2 Examples

**Example 5.1.** $f(x) = x^2$ or $y = x^2$ specifies that the computed value of $y$ should always be $x^2$, for any given value of $x$.

## 5.2 Functions in R

A function can be defined in R using the "function"command



## 5.3 Ranges and plots in R

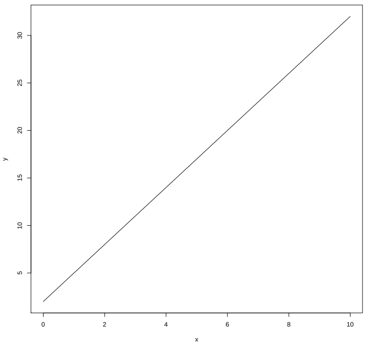
Functions in R can commonly accept a range of values and will return a corresponding vector with the outcome.

### 5.3.1 Examples

**Example 5.2.**
```
f <- function(x) {return(x*12)}
x <- seq (-5,5,0,1)
y <- f(x)
plot {(x,y) type= 'l'}
```

## 5.4 Plotting functions

In statistics, the function of interest is commonly called the response function. If we write Y=f(x), the outcome Y is usually called the response variable and x is the explanatory variable. Function values are plotted on vertical axis while x values are plotted on horizontal axis. This plots Y against x.




### 5.4.1 Examples

**Example 5.3.** The following R commands can be used to generate a plot for function; Y= 2+3x

```
x<- seq(0:10)
g <- function(x){
+ yhat <- 2+3*x
+ return(yhat)
+ }

x<-seq(0,10,0.1)
y<- g(x)
plot(x,y,type="l", xlab="x",ylab="y")
```

## 5.5 Functions of several variables

### 5.5.1 Examples

**Example 5.4.**

$$z = 2x + 3y + 4 \tag{2}$$

$$v = t^2 + 3x \tag{3}$$

$$w = t^2 + 3b * x \tag{4}$$

# 6 Polynomials

## 6.1 The general polynomial

The general polynomial:
$p(x) = a_0 + a_1x + a_2x^2 + \ldots + a_nx^n$
The simplest: $p(x) = a$

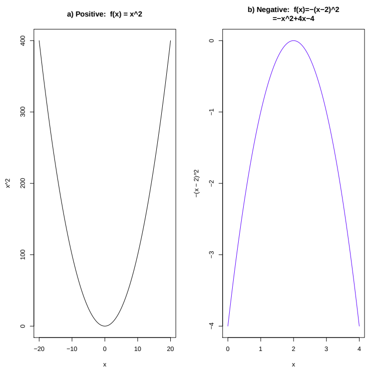
### 6.1.1 Details

**Definition 6.1.** A **polynomial** describes a specific function consisting of linear combinations of positive integer powers of the explanatory variable.

The general form of a polynomial is:
$p(x) = a_0 + a_1x + a_2x^2 + \ldots + a_nx^n$
The simplest of these is the constant polynomial $p(x) = a$.

## 6.2 The quadratic

The general form of the quadratic (parabola) is $p(x) = ax^2 + bx + c$.
The simplest quadratic is $p(x) = x^2$



Parabolas: Quadratic functions.

### 6.2.1 Details

The quadratic polynomial of the form $p(x) = ax^2 + bx + c$ describes a parabola when points $(x, y)$ with $y = p(x)$ are plotted.

The simplest parabola is $p(x) = x^2$ (Fig. a) which is always non-negative $p(x) \geq 0$ and $p(x) = 0$ only when $x = 0$.

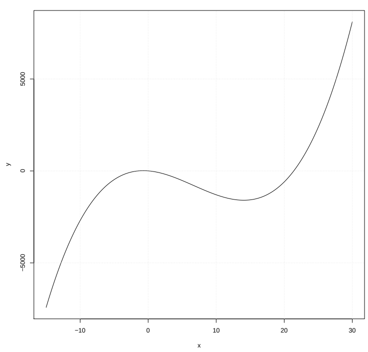
*Note 6.1.* Note that $p(-x) = p(x)$ since $(-x)^2 = x^2$.

If the coefficient at the highest power is negative, then the parabola is "upside down"(Fig. b).

This is sometimes used to describe a response function.

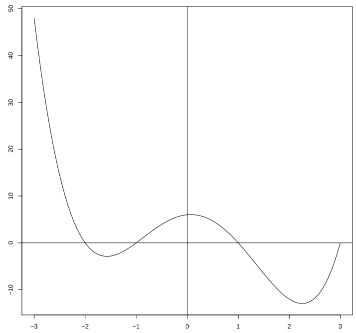
## 6.3 The cubic

The general form of a cubic polynomial is:
$p(x) = ax^3 + bx^2 + cx + d$


$y = x^3 - 20x^2 - 30x - 4$

## 6.4 The Quartic

The general form of the quartic polynomial is $p(x) = ax^4 + bx^3 + cx^2 + dx + e$



The general shape. Here we used the following equation $y = x^4 - x^3 - 7x^2 + x + 6$

## 6.5 Solving the linear equation

If the value of $y$ is given and we know that $x$ and $y$ are on a specific line so that $y = a + bx$, then we can find the value of $x$

### 6.5.1 Details

If a value of $y$ is given and we know that $x$ and $y$ lie on a specific straight line so that $y = a + bx$, then we can find the value of $x$ by considering $y = a + bx$ as an equation to be solved for $x$, since $y$, $a$ and $b$ are all known.

The general solution is found through the following steps:

- Equation: $y = a + bx$
- Subtract $a$ from both sides
  - $y - a = bx$
  - $bx = y - a$
- Divide by $b$ on both sides if $b$ is not equal to 0.
  - $x = \frac{1}{b}(y - a)$.

## 6.6 Roots of the quadratic equation

The general solution of $ax^2 + bx + c = 0$ is given by $x = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}$.

### 6.6.1 Details

Suppose we want to solve $ax^2 + bx + c = 0$, where $a \neq 0$.
The general solution is given by the formula

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

if $b^2 - 4ac \geq 0$. On the other hand, if $b^2 - 4ac < 0$, the quadratic equation has no real solution.

### 6.6.2 Examples

**Example 6.1.** Solve $x^2 - 3x + 2 = 0$
Putting this into the context of the formulation $ax^2 + bx + c = 0$, the constants are;
$a = 1, b = -3, c = 2$
Inserting this into the formula for the roots gives:

$$\begin{aligned}
x &= \frac{-(-3) \pm \sqrt{(-3)^2 - 4(1)(2)}}{2(1)} \\
x &= \frac{3 \pm \sqrt{9-8}}{2} \\
x &= \frac{3 \pm \sqrt{1}}{2} \\
x &= \frac{3+1}{2}, \frac{3-1}{2} \\
x &= \frac{4}{2}, \frac{2}{2} \\
x &= 2, 1
\end{aligned}$$

**Example 6.2.** Find the roots of the following polynomial

$$3x^4 + 14x^2 + 15$$

We can use the quadratic equation to solve for the roots of this polynomial if we substitute a variable for

$$x^2$$

Let's use the letter

$$a$$

$$3a^2 + 14a + 15$$

We then plug the constants in to the quadratic equation.

$$x = \frac{-(14) \pm \sqrt{14^2 - (4)(3)(15)}}{(2)(3)}$$

which simplifies to

$$\frac{-(14) \pm \sqrt{196-180}}{6}$$

which equals

$$-1\frac{2}{3}$$

and

$$-3$$

.

Then, since we substituted a for

$$x^2$$

we need to take the square root of these values to get the roots of the polynomial. So,

$$x_{1,2} = \pm\sqrt{-1\frac{2}{3}}$$

and

$$x_{3,4} = \pm\sqrt{3}$$

# 7 Simple data analysis in R

## 7.1 Entering data; dataframes

Several methods exist to enter data into R:

1. Enter directly: x<-c(4,3,6,7,8)
2. Read in a single vector: x<-scan("filename")
3. Use: x<-read.table("file address")

### 7.1.1 Details

The most direct method will not work if there are a lot numbers; therefore, the second method is to read in a single vector by x<-scan("filename"), "filename- text string, either a full path name or refers to a file in the working directory.

The scan() command returns a vector, but the read.table() command returns a dataframe, which is a rectangular table of data whose columns have names. A column can be extracted from a data frame, e.g., with x<- dat$a where"dat"is the name of the data frame and "a"is the name of a column.

*Note 7.1.* Note that for read.table("file address"), "file address"refers to the location of the file. Thus, it can be the URL or the complete file directory depending on where the table is stored.

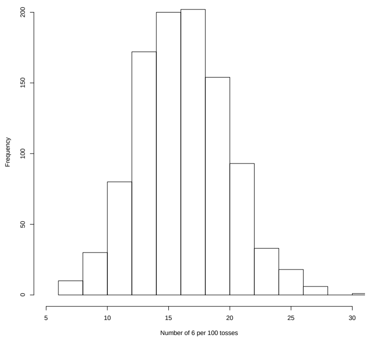
### 7.1.2 Examples

**Example 7.1.** Below are three examples using R code to enter data

1. x<-c(4,3,6,7,8)
2. x<-scan("lecture 70.txt")
3. x<-read.table("http://notendur.hi.is/ gunnar/kennsla/alsm/data/set115.dat", header=T)

## 7.2 Histograms

A histogram is a graphical display of tabulated frequencies, shown as bars.
In R use the command: hist()




### 7.2.1 Examples

A histogram is a graphical display of tabulated frequencies, shown as bars.

**Example 7.2.** If we toss a fair die 100 times and record the number of sixes, then we can view that as the outcome of a random variable $X$, which is binomial with $n = 100$ and $p = \frac{1}{6}$, i.e $X \sim b(n = 100, p = \frac{1}{6})$

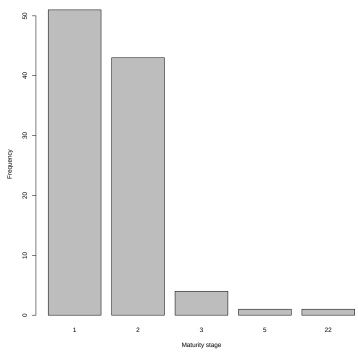
Now this can be done e.g. 1000 times to obtain numbers, $x_1, \ldots, x_{1000}$. Within R this can be simulated using

```
x <- rbinom(1000,100,1/6)
```

We would typically plot these using a histogram, e.g.
hist(x)
or
hist(x,nclass=50);l

## 7.3 Bar Charts

The bars in a bar chart usually correspond to frequencies in categories and are therefore kept apart.




### 7.3.1 Details

A bar chart is similar to the histogram but is used for categorical data.

## 7.4 Mean, standard error, standard deviations

### 7.4.1 Details

The most familiar measure of central tendency is the arithmetic mean.

**Definition 7.1.** An **arithmetic mean** is the sum of the values divided by the number values, typically expressed as:

$$\bar{y} = \frac{\sum_{i=1}^{n} y_i}{n}$$

**Definition 7.2.** The **sample variance** is a measure of the spread of a set of values from the mean value:

$$s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2$$

The sample standard deviation is more commonly used as a measure of the spread of a set of values from the mean value.

**Definition 7.3.** The **standard deviation** is the square root of the variance and may be expressed as:

$$s = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2}$$

**Definition 7.4.** The **standard error** is a method used to indicate the reliability of the sample mean:

$$SE_{\bar{y}} = \sqrt{\frac{s^2}{n}}$$

If a vector x in R contains an array of numbers then:
`mean(x)` returns the average, $\bar{x}$
`sd(x)` returns the standard deviation,$s$
`var(x)` returns the variance, $s^2$

We may also want to use several other related operations in R:
`median(x)`, the median value in vector x
`range(x)`, which list the range: `max(x)-\verb`$min(x)$;
If the variable `x` contains discrete categories, `table(x)` returns counts of the frequency in each category.

## 7.5 Scatter plots and correlations

If we have paired explanatory and response data we are often interested in seeing if a relationship exists between them. To do this, we first plot the data in a scatter plot.

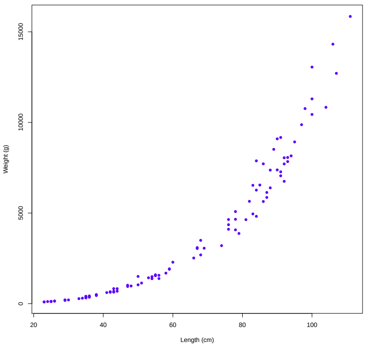



Figure: Scatter plot showing the length-weight relationship of fish species "X". Data source : Marine Resource Institution - Iceland.

### 7.5.1 Details

A first step in analyzing data is to prepare different plots. The type of variable will determine the type of plot. For example, when using a scatter plot both the explanatory and response data should be continuous variables.

The equation for the Pearson correlation coefficient is:

$$r_{x,y} = \frac{\sum_{i=1}^{n}(x_i - \overline{x})(y_i - \overline{y})}{\sum_{i=1}^{n}(x_i - \overline{x})^2 \sum_{i=1}^{n}(y_i - \overline{y})^2},$$

where $\overline{x}$ and $\overline{y}$ are the sample means of the x- and y-values.
The correlation is always between -1 and 1.

### 7.5.2 Examples

The following R commands can be used to generate a scatter plot for vectors x and y

**Example 7.3.**
```
plot(x,y)
```

# 8 Indices and the apply commands in R

## 8.1 Giving names to elements

We can name elements of vectors and data frames in R using the "names"command.

### 8.1.1 Examples

**Example 8.1.**
```
X<-c(41, 3, 73)
names(X)<-c("One", "Two", "Three")
```

View the results by simply typing "X"and the output of "X"is given as follows:

```
X
One Two Three
41 3 73
```

With this we can refer to the elements by name as well as locations using...

```
X[1]
One

X["Three"]
Three
73
```

## 8.2 Regular matrix indices and naming

A matrix is a table of numbers. Typical matrix indexing: mat[i,j], mat[1:2,] etc

A matrix can have row and column names Indexing with row and column names: mat["a","B"]

### 8.2.1 Details

**Definition 8.1.** A **matrix** is a (two-dimensional) table of numbers, indexed by row and column numbers.

*Note 8.1.* Note that a matrix can also have row and column names so that the matrix can be indexed by its names rather than numbers.

### 8.2.2 Examples

**Example 8.2.** Consider a matrix with 2 rows and 3 columns. Consider extracting first element (1,2), then all of line 2 and then columns 2-3 in an R session:

```
mat<-matrix(1:6,ncol=3)
mat
     [,1] [,2] [,3]
[1,] 1 3 5
[2,] 2 4 6

mat[1,2]
[1] 3

mat[2,]
[1] 2 4 6

mat[,2:3]
```

```
     [,1] [,2]
[1,] 3 5
[2,] 4 6
```

Next, consider the same matrix, but give names to the rows and columns. The rows will get the names "a"and "b"and the columns will be named "A", "B"and "C".
The entire R session could look like this:

```
mat<-matrix(1:6,ncol=3)
dimnames(mat)<-list(c("a","b"),c("A","B","C"))
mat
  A B C
a 1 3 5
b 2 4 6

mat["b",c("B","C")]
B C
4 6
```

## 8.3 The apply command

The apply command...
apply(mat,2,sum) – applies the sum function within each column
apply(mat,1,mean) – computes the mean within each row

## 8.4 The tapply command

Commonly one has a data vector and another vector of the same length giving categories for the measurements. In this case one often wants to compute the mean or variance (or median etc) within each category. To do this we use the tapply command in R.

### 8.4.1 Examples

**Example 8.3.** `z<-c(5,7,2,9,3,4,8)`

```
i<-c("m","f","m","m","f","m","f")
```

A. Find the sum within each group

```
tapply(z,i,sum)
 f  m
18 20
```

B.Find the sample sizes

```
tapply(z,i,length)
f m
3 4
```

C.Store outputs and use names

```
n<-tapply(z,i,length)
n
f m
3 4
n["m"]
m
4
```

## 8.5 Logical indexing

A logical vector consists of *TRUE* (1) or *FALSE* (0) values. These can be used to index vectors or matrices.

### 8.5.1 Examples

**Example 8.4.** `i<-c("m","f","m","m","f","m","f")`

```
z<-c(5,7,2,9,3,4,8)

i=="m"
[1] TRUE FALSE TRUE TRUE FALSE TRUE FALSE

z[i=="m"]
[1] 5 2 9 4

z[c(T,F,T,T,F,T,F)]
[1] 5 2 9 4
```

## 8.6 Lists, indexing lists

A list is a collection of objects. Thus, data frames are lists.

### 8.6.1 Examples

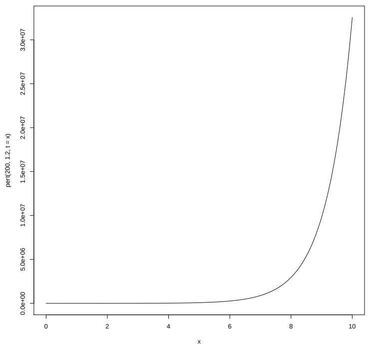
**Example 8.5.** `x<-list(y=2,z=c(2,3),w=c("a","b","c"))`

```
x[["z"]]
[1] 2 3
names(x)
[1] "y" "z" "w"
x["w"]
$w
[1] "a" "b" "c"
x$w
[1] "a" "b" "c"
```

# 9 Functions of functions and the exponential function

## 9.1 Exponential growth and decline

Exponential growth is typically expressed as:
$y(t) = Ae^{kt}$



Exponential growth curve

### 9.1.1 Details

**Definition 9.1. Exponential growth** is the rate of population increase across time when a population is devoid of limiting factors (i.e. competition, resources, etc.) and experiences a constant growth rate.

Exponential growth is typically expressed as:
$y(t) = Ae^{kt}$
where
$A$ (sometimes denoted $P$)=initial population size
$k$= growth rate
$t$ =number of time intervals

*Note 9.1.* Note that exponential growth occurs when $k > 0$ and exponential decline occurs when $k < 0$.

## 9.2 The exponential function

An exponential function is a function with the form: $f(x) = b^x$

### 9.2.1 Details

For the exponential function $f(x) = b^x$, $x$ is a positive integer and $b$ is a fixed positive real number. The equation can be rewritten as:

$$f(x) = b^x = b \cdot b \cdot b...b$$

.
When the exponential function is written as $f(x) = e^x$ then, it has a growth rate at time $x$ equivalent to the value of $e^x$ for the function at $x$.

## 9.3 Properties of the exponential function

Recall that the methods of the basic arithmetic implies that:

$$e^{a+b} = e^a e^b$$

for any real numbers $a$ and $b$.

## 9.4 Functions of functions

### 9.4.1 Details

Consider two functions, $f$ and $g$, each defined for some set of real numbers. Where $x$ can be solved in function $f$ using $Y = f(x)$ when $g(Y)$ exists for all such resulting $Y$. If $Y = f(x)$ and $g(Y)$ exist then we can compute $g(f(x))$ for any $x$.
If
$f(x) = x^2$ and
$g(y) = e^y$ then
$g(f(x)) = e^{f(x)} = e^{x^2}$

If we call the resulting function $h$;
$h(x) = g(f(x))$
Then $h$ is commonly written as
$h = g \circ f$

### 9.4.2 Examples

**Example 9.1.** If
$g(x) = 3 + 2x$ and
$f(x) = 5x^2$
Then
$g(f(x)) = 3 + 2f(x)$
$g(f(x)) = 3 + 10x^2$

$f(g(x)) = 5(g(x))^2$
$f(g(x)) = 5(3 + 2x)^2$
$f(g(x)) = 45 + 60x + 20x^2$

## 9.5 Storing and using R code

As R code gets more complex (more lines) it is usually stored in files. Functions are typically stored in separate files.

### 9.5.1 Examples

**Example 9.2.** Save the following file (test.r):

```
x=4
y=8
cat("x+y␣is", x+y, "\n")$
```

To read the file use:

```
source("test.r")
\end{lstllisting}
and the outcome of the equation is displayed in R
\end{xmpl}

%% Slide http://tutor-web.net/math/math612.1/lecture190/slide60
\subsection{Storing and calling functions in R}
\fbox{
\begin{minipage}{0.97\textwidth}
To save a function in a separate file use a command of the form "
   function.r".
\end{minipage}
}
\subsubsection{Examples}
\begin{xmpl}

\begin{lstlisting}
f<-function(x) {
    return (exp(sum(x)))
    }
```

can be stored in a file function.r and subsequently read using the source command.

# 10 Inverse functions and the logarithm

## 10.1 Inverse Function

If $f$ is a function, then the function $g$ is the inverse function of $f$ if

$$g(f(x)) = x$$

for all $x$ in which $f(x)$ can be calculated

### 10.1.1 Details

The inverse of a function $f$ is denoted by $f^{-1}$, i.e.

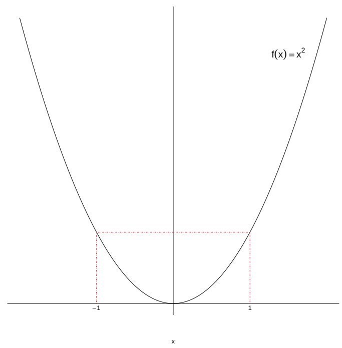
$$f^{-1}(f(x)) = x$$

### 10.1.2 Examples

**Example 10.1.** If $f(x) = x^2$ for $x < 0$ then the function $g$, defined as $g(y) = \sqrt{y}$ for $y > 0$, is not the inverse of $f$ since $g(f(x)) = \sqrt{x^2} = |x| = -x$ for $x < 0$.

## 10.2 When the inverse exists: The domain question

Inverses do not always exist. For an inverse of $f$ to exist, $f$ must be one-to-one, i.e. for each $x$, $f(x)$ must be unique.




The function $f(x) = x^2$ does not have an inverse since f(x)=1 has two possible solutions -1 and 1.

### 10.2.1 Examples

**Example 10.2.** $f(x) = x^2$ does not have an inverse since $f(x) = 1$ has two possible solutions -1 and 1.

*Note 10.1.* Note that iff $f$ is a function, then the function $g$ is the inverse function of $f$, if $g(f(x)) = x$ for all calculated values of $x$ in $f(x)$.

The inverse function of $f$ is denoted by $f^{-1}$, i.e. $f^{-1}(f(x)) = x$.

**Example 10.3.** What is the inverse function, $f^{-1}$, of $f$ if $f(x) = 5 + 4x$.

The simplest approach is to write $y = f(x)$ and solve for $x$:
With
$$f(x) = 5 + 4x$$
we write
$$y = 5 + 4x$$
which we can now rewrite as
$$y - 5 = 4x$$
and this implies
$$\frac{y-5}{4} = x$$
And there we have it, very simple:
$$f^{-1}(f(x)) = \frac{y-5}{4}$$

## 10.3 The base 10 logarithm

When $x$ is a positive real number in $x = 10^y$, $y$ is referred to as the base 10 logarithm of x and is written as:

$$y = \log_{10}(x)$$

or

$$y = \log(x)$$

### 10.3.1 Details

If $\log(x) = a$ and $\log(y) = b$, then $x = 10^a$ and $y = 10^b$, and

$$x \cdot y = 10^a \cdot 10^b = 10^{a+b}$$

so that

$$\log(xy) = a + b$$

### 10.3.2 Examples

**Example 10.4.**

$$\begin{aligned} log(100) &= 2 \\ log(1000) &= 3 \end{aligned}$$

**Example 10.5.** If

$$\log(2) \approx 0.3$$

then

$$10^y = 2$$

*Note 10.2.* Note that

$$2^{10} = 1024 \approx 1000 = 10^3$$

therefore

$$2 \approx 10^{3/10}$$

so

$$\log(2) \approx 0.3$$

## 10.4 The natural logarithm

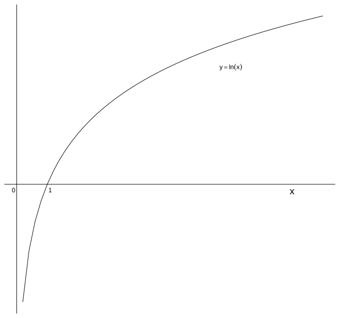
A logarithm with $e$ as a base is referred to as the natural logarithm and is denoted as $ln$ :

$$y = ln(x)$$

if

$$x = e^y = exp(y)$$

Note that $ln$ is the inverse of $exp$.




The curve depicts the fuction $y = \ln(x)$ and shows that $ln$ is the inverse of $exp$. Note that $\ln(1) = 0$ and when $y = 0$ then $e^0 = 1$.

## 10.5 Properties of logarithm(s)

Logarithms transform multiplicative models into additive models, i.e.

$$\ln(a \cdot b) = \ln a + \ln b$$

### 10.5.1 Details

This implies that any statistical model, which is multiplicative becomes additive on a log scale, e.g.

$$y = a \cdot w^b \cdot x^c$$
$$\ln y = (\ln a) + \ln(w^b) + \ln(x^c)$$

Next, note that

$$\begin{aligned} \ln(x^2) &= \ln(x \cdot x) \\ &= \ln x + \ln x \\ &= 2 \cdot lnx \end{aligned}$$

and similarly $\ln(x^n) = n \cdot \ln x$ for any integer n.
In general $\ln(x^c) = c \cdot \ln x$ for any real number c (for x>0).
Thus the multiplicative model (from above)

$$y = a \cdot w^b \cdot x^c$$

becomes

$$y = (\ln a) + b \cdot \ln w + c \cdot \ln x$$

which is a linear model with parameters $(\ln a)$, $b$ and $c$.
In addition, the log-transform is often variance-stabilizing.

## 10.6 The exponential function and the logarithm

The exponential function and the logarithms are inverses of each other

$$x = e^y \Leftrightarrow y = \ln x$$

### 10.6.1 Details

*Note 10.3.* Note the properties:

$$\ln(x \cdot y) = \ln(x) + \ln(y)$$

and

$$e^a \cdot e^b = e^{a+b}$$

### 10.6.2 Examples

**Example 10.6.** Solve the equation

$$10e^{1/3x} + 3 = 24$$

for $x$.
First, get the 3 out of the way.

$$10e^{1/3x} = 21$$

Then the 10.

$$e^{1/3x} = 2.1$$

Next, we can take the natural log of 2.1. Since *ln* is an inverse function of *e* this would result in

$$\frac{1}{3}x = \ln(2.1)$$

This yields

$$x = \ln(2.1) \cdot 3$$

which is

$$\approx 2.23$$