

Using Excel notation for mathematical formulae and equations ... and its advantages over conventional notation

Michael Wood (MichaelWoodSLG@gmail.com)
August 2015

Mathematics is there to make difficult things easy. Eugenia Cheng (Cakes, custard and category theory)

Perhaps the best known example of the elegance, simplicity and power of mathematics is Einstein's equation

$$E=mc^2$$

There are five symbols here but it actually incorporates seven concepts: E is the energy

equivalent in joules (concept 1) of a mass of m kilograms (concept 2), c is the speed of light in a vacuum in meters per second per second (concept 3), and 2 is the number two (concept 4). There is also the = sign indicating the numerical value of both sides is the same (concept 5), and there are also two other concepts which have *no* symbols to represent them but are implied by the position of the symbols which are there: the juxtaposition of m and c means that these two are to be multiplied (concept 6), and placing the 2 above and to the right of c implies taking a power of c (concept 7). Seven concepts, five symbols -- this is a really economical notation, and in this sense it is beautifully simple.

However, its simplicity is only useful if you know what it means and can use it for whatever purposes you might have in mind, such as for example, working out the energy equivalent of my 80Kg mass. This implies two principles for redesigning the notation for non-mathematicians unfamiliar with the conventions of algebraic notation. First, each concept should be represented by an explicit symbol which should be as easy to understand as possible. If someone does not understand that m and c are multiplied together they are likely to be stuck because there is no symbol whose meaning they can search for to find out what it means. Second, it would be nice if the whole equation, or the right hand side of it, could be pasted into a computer package to work out what it comes to in a particular scenario.

Applying these two principles to Einstein's equation I came up with

$$\text{energy} = \text{mass} * \text{soliv}^2$$

Simple Learning home page:
<http://woodm.myweb.port.ac.uk/SL/simplelearning.htm>

The suggestions made here are used in these articles:

Compound (exponential) growth and dedine:
<http://woodm.myweb.port.ac.uk/SL/compoundgrowth.pdf>

Sampling for possibilities:
<http://woodm.myweb.port.ac.uk/SampPoss.pdf> (see the formula at the end of the article on page 17).

where *soliv* is an acronym standing the speed of light in vacuum (about $3 \cdot 10^8$ meters per second), * is the multiplication sign, and ^ means to the power of. These are the symbols used in Excel formulae.

Less elegant, and longer, but this version has its advantages:

1. The names of the quantities are longer - *mass* instead of *m*, etc - making it easier to remember what the symbols stand for. I went for the acronym, *soliv*, because the full phrase is obviously too long to fit conveniently into an equation. I've restricted myself to lower case letters only for the reason explained under (3) below.

Computer programmers have long favoured long and hopefully meaningful variable names, but mathematical symbolism has not followed the idea for, I think, two reasons. First, the fact that computer languages are continually being invented means that change is an accepted part of life. Mathematics is different; the habit of reinvention does not exist, instead there is the slavish clinging to tradition which is customary in most formal academic disciplines (link to come). There is also a more practical reason: because juxtaposition of two symbols means that they are multiplied, *mass* would be interpreted as *m* times *a* times *s* times *s*. Using the symbol * for multiplication avoids this problem, but at the cost of making formulae involving a lot of multiplication less elegant to a mathematician's eye.

2. Formulae or equations can be pasted straight into cells in an Excel worksheet, provided that we tell Excel what the variables are. If we replace *mass* by 80, and *soliv* by $3 \cdot 10^8$, to get the formula

$$= 80 * (3 * 10^8)^2$$

and then paste this (including = which tells Excel to expect a formula) into a cell, the number returned in the cell is 7.2E18 or 7 200 000 000 000 000 000, or 7.2 million million million. This corresponds to the number of joules generated if I were to be completely converted to energy. (Excel allows you to reformat numbers - right click on a cell and click format - and define variables like *mass* - check Help on Define name for how to do this.)

3. There is a further advantage in using Excel notation: it is possible to make use of the large number of built in functions. Many of these are standard mathematical functions (e.g. trigonometry functions like sine - SIN in Excel), but some aren't (e.g. the financial and date & time functions, and logical functions like IF). Excel functions are denoted by capital (upper case) letters, which is why it's important to lower case for the variable names. For some examples see <http://woodm.myweb.port.ac.uk/SL/compoundgrowth.pdf>.

However, against these advantages, the traditional notation is wonderfully economical, and the variety of symbols used, and the fraction line which gives the beginning of a second dimension to the notation, mean that it is possible to express complicated formulae in such a way that it is easy for someone familiar with the notation to see how they work. For example the probability density function of the normal distribution is usually written like this (this is copied from Wikipedia):

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

This is the function you need to plot to draw the so called normal distribution in statistics (link to come). The formula is based on some fairly advanced mathematics.

The Excel version of this formula is

$$= \text{EXP}(-((x-\text{mean})^2)/(2*\text{sd}^2))/\text{SQRT}(2*\text{PI}()*\text{sd}^2)$$

Where *mean* in the second formula is μ in the first, *sd* is σ . PI() is the built in Excel function for π (the circumference of a circle divided by its diameter which comes to 3.14159...) and EXP is the built in function for the exponential function. If, for example, $x=11$, $\text{mean}=10$ and $\text{sd}=2$, both formulae give 0.176 .

To a trained mathematician, the first version will look better; to anyone else the second version is probably preferable. But there is a problem with the brackets in the Excel version. The open bracket after EXP obviously has another (dose) bracket to tell Excel that the EXP function is worked out from the bit between the brackets, but it may not be immediately obvious that the close bracket in question is the one directly before the division sign (/). This makes it difficult to understand how the formula "works".

You can improve the Excel formula a bit by inserting spaces like this:

$$=\text{EXP} (-(x-\text{mean})^2) / (2*\text{sd}^2)) / \text{SQRT} (2*\text{PI}()*\text{sd}^2)$$

Excel can read this one (provided you tell Excel what x, mean and sd are using Formulas -- Name manager).

To make it even clearer you could use different font sizes:

$$=\text{EXP}\left(-\left(x-\text{mean}\right)^2\right) / \left(2*\text{sd}^2\right) / \text{SQRT}\left(2*\text{PI}()\right)*\text{sd}^2$$

The EXP brackets are now big ones, so you can see at a glance which they are. You will be able to paste this formula into Excel without the = sign, but when you insert the = sign to tell Excel to interpret it as a formula, Excel will insist on removing the different font sizes.

Excel has a final trick up its sleeve. Because the normal distribution function is widely used there is a built in function for it:

$$=\text{NORMDIST}(x,\text{mean},\text{sd},\text{FALSE})$$

So with $x=11$, $\text{mean}=10$ and $\text{sd}=2$, this fomula will also return 0.176 . (Check the Help in Excel to see exactly how this works, and what the FALSE indicates.)

Mathematical notation is, of course, a much broader and more powerful language than these crude suggestions might imply. For example, the symbol \int represents the process of integration which results in another function, not a number. But I do think it is worth trying to adapt notation so that it is as useful and understandable as possible from the perspective of its audience.