

Scientific Data Analysis

An introduction to the principles of scientific measurement and error analysis

By David W. Knight

Version 1.01. 12th March 2010.

© D. W. Knight, 2010.

The most recent version of this document can be obtained from the author's website:

<http://www.g3ynh.info/>

Table of Contents

Introduction.....	1
1. Errors and Uncertainties.....	2
2. Significant figures and rounding.....	5
3. Reading digital meters.....	6
4. Instrumental defects.....	7
5. Calibration of instruments.....	8
6. Propagation of Uncertainties.....	8
7. Combination of uncorrelated measurements.....	10
8. Weighted average.....	11
9. Correlation.....	12
10. Linear Regression.....	14
11. Weighted Linear Regression.....	19
12. The reduced chi-squared test.....	22

Introduction

Winston Churchill famously graded misleading information according to the three categories: "Lies, damned lies, and statistics". The quip is shocking for its apparent rejection of scientific principles; but it is also, we must admit, fair comment on the background to many a dubious conclusion. So what is it about statistics, those numbers which purport to supply facts, which gives them the ability to supply nonsense and political spin with equal facility? The answer lies in the two hidden attributes of any bare statistic, *bias* and *uncertainty*, of which politicians rarely speak, but to which scientists must always pay attention if they are to avoid fooling themselves.

Here we concern ourselves with the business of extracting scientific information from physical measurements; with the emphasis on getting the best out of the measuring equipment, and then working out how good the resulting information is. This is somewhat harder than the comfortable business of recording instrument readings and then believing them; but there is a prize which should please both amateur and professional experimenters alike: the almost magical ability to obtain accurate information using commonplace (and relatively inexpensive) laboratory equipment.

1. Errors and Uncertainties

The outcome of a scientific measurement is called a **random variable**. We cannot know in advance what value a random variable will take on, we can only know that it will probably fall within a certain range.

The **error** in a random variable is defined as the difference between an observed or deduced value and the true value. We do not usually know what the 'true' value is. Experimentally we can never obtain it, we can only converge upon it in a statistical sense, i.e., by making repeated measurements and averaging the results.

We can divide sources of error into 3 groups:

- **Illegitimate errors:** i.e., mistakes in measurement or calculation.
- **Systematic errors (bias):** i.e., poor experimental design or method, failure to calibrate equipment.
- **Random errors:** i.e., noise and fundamental uncertainty.

Here we use the term "noise" in the statistical or scientific sense, to mean scatter or unpredictability in the observed or instantaneous value of a random variable. The audible hissing noise in the output of a radio receiver is due to the inherent unpredictability of the output power, and so the vernacular and the statistical meanings are perfectly analogous.

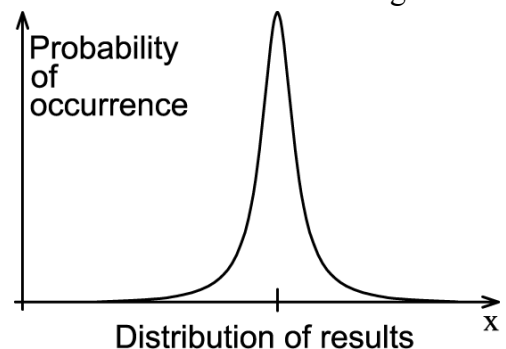
An experimental result is **accurate** if all illegitimate and systematic errors have been eliminated. An experimental result is **precise** if noise and random error have been minimised.

- **Accuracy** = nearness to the truth.
- **Precision** = narrowness of scatter in repeated measurements.

A measurement can be *precise but not accurate* if the scatter of results is small but there is a systematic error in the method. A measurement can be *accurate but not precise* if it is free from systematic errors but the data are noisy (Note that the Latin word "data" is plural. The singular of "data" is "datum").

Precision:

The value of a random variable is meaningless without some measure of its precision or *uncertainty*. One possible way to determine the uncertainty in an experimental result is to repeat the measurement a large number of times. This will only work if the **resolution** of the measuring device is considerably greater than the uncertainty of the measurement; e.g., it doesn't matter how many times you connect a digital multimeter to a resistor, it will always give the same number (provided that the temperature doesn't change). The resolution of a digital measuring instrument is usually the same as its precision (i.e., provided that the least significant digit remains stable). In many situations however, the noise in the measurement is clearly apparent to the experimenter, in which case it is possible to determine a **distribution** of results.



The complete set of all possible measurements of a random variable is called the **parent population**. The parent population may be finite or infinite. An example of a finite parent population is the permeability of type K transformer cores in last week's production batch. An example of an infinite parent population is the noise voltage of an RF amplifier. In most cases we do not have access to the parent population, or it is impractical or too expensive to measure the whole population. We must therefore take sample measurements and work with a smaller **sample population**.

Sample mean:

For measurements with randomly distributed errors, the best estimate for the true value of the variable is the average of the sample population. This is usually defined as the mean value:

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i \quad \text{Sample mean}$$

The equation above implies the convention that we have made a set of measurements and have numbered them from 1 to n. The subscript i is used to refer to the measurements generically, i.e., x_i is an individual measurement with a unique identifier i which has a value between 1 and n. The symbol Σ (capital Sigma) indicates that a summation is to be performed on objects of the type indicated to its right; with the range over which the summation is to be carried out written below and above. The mean of the sample population (pronounced "x bar") is thus the sum of all the individual measurements divided by the total number of measurements.

Note that in some bodies of work, the subscript i may already have been used to indicate (say) parameters associated with a current sampling network. In that case, as a precaution against ambiguity, another letter can be used as the sample number when carrying out data analysis. Strictly it should not be necessary to take such precautions, because the i below the summations symbol relates only to the subscripts on the variables in the summation itself and has no global significance, but we can just as well write:

$$\bar{x} = \frac{1}{n} \sum_{k=1}^n x_k \quad \text{Sample mean}$$

without changing the meaning of the statement in any way.

Parent mean:

The mean of the parent population is defined as the sample mean in the limit that the number of samples goes to infinity (i.e., encompasses everything).

$$\mu = \lim_{n \rightarrow \infty} \left[\frac{1}{n} \sum_{i=1}^n x_i \right]$$

The parent mean is usually given the symbol μ ("mu"), which is italicised here in an attempt to distinguish it from magnetic permeability.

Deviation:

All data are meaningless in the absence of some implied or actual measure of uncertainty. We may guess that real-estate agents are likely to get the dimensions of rooms right to within an inch or so; but for scientific and engineering purposes, we require some kind of rigorously defined interval which can be carried through into other calculations. The most widely used definition is that involving the mean of the squares of the deviations of the individual measurements. This gives rise to a quantity known as the **variance** of the data, which is given the symbol σ^2 ("sigma squared"):

$$\sigma^2 = \lim_{n \rightarrow \infty} \left[\frac{1}{n} \sum_{i=1}^n (x_i - \mu)^2 \right] \quad \text{Variance}$$

Notice that the variance is defined in terms of the parent mean μ , because we can only know it

absolutely when we have included the deviation of every member of the population. Note also the reason for averaging the *squares* of the deviations, the point being that the deviations themselves can be positive or negative, but the square is always positive. The average of the deviations themselves (unsquared) is by definition zero.

The problem with the rigorous definition of variance, is that we do not usually know the parent mean. All we have is an estimate of it in the form of the sample mean \bar{x} . We must therefore make do with an estimated variance based on the assumption that $\bar{x} \approx \mu$. The most naive attempt to find an estimate of variance would be:

$$\sigma^2 = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2 \quad (\text{only true for large } n)$$

The problem with this definition is that if the number of samples is only 1 it tells us that $\sigma^2=0$, i.e., it implies that there is no uncertainty if only one measurement is made. The reality is that we can have no confidence whatsoever in a single uncorroborated measurement, and so this definition is false. The flaw, expressed in mathematical language, is that the equation has an incorrect *boundary condition*, i.e., it does the wrong thing when pushed to one of the limits of its range of applicability. There only way in which we can repair it is by changing it to:

$$\sigma^2 = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2 \quad \begin{array}{l} \text{Estimated variance} \\ \text{(true for any } n\text{)}. \end{array}$$

Dividing by $n-1$ instead of n has no effect if n is large, but when $n \rightarrow 1$, $\sigma^2 \rightarrow 0/0$, i.e., the corrected expression tells us that uncertainty is *undefined* if we only make one measurement. The quantity $n-1$ is the number of **degrees of freedom** in the data, i.e., it is the number of independent data or measurements at our disposal. The reason why we must divide the square error sum by $n-1$ instead of n is that we have lost one degree of freedom by using the data themselves as the source of the estimate of the population mean.

The quantity σ , which is defined as:

$$\sigma = +\sqrt{\sigma^2}$$

is known as the **standard deviation** of the data, and is the RMS average deviation. If it is derived from a variance using \bar{x} instead of μ it should strictly be referred to as the **estimated standard deviation** or ESD. In the event that the errors in the data are distributed according to a **normal** or **Gaussian distribution** (actually first described by de Moivre in 1733¹), the region $\pm \sigma$ is the **68% confidence interval**, i.e., for normally distributed data there is a 68% probability that the population mean will lie somewhere between $\bar{x} - \sigma$ and $\bar{x} + \sigma$. The word "normal" in a mathematical sense means 'tending towards a norm (or mean)', but it is also 'normal' in a vernacular sense for data to be normally distributed.

It is good practice to state all measurements (or at least those with normally distributed errors) in the form $\bar{x} \pm \sigma$. There are two commonly used notations for doing so, one is to put σ after a \pm symbol, and the other is to express σ as a change in the right-most digit of a number and place it in brackets after the number. Hence:

1.231(2) is the same as 1.231 ± 0.002

1.4097(19) is the same as 1.4097 ± 0.0019

¹ Introduction to Probability and Statistics for Engineers and Scientists. Sheldon M Ross. Wiley 1987. Normal Random Variables. Ch. 3, Sec. 5, p96.

Note that a standard deviation usually has the same units as the quantity to which it relates, hence it is appropriate to write:

$$I = 632 \pm 3 \text{ mA} \quad [\textit{not } I = 632\text{mA} \pm 3].$$

One situation in which the uncertainty is not in the same units as the measurement however, is when it is expressed as a percentage. In this case it is correct to write:

$$I = 632\text{mA} \pm 0.5\% \quad [\textit{but definitely not } I = 632 \pm 0.5\% \text{ mA}].$$

To convert an uncertainty in % back into its native units, divide it by 100 and multiply it by the measurement to which it relates:

$$632\text{mA} \pm 0.5\% = 632 \pm (0.005 \times 632) \text{ mA} = 632 \pm 3\text{mA}$$

Manufacturing tolerances:

Resistors, capacitors, and other manufactured components usually have "brick wall" tolerances associated with their nominal values; i.e., the manufacturer 'guarantees' that the true value will lie within the stated range (provided that other conditions such as operating temperature are also within the stated range). Hence, the manufacturer's tolerance is (with some reservations) notionally a 100% confidence interval, and the associated errors are not normally distributed. The actual distribution of errors then depends on the manufacturing process.

If the manufacturing process is imprecise, components are tested and binned according to the preferred-value range in which the actual measured value lies. In that case, the errors have a flat distribution and the most probable error is 50% of the tolerance. For a normally distributed population on the other hand, the most probable error is σ . Hence, when analysing data which are subject to both measurement errors and component tolerance errors, a reasonable approximation to the standard deviation of a component value is given by taking half of the manufacturer's tolerance.

Sometimes a manufacturing process has considerably greater precision than the stated component tolerances seem to imply. This is a danger area for those attempting to deduce the most probable error. The problem here is that for some components, particularly rolled capacitors and wirewound resistors, the manufacturer can save materials by always producing components with values which are close to the lower tolerance limit. Hence, if working with stated tolerances, it is always necessary to ask the question "can the manufacturer maximise profits by engineering a difference between the most probable value and the nominal value?" If the answer is "yes", then the nominal value cannot be trusted and independent measurement is the only sensible course of action.

2. Significant figures and rounding

There is usually no point in recording a number with significant digits beyond the point where the standard deviation renders those digits meaningless. We might, for example, perform a calculation which gives the result:

$$x = 5.13782041 \pm 0.00362148$$

Taking the most austere attitude towards rounding, this number should be reported as:

$$x = 5.138 \pm 0.004$$

although, for those who dislike throwing anything away:

$$x = 5.1378 \pm 0.0036$$

is a little less draconian.

The usual rounding rule taught in schools is that, if the digit after the last significant digit to be reported is between 5 and 9, round up; but if it is between 0 and 4, round down. This rule however only works unambiguously if we know several more digits than are to be reported. If, for example, if we wish to round the number 21.351723 to one decimal place, it is obvious that the result should

be 21.4; and if we wish to round the number 21.347816 likewise, the result should be 21.3. A difficulty will only occur if we have to round a number like 21.350000, because then we have no way of deciding whether to go up or down. In that case we should decide randomly (tossing a coin will do the trick), or we can use a quasi-random rule like 'round up if the digit to be rounded is odd' (or vice versa). Fortunately, the cases where such decisions have to be made are very rare, and so the old school rule 'always round up' will not introduce significant bias into the data.

A more serious problem arises when someone or something has already truncated the number we wish to round. Take for example the case where we wish to round the number 21.35 to one decimal place. How do we know whether the last 5 arose from a string like 51723 or from a string like 47816? If we always round up, we will on average add a bias of +0.25 to the least significant digit, which will introduce a systematic error admittedly smaller than, but nevertheless of the order of, the standard deviation. The solution, if *forced* into this position, is once again to round up or down on a random or pseudo-random basis. The solution if using a spreadsheet or writing a computer program, is to prevent the problem from occurring by formatting the output so that numbers are always given to a few more places than are significant.

There is another situation in which it is not a good idea to round a number too drastically. This is when the number is to be used as a seed for the regeneration of a data set or for the prediction of new data. A typical situation is when the data in a table have been fitted to a mathematical function. The parameters determined from the fitting process may have uncertainties which suggest that they should be rounded; but if the numbers are rounded and put back into the function, they will no longer give the best estimates of the original data or good predictions for new data. There is no harm in reporting:

$$x = 5.13782041 \pm 0.0036$$

and it is sometimes useful to retain a long string of computed digits in this way. The only problem is that it is sometimes necessary to argue with editors and reviewers who think that they know better.

The last word on rounding is: only round at the end of a computation. *Never* round intermediate results.

3. Reading digital meters

When using digital measuring instruments, the best we can usually manage is to assume that the designer of the instrument has incorporated statistically reasonable rounding rules into the firmware. This raises the question: 'what do you write down when the last digit of the reading flickers between two numbers?' Take for example, the situation where a reading flickers between 100.1 and 100.2 when measuring a resistance. If the reading is steady at 100.1, then it is reasonable to assume that a number between 100.05000... and 100.14999... is implied. If the reading is steady at 100.2, then a number between 100.15000... and 100.24999... is implied. Hence if the instrument cannot decide, then logically the number is exactly 100.15. There is no convention for recording the extra phantom digit obtained in this way, and indeed most workers round up or down arbitrarily. It is fallacious to argue that arbitrary rounding does not matter however, because many digital instruments are designed in such a way that the uncertainty is only in the last digit. This ensures that arbitrary rounding will introduce bias of the order of the standard deviation. Thus the extra digit should be included in any calculations which use the measurement; and the author's recommendation (in the event that other users of the data will wonder where it came from) is to put it in *italics* or a reduced font size when reporting; i.e, for the above example, the measurement can be reported as 100.15 or 100.15.

4. Instrumental defects

Measuring instruments suffer from some or all of the following defects:

- Random error
- Scale error
- Offset error
- Linearity errors
- Range-switching error
- Monotonicity errors

In electrical measuring instruments, **scale error** is due to the incorrect adjustment of a shunt or series resistor, or some other gain or sensitivity control. A scale error implies that all readings of the instrument are multiplied by some (perhaps unknown) scaling factor.

Offset error occurs when an instrument does not read zero for zero input. Its effect is to introduce a systematic error which is proportionately greater for low readings than it is for high readings. For digital instruments, offset nulling is part of the calibration procedure. For moving-coil meters, the offset may change according to the physical orientation of the instrument, and it is important to check the zero adjustment before making readings.

Linearity errors arise from a variety of causes and are often difficult to quantify. Sometimes the cause of non-linearity is well understood (e.g., diode detectors), and a correction function can be applied to the readings. Often however, non linearity is due to instrumental idiosyncrasies, or failure to care for the instrument correctly. For moving coil instruments, linearity can be affected by steel swarf or stray parts adhering to the magnet and upsetting the uniformity of the magnetic field. Variable reference capacitors are often calibrated by means of knifing-vanes, and kinks in the relationship between dial markings and actual capacitance can occur at the boundaries between vanes.

Range-switching error is due to the change in scale error which can occur on switching an instrument from one sensitivity setting to another. It manifests itself as a jump or discontinuity in the graph of a set of measurements, but can be controlled by careful calibration or application of a range-dependent correction factor to the readings. Note incidentally, that even if an instrument is adjusted so that graphs plotted using different ranges are perfectly contiguous; changing range can still introduce a change in the proportionate (%) standard deviation of a measurement, and statistical bias (unfair weighting) can result if this is not accounted for.

Monotonicity error can occur in digital instruments when a change takes place in one of the more significant (left-most) digits of the binary representation. It manifests itself as the reading or the output going down when it should go up, or vice versa. Monotonicity error limits the effective resolution of A-D and D-A converters, e.g., a converter might have an physical resolution of 16 bits, but it might only be monotonic at the 14 bit level (the last two digits must be rounded off to guarantee a monotonic output).

5. Calibration of instruments

In the user manual provided with a measuring instrument, there will be a section specifying the accuracy of the instrument on its various ranges. Such specifications usually refer to scale errors and are normally expressed in %. At time of calibration, there are two possible interpretations of the specification for a particular instrument range:

- 'If it is within the specified limits it is good enough'.
- 'If it is adjusted exactly to the standard, then there is a good chance that, if it is treated carefully, it will still be within its specified limits at the end of the recommended recalibration interval'.

There are no prizes for guessing which interpretation is correct. Instruments calibrated according to the first philosophy *always* introduce systematic errors. The fact that there will be some uncertainty in the value of the standard, and some error in the reading of the standard, means that an instrument which appears to be working within its tolerance may not be. Exact agreement with the standard maximises the probability that the instrument will perform according to its specification.

6. Propagation of Uncertainties

A common situation when making use of scientific measurements is the need to compute the value of some other quantity from those measurements. To do that, we must use a mathematical formula, and the first question we need to ask is: 'how is the uncertainty in a measured quantity transmitted through a formula into the final result?'. Let us begin by considering a generalised formula with a single variable parameter:

$$y = f(x)$$

where the parameter x has an uncertainty δx (here we will start by using δ instead of σ because σ is always positive by definition, whereas an uncertainty transmitted through a formula can change sign). We want to know the uncertainty of y ; to which, by obvious convention, we will assign the symbol δy . The solution to this problem is given by straightforward logic: The uncertainty of y is the rate of change of y with respect to x multiplied by the uncertainty in x , i.e.:

$$\delta y = \delta x \, dy/dx$$

All we have to do is find the derivative dy/dx and multiply it by δx . So far, so good, but what we have at this stage usually relates to very trivial situations. An example might be that of measuring a current in Amps, and wanting to know it and its standard deviation in milliAmps. This is so easy it barely warrants a thought, but if we state the problem formally we can examine the underlying process:

Let y be the current in mA and x the current in A. Then:

$$y = 1000x$$

$$dy/dx = 1000$$

$$\delta y = 1000\delta x$$

In this case the derivative is positive, so we can substitute standard deviations directly in place of the deltas:

$$\sigma_y = 1000\sigma_x$$

Things get a little more complicated when, for example, powers of x or logarithms are involved, because then the uncertainty which x transmits into y depends on the value of x , but the general principle is the same. Take for example:

$$y = 1/x$$

In this case:

$$dy/dx = -1/x^2$$

hence:

$$\delta y = -\delta x/x^2$$

which means, for this particular relationship, that large values of x transmit less uncertainty into y than small values of x . We can also see from this example why the deltas are not the same as standard deviations; because we have the situation in this case that if δx is positive, then δy is negative. We can force the general technique to work with standard deviations however, either by using the magnitude of the derivative (i.e., by just ignoring the sign), or algebraically by using the RMS value of the transmitted uncertainty, i.e.:

$$\sigma_y = +\sqrt{[(\sigma_x dy/dx)^2]}$$

The point being to make the answer positive by squaring it and the taking the square root.

Now let us consider the general situation, which is frequently encountered when the problem at hand is sufficiently difficult to warrant some programming or the production of a spreadsheet. This is when several measured quantities of different origin must be combined in the calculation of a desired quantity. We can express this situation generically by writing:

$$y = f(p, q, r, s, \dots)$$

where p, q, r, s , etc. are variables with uncertainties $\delta p, \delta q, \delta r, \delta s$, etc.. If we consider each variable in turn, we can find its contribution to the uncertainty in y by taking the *partial derivative* of the function with respect to the variable (i.e., we differentiate the function with all of the other variables held constant). If we assign the symbol δy_p to the contribution which the uncertainty in p makes to the uncertainty in y , then:

$$\delta y_p = \delta p \partial y / \partial p$$

and so on. Thus we can determine all of the partial error contributions from all of the parameters, but we are left with the problem of how to add them together. Our first guess might be that we should simply add them directly, i.e.:

$$\delta y = (\delta p \partial y / \partial p) + (\delta q \partial y / \partial q) + (\delta r \partial y / \partial r) + \dots$$

The failing here is that derivatives can be positive or negative; hence for some formulae and data there may be total cancellation of all errors, and this is physically unreasonable. We might therefore try to remove this flaw by using magnitudes:

$$\delta y = |\delta p \partial y / \partial p| + |\delta q \partial y / \partial q| + |\delta r \partial y / \partial r| + \dots$$

This looks better, but if we translate the statement into ordinary language, we will find that it is still not correct. The statement tells us to compute the uncertainty in y on the assumption that each of the parameters is most likely to be displaced from its parent mean by an amount equal to its uncertainty (which is reasonable), but then assumes that the displacements will always conspire in such a way that their contributions increase the total. This amounts to a worst-case combination of errors, and the probability that it will occur in practice is very small. The reality is that the difference between a variable and its parent mean is just as likely to be negative as it is to be positive. Hence there is just as much chance that the uncertainties of a pair of variables will partially cancel as there is that they will augment. We must deal with this issue by combining the parameter uncertainties in the way which is *most probable*; and indeed, if the uncertainties are standard deviations, we must combine them to produce a result which is also a standard deviation. This requires us to consider the extent to which the parameter uncertainties are *correlated* (i.e., literally, co-related).

If an experiment is properly designed, then by definition, all possible steps have been taken to eliminate systematic errors. This does not necessarily mean that all systematic errors *have been* eliminated, but it does mean that such errors have been eliminated as far as is known. In that case, it is sensible to assume that all of the remaining errors are random. If an error in a parameter is random, then its magnitude or direction cannot be influenced by the error in any other parameter. If the deviations in two or more variables are able to change independently, each without affecting any

other, then they are *uncorrelated*; and, by virtue of their independence they must exist in different dimensions. Hence the uncertainties in any pair of parameters can be considered to be like vectors which move at right angles to each other. Objects having this property of mutual independence are said to be *orthogonal*. If we transmit the uncertainties through a formula and turn them into partial error contributions, the partial errors are still orthogonal, but are now all measured in the same units and can be treated like the components of an actual vector. The the proper way to combine them therefore is to take the *vector sum*, i.e., to calculate the magnitude of the resultant by using Pythagoras' theorem. Hence the correct way to find the sum of a set of uncorrelated error contributions is:

$$\delta y = +\sqrt{[(\delta p \partial y/\partial p)^2 + (\delta q \partial y/\partial q)^2 + (\delta r \partial y/\partial r)^2 + \dots]}$$

Notice that this sum is an RMS value. Therefore it can be used with standard deviations instead of arbitrarily defined uncertainties, and a standard deviation will be obtained as the result:

$$\sigma_y = +\sqrt{[(\sigma_p \partial y/\partial p)^2 + (\sigma_q \partial y/\partial q)^2 + (\sigma_r \partial y/\partial r)^2 + \dots]}$$

A practical example of the use of an error function of this type is given in an accompanying article².

7. Combination of uncorrelated measurements

There is one application of the error analysis procedure just outlined which should be memorised by anyone engaged in the business of making measurements. It is the answer to the question: 'how much improvement in precision can be obtained by repeating a measurement several times and taking the average?' This question of course only relates to measuring devices which are capable of resolving the noise in the quantity being measured, or which in some sense add noise to a level which is greater than the instrumental resolution.

Let us first consider the case where two measurements are made of a quantity x , and the average is taken:

$$x = (x_1/2) + (x_2/2).$$

If both measurements have the same standard deviation σ , and the errors are uncorrelated, then:

$$\sigma_x = \sqrt{[(\sigma \partial x/\partial x_1)^2 + (\sigma \partial x/\partial x_2)^2]}$$

where $\partial x/\partial x_1 = 1/2$ and $\partial x/\partial x_2 = 1/2$, hence:

$$\sigma_x = \sqrt{[(\sigma/2)^2 + (\sigma/2)^2]}$$

which simplifies to:

$$\sigma_x = \sigma/\sqrt{2}$$

Averaging two uncorrelated measurements improves the precision by a factor of $1/\sqrt{2}$.

Now consider the general case when x is the average of n measurements:

$$x = (x_1/n) + (x_2/n) + (x_3/n) + \dots + (x_n/n)$$

The derivatives are all the same, i.e.:

$$\partial x/\partial x_i = 1/n$$

Hence:

$$\sigma_x = \sqrt{[n(\sigma/n)^2]}$$

which simplifies to:

$$\sigma_x = \sigma/\sqrt{n}$$

Averaging n uncorrelated measurements improves the precision by a factor of $1/\sqrt{n}$.

Note that these techniques will only work with measurements which have genuine randomness in

² "AC Theory", D W Knight 2009, section 39. Available from g3ynh.info.

the least significant digits of the result. There is no use in writing down the unchanging reading of a digital multimeter several times (the uncertainty in that case must be taken from the manual or deduced by means of some calibration process), but it is valid and sensible to make the same measurement with two or more different digital multimeters and take the average.

8. Weighted average

Another important problem related to the propagation of uncertainties is: 'how do you combine two measurements having different standard deviations to obtain a meaningful average?' If one measurement has a significantly smaller uncertainty than the other, then it is not fair to the better measurement to take the straight average and bias will be introduced by doing so. On the other hand, it is a waste of data to throw away the inferior measurement. The solution is to take a weighted average, the idea being that if measurement A is n times better than measurement B, we should add n of A to 1 of B and divide by $n+1$. This leaves us with the problem of defining what is meant by 'better'; but it transpires that, for reasons of dimensional consistency which will become apparent later, the *variance* of an observation from a normally distributed population is the proper measure of its goodness. The *weight* of an observation is consequently a number which gets bigger as the variance gets smaller, i.e.:

$$w_i = 1/\sigma_i^2$$

The weight of an observation is a type of frequency, it is the number of times it should be included in an average relative to an observation with *unit weight* (i.e., a weight of 1). It does not matter that it is not an integer. There is no problem in adding 1.7 of A to 3.1 of B and dividing by 4.8. Hence the general form for a weighted average is:

$$x = \frac{(w_1 x_1) + (w_2 x_2) + (w_3 x_3) + \dots + (w_n x_n)}{\sum w_i}$$

where: $w_i = 1/\sigma_{xi}^2$

Here we have used the summation symbol without indicating the range, an informal shorthand which means "sum over the whole range". The weighted average is more compactly written:

$$x = (\sum w_i x_i) / (\sum w_i)$$

Now to work out the uncertainty in x , we note that the derivatives are given by:

$$\partial x / \partial x_i = w_i / (\sum w_i)$$

Hence:

$$\sigma_x = \sqrt{[(x_1 w_1 / \sum w_i)^2 + (x_2 w_2 / \sum w_i)^2 + (x_3 w_3 / \sum w_i)^2 + \dots + (x_n w_n / \sum w_i)^2]}$$

but $1/(\sum w_i)^2$ can be factored out of the square root bracket:

$$\sigma_x = (1/\sum w_i) \sqrt{[(x_1 w_1)^2 + (x_2 w_2)^2 + (x_3 w_3)^2 + \dots + (x_n w_n)^2]}$$

Hence:

$$\sigma_x = [\sqrt{(w_i x_i)^2}] / (\sum w_i)$$

Relative weight:

The absolute weight of an observation is given by the reciprocal of its variance. There are occasions however, when the variances of the observations are not known in advance of the averaging process but there are instrumental or mathematical considerations which permit the *relative* variances to be computed. In such cases we can take an average using relative weights in place of absolute weights and still obtain an unbiased result. This follows by inspection of the expression for the weighted

average (above), where it is apparent that multiplying all of the weights by the same arbitrary constant k (say) will not change the result:

$$x = [\Sigma(k/\sigma_{xi}^2)x_i] / [\Sigma(k/\sigma_{xi}^2)] = [k\Sigma(1/\sigma_{xi}^2)x_i] / [k\Sigma(1/\sigma_{xi}^2)]$$

It is for this reason that we give the weights their own symbol rather than writing them explicitly as reciprocal variances: they do not have to be variances as long as they are proportional to the variances of the observations.

A situation in which a weighted average is required but absolute variance might not be known in advance is that which arises when data have to be processed by means of some mathematical formula before they can be averaged. This statement is actually a description of the generalised process of mathematical modelling or *data reduction*. It is often the case that the input data all have the same uncertainty, but are scaled by some non-linear function. This means that the amount by which the uncertainty is scaled will vary depending on the value of the measurement. If we give the original measurements a relative weight of 1, we can easily work out the new weight after scaling and so take the weighted average. Furthermore, we can then work out the variance of the average, and from it deduce the standard deviation of an observation of unit weight. This subject is examined in more detail in section 11.

9. Correlation

In the event that uncertainty contributions are uncorrelated, we treat them as the orthogonal components of a vector and find the overall uncertainty by taking the magnitude. There are however situations in which uncertainties will be correlated, in which case it is appropriate to add the error contributions directly. A typical example of correlation is when a formula is used to combine two measurements taken from the same meter, and the meter has a scale calibration error, i.e., there is some fixed proportionate error or *scale factor* which affects all readings. The situation we have is:

$$y = f(x_1, x_2)$$

and due to the correlation (if x_1 is above its parent mean then so will be x_2 and vice versa), the error in y is:

$$\delta y = (\delta x_1 \partial y / \partial x_1) + (\delta x_2 \partial y / \partial x_2)$$

The effect of the correlation depends entirely on the derivatives. If both derivatives have the same sign, then the two errors will add. This will mean that the overall error will be greater than if the errors were unrelated, and there is said to be a **positive correlation**. If on the other hand, the derivatives have opposite signs, then the two errors will tend to cancel, and there is said to be **negative correlation**. Hence, if there is positive correlation, the error contributions can be considered to be both pointing in the same direction; if there is negative correlation, the error contributions can be considered to be pointing in opposition; and if there is no correlation then, as we have seen before, the error contributions should be considered to be pointing at right angles.

Now consider the example of a moving-coil meter with a scale calibration error. For each reading there will also be a random error, due to the friction of the bearings and the need to interpolate between the scale markings by eye. Hence every reading will have two error components; the calibration error and the noise. The upshot is that it will not be correct to consider the error vectors to be orthogonal, but neither will it be correct to consider them to be aligned. Consequently they must be considered to be at some intermediate angle. The cosine of this angle is known as the **correlation coefficient**. It is often given the symbol ρ (rho), and is defined in such a way that $\rho = \text{Cos}(180^\circ) = -1$ when the correlation is positive, $\rho = \text{Cos}(0) = +1$ when the correlation is

negative, and of course, $\rho = \cos(90^\circ) = 0$ when there is no correlation. The correlation coefficient is however a very formal way of dealing with the present problem, and is best reserved as a measure of the independence of the parameters obtained by fitting data to a mathematical model. For the purpose of dealing with the correlation between instrumental measurements, it is far simpler to take the view that the noise and the calibration errors can be treated separately. Thus we might have two error components for a single measurement, one being the calibration error which we will call $\delta x_{(cal)}$, and the other being the precision, which we will call $\delta x_{(rnd)}$ (random). If we just take a single measurement from the meter, then, since the random error and the calibration error are not correlated, the standard deviation will be:

$$\sigma_x = \sqrt{[\delta x_{(rnd)}^2 + \delta x_{(cal)}^2]}$$

If we use two measurements from the meter however (x_1, x_2), and put them into a formula to obtain a quantity y , then:

$$\sigma_y = \sqrt{(\delta x_{1(rnd)} \partial y / \partial x_1)^2 + (\delta x_{2(rnd)} \partial y / \partial x_2)^2 + \{ (\delta x_{1(cal)} \partial y / \partial x_1) + (\delta x_{2(cal)} \partial y / \partial x_2) \}^2}$$

In this case, we have added the calibration error components directly; but then, since the total calibration error is not correlated with the random errors, we have squared it and included it in the ordinary RMS error sum.

There is one obvious way to improve the accuracy of the derived quantity, and that is to calibrate the measuring instrument against a reference standard. There are occasions however, when correlation works to the advantage of the experimenter, provided that the measurements are made in a sensible way. Such is the case when there is a need to measure a *ratio*.

Measuring a ratio:

Consider the commonly encountered transfer function:

$$\eta = V_2 / V_1$$

If both voltages are measured using the same meter, and the meter has an unknown scale calibration error s , then for every reading taken from the meter:

$$V_i = s V_i'$$

where V_i' is the reading which would have been given had the meter been calibrated. Hence:

$$V_i = V_i' + (s-1) V_i'$$

and so the calibration error in V_i is:

$$\delta V_{i(cal)} = (s-1) V_i' = V_i (s-1) / s = V_i (1-1/s)$$

The error transferred into the voltage ratio is:

$$\delta \eta = (\delta V_{1(cal)} \partial \eta / \partial V_1) + (\delta V_{2(cal)} \partial \eta / \partial V_2)$$

$$\text{where: } \partial \eta / \partial V_1 = -V_2 / V_1^2 \quad \text{and} \quad \partial \eta / \partial V_2 = 1 / V_1$$

Hence:

$$\delta \eta = [-V_1 (1-1/s) V_2 / V_1^2] + (V_2 (1-1/s) / V_1) = 0$$

Hence, if a ratio is obtained by using two measurements from the same meter, and also using the same range if range switching is possible, then any scale calibration errors are *cancelled*. The random errors in the measurements are, of course, still present, and can be combined in the usual way.

10. Linear Regression

It is often possible to discover a linear relationship between two variables, or to rearrange a formula involving those variables and thereby extract a linear relationship from it. Data from a series of measurements can then be analysed using a model of the general form:

$$y = a + bx$$

This is the equation of a straight-line graph, where b is the slope or gradient of the line, and a is the y -axis intercept point, i.e., the value of y when $x=0$.

Consider an experiment or measurement procedure for which it is known or suspected that a linear relationship exists between two variables x and y . Traditionally, y is the dependent variable and x is the independent variable; and so we will assume that a measurement involves adjusting x to a particular value and recording the corresponding value of y . If a set of such measurements is obtained for a sufficiently broad range of x values, we will at some stage acquire sufficient data to test whether or not the linear relationship hypothesis is reasonable; and if it is reasonable, we will be in a position to obtain usefully accurate values for the slope and the intercept point.

We will assume that measurements are numbered from 1 to n , and that we have at our disposal a list of paired values of x and y , i.e.:

$$(x_1, y_1), (x_2, y_2), (x_3, y_3), \dots, (x_n, y_n)$$

The process of finding a straight-line graph to fit a data set is called **linear regression**. The term 'regression' implies the act of reducing the data to some underlying law; the point being that, if it works, we will in this case replace an entire data set with just two numbers, a and b . To find the procedure, we will start with a special case, which is rarely valid but commonly applied regardless of tiresome details like appropriateness. The special simplifying assumptions are:

- The x_i can be fixed with arbitrary precision.
- The y_i all have the same uncertainty σ .

Obviously, except in rare cases when the x_i are exact by definition (e.g., the number of turns of wire on a magnetic toroid), there will be some uncertainty in the setting of x . It is however acceptable to assume that the x_i are exact provided that:

$$\sigma_{x_i} / x_i \ll \sigma_{y_i} / y_i$$

Note that an absolute uncertainty has the same units as the variable to which it belongs, and that the two ratios compared in the inequality above are therefore dimensionless. The expression says that we may assume the x_i to be exact if their *relative uncertainties* are very small compared to the relative uncertainties of the y_i . With regard to the assumption that the uncertainties in the y_i are all the same: that can sometimes be the case, but we will need to extend the procedure later to cope with situations in which it is not.

We begin by assuming that the linear relationship is true, i.e.:

$$y = a + bx$$

On account of random measurement errors however, a given y_i will not fit the expression exactly. Hence, for the individual measurements, we must write the relationship with an additional error term:

$$y_i = a + bx_i + e_i \quad \dots \quad (10.1)$$

where e_i is known as the **residual error** in y_i .

Now imagine the process of trying to fit the data by guesswork. To do that we would plot the data as points on a graph, and then, using a ruler and pencil, we would try to draw the best straight line through the points. That would involve sliding the ruler up and down and rotating it until the errors appeared to be scattered equally above and below the line all the way along it. To do the same mathematically (and without guesswork), we satisfy two conditions. The first is that the sum of all

the errors should add up to zero (equivalent to sliding the ruler up and down), i.e.:

$$\sum_{i=1}^n e_i = 0$$

The second is that the sum of the squares of the errors should be as small as possible (equivalent to rotating the ruler). The point here is that by making the errors add up to zero, we have found a line which passes through the mid-point of the data set. We still need to rotate the line however, because it may be that all of the residuals on one side of the mid point are positive while all of those on the other side are negative. By squaring the errors, we make them all positive, so the square error sum will only be minimised when the line is optimally embedded in the data. The process of fitting data by minimising the square error sum is called the method of *least squares*.

To minimise the square error sum, we note that it will vary as each parameter is varied, and that a minimum occurs when its rate of change with respect to a parameter is zero. Hence we want to find the two conditions:

$$\partial \Sigma e_i^2 / \partial a = 0 \quad \text{and} \quad \partial \Sigma e_i^2 / \partial b = 0$$

To find the derivatives, we start by writing the square error sum explicitly using equation (10.1):

$$\Sigma e_i^2 = \Sigma (y_i - a - bx_i)^2$$

which expands to:

$$\Sigma e_i^2 = \Sigma (y_i^2 + a^2 + b^2 x_i^2 - 2ay_i - 2bx_i y_i + 2abx_i)$$

Differentiating a summation is just a matter of differentiating the individual terms and adding them together, and so:

$$\partial \Sigma e_i^2 / \partial a = \Sigma (2a - 2y_i + 2bx_i) = 0$$

hence:

$$\Sigma (a - y_i + bx_i) = 0$$

The term Σa just means 'add a to itself n times', and so $\Sigma a = na$, hence:

$$\Sigma y_i = na + b \Sigma x_i \quad \dots \dots \dots (10.2)$$

Also:

$$\partial \Sigma e_i^2 / \partial b = \Sigma (2bx_i^2 - 2x_i y_i + 2ax_i) = 0$$

hence:

$$\Sigma (bx_i^2 - x_i y_i + ax_i) = 0$$

$$\Sigma x_i y_i = a(\Sigma x_i) + b \Sigma x_i^2 \quad \dots \dots \dots (10.3)$$

This gives a solution for b in terms of a:

$$b = [(\Sigma x_i y_i) - a(\Sigma x_i)] / \Sigma x_i^2$$

Substituting this into (10.2) gives:

$$\Sigma y_i = na + \{ [(\Sigma x_i y_i) - a(\Sigma x_i)](\Sigma x_i) / \Sigma x_i^2 \}$$

Multiplying throughout by Σx_i^2 then gives:

$$(\Sigma x_i^2) \Sigma y_i = na(\Sigma x_i^2) + (\Sigma x_i y_i)(\Sigma x_i) - a(\Sigma x_i)^2$$

Which gives a full solution for a by rearrangement:

$$a = \frac{(\Sigma x_i^2)(\Sigma y_i) - (\Sigma x_i y_i) \Sigma x_i}{n(\Sigma x_i^2) - (\Sigma x_i)^2} \quad (10.4)$$

Equation (10.3) also gives a solution for a in terms of b:

$$a = [(\Sigma x_i y_i) - b \Sigma x_i^2] / \Sigma x_i$$

Substituting this into (10.2) gives:

$$\Sigma y_i = n \{ [(\Sigma x_i y_i) - b \Sigma x_i^2] / \Sigma x_i \} + b \Sigma x_i$$

and multiplying throughout by Σx_i gives:

$$(\Sigma x_i) \Sigma y_i = n(\Sigma x_i y_i) - nb(\Sigma x_i^2) + b(\Sigma x_i)^2$$

Which gives a full solution for b by rearrangement:

$$b = \frac{n(\sum x_i y_i) - (\sum x_i) \sum y_i}{n(\sum x_i^2) - (\sum x_i)^2} \quad (10.5)$$

Note that equations (10.4) and (10.5) share the same denominator. Also notice that solutions for a and b can be obtained by calculating four sums: $\sum x_i$, $\sum y_i$, $\sum x_i y_i$, and $\sum x_i^2$; i.e., all it takes is four columns of a spreadsheet and two simple equations involving the totals to find the optimum regression line.

Example: Diode correction function.

When using a diode detector to measure the level of an AC signal, the output voltage of the detector is always slightly lower than that predicted on the basis that the diode is a perfect rectifier.

Consequently, in order to obtain accurate results, it is necessary to correct the data for the diode forward voltage drop. Correction can be achieved by making a set of measurements of diode forward voltage vs. forward current and using them to determine a correction function. If a signal diode with low junction resistance, such as the 1N5711 is used, the diode forward characteristic can be fitted to the following expression:

$$V_f = V_0 + V_1 \text{Log}_e(I_f / I_{ref})$$

This relationship holds good over several decades, and is therefore more than adequate for the correction of readings taken from a single meter range.

In the spreadsheet shown below, a set of diode measurements has been entered in the first two columns, and a linear regression analysis (explained below) has determined the formula:

$$V_f = 0.15834185 + 0.02906031 \text{Log}_e(I_f / [\mu\text{A}]) \quad [\text{Volts}]$$

Note that the units of the current are shown as a divisor in the logarithm bracket, i.e., the reference current is $1\mu\text{A}$ and so every current value put into the formula must be stated in μA . The significance of this notation is that the quantity enclosed by the log bracket must be dimensionless, and a measurement is rendered dimensionless by dividing it by its units. Hence " $I_f / [\mu\text{A}]$ " means "be sure that the number you put here represents a value measured in μA ."

The microAmmeter readings in the first column are nominal (i.e., uncalibrated), because this is a calibration function for a particular diode used in conjunction with a particular meter. The meter should be the same as that used for the acquisition of the experimental data to be corrected, the point being to cancel calibration errors by exploiting negative correlation as discussed in section 9.

A	B	C	D	E	F	G	H	
IN5711 Vf vs If data, fitted to regression line.								1
If / μ A	Vf = y	ln(If) = x	xy	x ²	Vf calc	obs-calc	(obs-calc) ²	2
50	0.2720	3.9120	1.0641	15.3039	0.272026	-0.000026	0.000000001	3
55	0.2750	4.0073	1.1020	16.0587	0.274796	0.000204	0.000000042	4
60	0.2770	4.0943	1.1341	16.7637	0.277325	-0.000325	0.000000105	5
65	0.2795	4.1744	1.1667	17.4255	0.279651	-0.000151	0.000000023	6
70	0.2820	4.2485	1.1981	18.0497	0.281804	0.000196	0.000000038	7
75	0.2840	4.3175	1.2262	18.6407	0.283809	0.000191	0.000000036	8
80	0.2860	4.3820	1.2533	19.2022	0.285685	0.000315	0.000000099	9
85	0.2870	4.4427	1.2750	19.7372	0.287447	-0.000447	0.000000200	10
90	0.2890	4.4998	1.3004	20.2483	0.289108	-0.000108	0.000000012	11
95	0.2910	4.5539	1.3252	20.7378	0.290679	0.000321	0.000000103	12
100	0.2920	4.6052	1.3447	21.2076	0.292170	-0.000170	0.000000029	13
sums →	3.1145	47.2376	13.3898	203.3752		0.000000	0.000000687	14
						esd of fit =	0.000276340	15
							esd	16
Open document spreadsheet					a = V0 =	0.158341851	0.000004092	17
IN5711.ods					b = V1 =	0.029060313	0.000000952	18

In this case, defining the regression line as $y=a+bx$, we identify:

$$y=V_f, \quad a=V_0, \quad b=V_1 \quad \text{and} \quad x=\text{Log}_e(I_f/\mu\text{A})$$

The subscript i has been dropped in the spreadsheet column headings, because it should be obvious that an x or y value associated with a particular measurement is unique to that measurement (i.e., the ' i ' is implicit).

In columns C, D, and E we calculate the quantities x_i , $x_i y_i$, and x_i^2 . The sums in row 14 are then calculated using the spreadsheet Σ function, so that:

$$B14 = \Sigma y_i, \quad C14 = \Sigma x_i, \quad D14 = \Sigma x_i y_i, \quad \text{and} \quad E14 = \Sigma x_i^2.$$

The fitting parameters a and b are then calculated using equations (10.4) and (10.5); where, for the data above, the number of observations $n=11$.

A regression line has now been determined, but it would be unwise to discontinue the analysis at this point. If we use the fitting function to produce a set of calculated y_i values and compare them with the observed values, the exercise becomes self-policing with regard to any mistakes which might have been made in entering the formulae. Note that the formula used in column F must use absolute addressing to pick-up a and b from cells G17 and G18. In Open Office Calc (and other spreadsheets), this is done by placing a $\$$ symbol in front of any row or column designators which must be held constant; i.e., the seed formula for row 3 is:

$$F3 = G\$17 + G\$18 * C3$$

The residuals (observed - calculated) are then obtained in column G using the seed formula:

$$G3 = B3 - F3$$

The sum in cell G14 is just a check to show that the least-squares fitting procedure has indeed established the condition $\Sigma e_i = 0$.

Standard deviation of fit:

The calculated values from a least-squares fit are, in effect, a set of estimates for the parent means (i.e., true values) of the individual observations. Consequently, if all of the observations have the same precision, the residual errors can be treated as though they are all deviations from the same mean. This implies that we can calculate the variance of the fit and relate it to the variance of an observation as if we had made repeated measurements at that particular setting of x . To this end, column I of the spreadsheet shown above is used to calculate the quantity $\sum e_i^2$ (cell H14), and the standard deviation of an observation is estimated (cell H15) as:

$$\sigma_{\text{fit}} = \sqrt{(\sum e_i^2) / (n-2)}$$

The divisor $n-2$ is the number of degrees of freedom in the data, and is two less than the number of observations because the estimate is based on the use of two parameters (a and b) which have been extracted from the data. Once again, we can see that this divisor is correct by considering boundary conditions. If a data set consisting of only two observations is fitted to a function having two variable parameters, then the fit will be exact and the square error sum will be zero. In that case, all knowledge of uncertainties vanishes, and so the number of degrees of freedom must be $n-2$. In general, the number of degrees of freedom for any fitting process is the number of observations included in the fit minus the number of variable parameters.

In the example above, we obtained $\sigma_{\text{fit}}=0.00028$ Volts, and it is interesting to consider how this number came about. The experiment was conducted using a moving-coil microAmmeter, and a digital Voltmeter giving 3 decimal places on its 2V range. By gently tapping the case of the moving-coil meter with a pencil to minimise errors due to bearing friction, it was estimated that the current could be set with a precision of about $\pm 0.2\mu\text{A}$. These current settings were then destined to be converted into logarithms; a process which, although non-linear, has the effect of scaling-down the uncertainties. Hence it was considered reasonable to ignore the setting errors, and since the voltage measurements all have near identical precision, the data were deemed suitable for a simple regression analysis.

The digital voltmeter used had, in its manual, a stated *accuracy* of $\pm 0.5\%$, ± 0.001 for the 2V DC range. The first uncertainty is however primarily a measure of the expected scale error, and the second uncertainty (± 1 in the last digit) is primarily the offset error. These quantities say very little about the actual *precision* of the measurements under the conditions encountered, and σ_{fit} is a measure of precision, not accuracy. We can however deduce the likely precision by noting that the range of voltages involved was very small (0.272 to 0.292), and that the meter is unlikely to have exhibited significant non-linearity over such a range. Therefore, the principal causes of deviation between observed and calculated values are either rounding error, or invalid assumptions about setting error, or failure of the data to conform to the model. Now, the meter reads to three decimal places, which means that there will be a maximum rounding error of ± 0.0005 in any reading. On average however, we should expect a rounding error of about half the maximum, on which basis we should expect $\sigma_{\text{fit}}=0.00025$ if the deviations are due entirely to rounding. We got $\sigma_{\text{fit}}=0.00028$, which tells us that the data do conform to the model within the uncertainty of the method, and that rounding error is the primary cause of deviation.

11. Weighted Linear Regression

The problem with the simple linear regression process described above is that it is only valid when the precision of the dependent variable is the same for all samples. This is often not the case, and it definitely ceases to be the case if the dependent variable must be subjected to some non-linear transformation before it can be equated with y . Take, for example, a situation in which the dependent variable is governed by an exponential relationship:

$$z = k \exp(bx)$$

where 'exp' means 'e to the power of', z is the dependent variable, x is the independent variable, and k and b are parameters. We can linearise this relationship by taking the logarithms of both sides:

$$\text{Log}_e(z) = \text{Log}_e(k) + bx$$

which means that we can determine a regression line $y=a+bx$, with $y=\text{Log}_e(z)$ and $a=\text{Log}_e(k)$. Now let us suppose that all of the measurements of z have the same uncertainty σ_z (which may or may not be known). The uncertainty of a given value of y will be:

$$\sigma_i = \sqrt{[(\sigma_z \, dy/dz)^2]}$$

where, in this case, $dy/dz = 1/z$

Hence, expressed in terms of variances:

$$\sigma_i^2 = \sigma_z^2 / z_i^2$$

The variance of a given y_i will depend strongly on the value of the corresponding z_i . Hence the deviations of the y_i can no longer be treated as though they relate to the same mean. The immediate practical consequence is that simple linear regression analysis will not give the best fitting parameters because it will give equal weight to all observations.

The solution to this problem is related to dimensional analysis. No two quantities can be added unless they are expressed in the same units, and so all terms in the error sum and in the square error sum must be measured in the same units. If the data are all subjected to some non-linear scaling process, then each element of each sum should be subjected to the reciprocal of this scaling process, or the arithmetic will be invalid. The most consistent action therefore is to divide each error by some external measure of relative error which has the same units. The resulting *dimensionless* square error sum is known as the **goodness of fit** (GooF) and is famously referred to as **Chi squared**:

$$\chi^2 = \Sigma(e_i^2/\sigma_i^2)$$

This can also be written:

$$\chi^2 = \Sigma(w_i \, e_i^2)$$

where

$$w_i = 1/\sigma_i^2$$

is called the **weight** of an observation. We have met this quantity before in the discussion of weighted averages, but now we can see that it is a normalisation parameter, i.e., it serves to scale deviations so that they all behave as though they belong to the same mean. Hence χ^2 , in this context, is the **weighted square error sum**.

Now, for the weighted regression analysis, using equation (10.1) we have:

$$\chi^2 = \Sigma w_i (y_i - a - bx_i)^2$$

which expands to:

$$\chi^2 = \Sigma w_i (y_i^2 + a^2 + b^2 x_i^2 - 2ay_i - 2bx_i y_i + 2abx_i)$$

To find the *weighted* least squares fit, we minimise the *weighted* square error sum with respect to each variable parameter, i.e.,

$$\partial\chi^2/\partial a = 0 \quad \text{and} \quad \partial\chi^2/\partial b = 0$$

hence:

$$\partial\chi^2/\partial a = \Sigma w_i (2a - 2y_i + 2bx_i) = 0$$

Which gives:

$$\sum w_i y_i = a(\sum w_i) + b\sum w_i x_i \quad \dots \dots \quad (11.1)$$

and

$$\partial \chi^2 / \partial b = \sum w_i (2bx_i^2 - 2x_i y_i + 2ax_i) = 0$$

Which gives:

$$\sum w_i x_i y_i = a(\sum w_i x_i) + b\sum w_i x_i^2 \quad \dots \dots \quad (11.2)$$

Solution of these two simultaneous equations can be obtained by using (11.2) to substitute for b in (11.1) and rearranging for a, then using (11.2) to substitute for a in (11.1) and rearranging for b. The result is:

$$a = \frac{(\sum w_i x_i^2)(\sum w_i y_i) - (\sum w_i x_i)(\sum w_i x_i y_i)}{(\sum w_i)(\sum w_i x_i^2) - (\sum w_i x_i)^2} \quad (11.3)$$

$$b = \frac{(\sum w_i)(\sum w_i x_i y_i) - (\sum w_i x_i)(\sum w_i y_i)}{(\sum w_i)(\sum w_i x_i^2) - (\sum w_i x_i)^2} \quad (11.4)$$

Once again, the expressions for the parameters have the same denominator. Also notice that if the weights are all set to 1, the expressions revert to the non weighted form (equations 10.4 and 10.5).

The weighted fitting procedure will work just as well with relative weights, and these can be deduced from the rule by which raw data are converted into y values. Also notice that an observation can be excluded from the fit by setting its weight to zero. This is useful when a particular observation is found to have a large residual and an illegitimate error (i.e. a mistake) or a breakdown of the model is suspected. Note that if an observation is excluded in this way, then the number of observations is reduced by one, and the number of degrees of freedom used in calculating the variance of the fit (see below) must be reduced accordingly. A technique which allows data to be excluded and re-included quickly in spreadsheet calculations (to see what happens) is to define a fitting flag for each observation. The fitting flag can have a value of 1 or 0. The theoretical weight of an observation is multiplied by the flag to decide whether it will be included, and the number of observations is determined by taking the sum of the flags. If fitting flags are defined in a computer program; users can be prevented from entering values other than 0 or 1 by including a statement to the effect that:

If $w_i \neq 0$, $w_i = 1$

For weighted linear regression, the input data must be given as a three-column list (x_i , y_i , w_i); although these quantities may of course be computed according to some transformation of the raw data. We then need to compute the quantities:

$\sum w_i$, $\sum w_i x_i$, $\sum w_i y_i$, $\sum w_i x_i y_i$, and $\sum w_i x_i^2$

If using a spreadsheet, $\sum w_i$ is the sum at the bottom of the weights column, and so four extra columns are needed to assemble the quantities needed for equations (11.3) and (11.4).

By incorporating weights, we have effectively scaled all of the sample populations so that they have the same variance. Hence, we can estimate the standard deviation of an observation of unit weight by treating all of the *normalised* deviations as though they are deviations from the same mean. A deviation is normalised by dividing it by its standard deviation (or by a quantity proportional to its standard deviation), i.e.:

$$e_i / \sigma_i = [y_{i(\text{observed})} - y_{i(\text{calculated})}] / \sigma_i$$

Note that this expression says that the deviation of a y value with a large uncertainty should be

taken less seriously than the deviation of a y value with a small uncertainty. Hence:

$$\sigma_{\text{fit}}^2 = (\sum e_i^2 / \sigma_i^2) / \nu$$

where ν (nu) is the number of degrees of freedom, defined as:

$$\nu = n - p$$

n being the number of observations of *finite weight*, and p the number of *variable* parameters. For a linear regression analysis, $\nu = n - 2$.

Now observe that:

$$\sigma_{\text{fit}}^2 = \chi^2 / \nu$$

This quantity is known as "**reduced chi-squared**" and has a very special property. It is the variance of an observation of unit weight. If the weights used in the fit are absolute, i.e., they are derived from the known variances of the observations, then:

$$\chi^2 / \nu \approx 1$$

If the fit is based on absolute weights, then there must be something wrong if the variance of an observation of unit weight is not found to be approximately 1. This relationship is the basis of the *reduced chi-squared test*, which will be discussed in section 12.

If we use relative weights, then the true variance of an observation of unit weight is not expected to be 1. In this case, the quantity:

$$\sigma_{\text{fit}} = \sqrt{[\sum w_i e_i^2] / (n - p)}$$

becomes the estimated standard deviation (ESD) of an observation of unit weight, and the ESD of any particular observation is given by:

$$\sigma_i = \sigma_{\text{fit}} / \sqrt{w_i}$$

Note that the σ_i is the ESD of a y_i , not of a quantity (z_i say) from which y_i was obtained by transformation. To obtain the uncertainty of z_i , it is necessary to apply the reverse transformation, i.e., if:

$$z_i = f(y_i)$$

then

$$\sigma_{z_i} = \sigma_i dz/dy$$

Uncertainties in the fitting parameters:

If the parameters obtained from a least-squares fit are to be used solely for the purpose of recreating the data, then we have little interest in them apart from plugging them into a formula to make use of the values they produce. In many situations however, the parameters themselves can be related to physical quantities, in which case we are also interested in their uncertainties. These uncertainties can be obtained by applying the normal rules relating to the propagation of uncertainties, i.e., we calculate the error contribution from the uncertainty in each observation, and add the squares of these contributions to find the variance of the parameter. This, of course, will only apply if the error contributions are uncorrelated, i.e., the calculated variance will only tell us about the precision of the parameter, it will not tell us its accuracy. Hence, for the parameters from a linear regression analysis:

$$\sigma_a^2 = \sum (\sigma_i \partial a / \partial y_i)^2 \quad \text{and} \quad \sigma_b^2 = \sum (\sigma_i \partial b / \partial y_i)^2$$

but $\sigma_i = \sigma_{\text{fit}} / \sqrt{w_i}$, hence:

$$\sigma_a^2 = \sigma_{\text{fit}}^2 \sum [(\partial a / \partial y_i)^2 / w_i] \quad \text{and} \quad \sigma_b^2 = \sigma_{\text{fit}}^2 \sum [(\partial b / \partial y_i)^2 / w_i]$$

The required derivatives are of course obtained by differentiating equations (11.3) and (11.4), a task facilitated by the fact that the expressions for a and b both have the same denominator, and that y does not appear in the denominator. This means that the denominator will remain constant in the differentiation; and it will already have been evaluated for the purpose of obtaining a and b . We will

therefore give it the symbol D, a local constant, defined as:

$$D = (\sum w_i)(\sum w_i x_i^2) - (\sum w_i x_i)^2 \quad (11.5)$$

Hence:

$$a = [(\sum w_i x_i^2)(\sum w_i y_i) - (\sum w_i x_i)(\sum w_i x_i y_i)] / D$$

$$\partial a / \partial y_i = [(\sum w_i x_i^2) w_i - (\sum w_i x_i) w_i x_i] / D$$

$$= w_i [(\sum w_i x_i^2) - x_i (\sum w_i x_i)] / D$$

$$\sigma_a^2 = \sigma_{\text{fit}}^2 \sum (\partial a / \partial y_i)^2 / w_i$$

$$= (\sigma_{\text{fit}}^2 / D^2) \sum w_i^2 [(\sum w_i x_i^2) - x_i (\sum w_i x_i)]^2 / w_i$$

$$= (\sigma_{\text{fit}}^2 / D^2) \sum w_i [(\sum w_i x_i^2)^2 - 2x_i (\sum w_i x_i)(\sum w_i x_i^2) + x_i^2 (\sum w_i x_i)^2]$$

$$= (\sigma_{\text{fit}}^2 / D^2) [(\sum w_i x_i^2)^2 (\sum w_i) - 2(\sum w_i x_i)^2 (\sum w_i x_i^2) + (\sum w_i x_i^2)(\sum w_i x_i)^2]$$

$$= (\sigma_{\text{fit}}^2 / D^2) (\sum w_i x_i^2) [(\sum w_i x_i^2)(\sum w_i) - (\sum w_i x_i)^2]$$

$$= (\sigma_{\text{fit}}^2 / D^2) (\sum w_i x_i^2) D$$

$$\sigma_a^2 = \sigma_{\text{fit}}^2 (\sum w_i x_i^2) / D \quad (11.6)$$

$$b = [(\sum w_i)(\sum w_i x_i y_i) - (\sum w_i x_i)(\sum w_i y_i)] / D$$

$$\partial b / \partial y_i = [(\sum w_i) w_i x_i - (\sum w_i x_i) w_i] / D$$

$$= w_i [x_i (\sum w_i) - (\sum w_i x_i)] / D$$

$$\sigma_b^2 = \sigma_{\text{fit}}^2 \sum [(\partial b / \partial y_i)^2 / w_i]$$

$$= (\sigma_{\text{fit}}^2 / D^2) \sum w_i^2 [x_i (\sum w_i) - (\sum w_i x_i)]^2 / w_i$$

$$= (\sigma_{\text{fit}}^2 / D^2) \sum w_i [x_i^2 (\sum w_i)^2 - 2x_i (\sum w_i)(\sum w_i x_i) + (\sum w_i x_i)^2]$$

$$= (\sigma_{\text{fit}}^2 / D^2) [(\sum w_i x_i^2)(\sum w_i)^2 - 2(\sum w_i x_i)^2 (\sum w_i) + (\sum w_i x_i)^2 (\sum w_i)]$$

$$= (\sigma_{\text{fit}}^2 / D^2) (\sum w_i) [(\sum w_i x_i^2)(\sum w_i) - (\sum w_i x_i)^2]$$

$$= (\sigma_{\text{fit}}^2 / D^2) (\sum w_i) D$$

$$\sigma_b^2 = \sigma_{\text{fit}}^2 (\sum w_i) / D \quad (11.7)$$

Thus the parameter uncertainties are easily computed from quantities already determined during the fitting process.

If the weights of all of the observations are set to 1 (except for any exclusions) then:

$$\sigma_a^2 = \sigma_{\text{fit}}^2 (\sum x_i^2) / D$$

and

$$\sigma_b^2 = \sigma_{\text{fit}}^2 n / D$$

where n is the number of observations included in the fit.

12. The reduced chi-squared test

In sections 10 and 11 we alluded to the idea that the goodness of a fitting process can be evaluated on the basis of a comparison between the observed and the expected variances of the measurements. This test is known as the reduced χ^2 test and is a very powerful tool for the analysis of scientific data (and also a method for detecting scientific fraud). χ^2 in this case is defined as the weighted square error sum based on absolute weights. 'Reduced χ^2 ' is χ^2 divided by the number of degrees of freedom in the data ($\nu = n - p$) and is given by:

$$\chi^2 / \nu = (1/\nu) \sum_{i=1}^n [(y_{i(\text{obs})} - y_{i(\text{calc})}) / \sigma_i]^2 \quad (12.1)$$

This quantity is, of course, the variance of the fit based on absolute weights, and is therefore the variance of an observation of unit weight. Since an absolute weight is a reciprocal variance, the variation of an observation with a weight of 1 should be 1.

The point about the χ^2 test is that, if the variance of the fit is considerably *larger* than the estimated variance of an observation of unit weight, then the model does not agree with the data and should be modified or discarded. If the variance of the fit is *about the same* as the variance of an observation of unit weight, then the model accounts for the data and is valid within the accuracy of the experimental technique. If the variance of the fit is *considerably smaller* than the variance of an observation of unit weight, then the data may have been faked, or too many variable parameters have been used and the model is fitting noise, or the estimates of the σ_i are too pessimistic. The corresponding ranges of χ^2/ν for the three possible outcomes are accordingly:

$\chi^2/\nu \gg 1$	The model is incorrect, the errors are not random, or the σ_i are underestimated.
$\chi^2/\nu \approx 1$	The data agree with the model.
$\chi^2/\nu \ll 1$	The data are fake, there are too many variable parameters, or the σ_i are overestimated.

Note, that if all observations (y_i) have the same standard deviation σ , then the weights can be factored out of the summation in equation (12.1) above and we have:

$$\chi^2/\nu = (1/\sigma^2)(1/\nu) \sum_{i=1}^n (y_{i(\text{obs})} - y_{i(\text{calc})})^2$$

Hence, in this simplified case, we can define the variance of an observation as estimated from the fit as:

$$\sigma_{\text{fit}}^2 = \sigma^2 \chi^2/\nu = (1/\nu) \sum_{i=1}^n (y_{i(\text{obs})} - y_{i(\text{calc})})^2$$

To obtain χ^2/ν , all we have to do is calculate σ_{fit}^2 from the data and divide it by σ^2 .

In the example at the end of section 10 we obtained $\sigma_{\text{fit}}=0.00027634$ Volts and deduced that if the errors were entirely due to rounding of the readings of the DVM used to acquire the data then a standard deviation of 0.00025 V should be expected. For this example:

$$\chi^2/9 = (0.00027634 / 0.00025)^2 = 1.22$$

The reduced χ^2 test can be applied with considerable sophistication by reference to the χ^2 distribution³, but for a simple qualitative interpretation, it is fair to say that the model gave an excellent fit to the data, and the main effect of least-squares fitting was to drive a line between the DVM rounding errors.

3 **Data Reduction and Error Analysis for the Physical Sciences**, Philip R Bevington. McGraw-Hill, 1969. Library of Congress cat. card # 69-16942. 5-4: χ^2 test of distribution.

Optimisation of multi-parameter empirical fitting functions

By David Knight

Version 1.01, 30th March 2010.

© D. W. Knight, 2010.

Check the author's website to ensure that you have the most recent version of this document:

<http://www.g3ynh.info/> .

Introduction

For some types of data reduction problem, the 'least-squares' fitting criterion is inappropriate. This is the case (for example) when determining an approximate empirical formula to be used in place of a more complicated calculation procedure, where it is better to minimise the peak error (the runout) than it is to minimise the RMS error. There is also a choice over whether to define the fitting residuals as absolute (observed minus calculated) or as proportionate errors; and the average used in determining the proportion can be weighted anywhere between the observed and the calculated value. This article discusses some of the possibilities and their effects in the context of non-linear empirical modelling using downhill search methods.

Downhill search methods

When fitting data to a theoretical model, and presuming that there is no restriction on the number of variable parameters; the usual procedure for determining the optimal parameter values is to configure the problem as a set of simultaneous equations which can be solved numerically by inverting a matrix¹. One reason for using this approach is that it provides intermediate results for the estimation of parameter uncertainties and correlation coefficients. If the model is empirical however, then the parameters have no physical significance and so cannot be used in other calculations; which means that the full statistical analysis is not required. The matrix inversion procedure can still be used of course, and has the advantage of computational efficiency; but for one-off fitting problems, such as developing approximate formulae or converting tabulated data into functions, a great deal of programming can be avoided by using a search method.

For the solution of any fitting problem, it is first necessary to define the fitting criterion. This is usually embodied in a composite error-function which produces a single number called the 'goodness of fit' or "Goof". The Goof is calculated from the fitting residuals in such a way that it takes every residual of finite weight into account and diminishes smoothly (i.e., without discontinuities) as the parameters are shifted towards their optimum values.

When solving by means of a search method, the variable parameters can be considered as the co-ordinates a point on an N-dimensional hypersurface (this is easiest to imagine when N, the number of variable parameters, is 2, in which case the surface is simply a sheet). The Goof provides an extra dimension, the height of the terrain, and the job of the search algorithm is to walk downhill, using information obtained by applying increments to the parameters, until it finds itself at the bottom of a basin. Any minimum it finds is a possible solution; but note that for complicated non-linear systems, it could be just a local minimum, and there might be a better solution elsewhere. Hence, for many problems, there should be an initial hand optimisation phase (to get close to a good solution), before the parameters are passed over to the search routine. It is also sometimes useful to give the system a kick, to see if it can be made to jump out of one minimum and find another; and that can be done by adjusting the initial parameter shifts, or by changing parameters manually

¹ The procedure is described in numerous references. See for example: **Data Reduction and Error Analysis for the Physical Sciences**, Philip R Bevington. McGraw-Hill, 1969. Library of Congress cat. card # 69-16942.

A brief explanation and Fortran code is given in **D W Knight PhD 1985**, Appendix 6. [www.g3ynh.info].

between fitting runs.

The simplest search strategy is that of cycling through the parameters one at a time and adjusting them to minimise the Goof. If the problem is non-linear (i.e., the parameters are not orthogonal), then, after one parameter has been adjusted, all of the others may require re-adjustment, and so it is necessary to repeat the cycle a large number of times until no further significant improvement can be obtained. This is known as a 'grid search', and the process can be painfully slow. The inefficiency can be overcome, to some extent, by using methods which involve shifting all (or, at least, some) of the parameters simultaneously.

There are numerous ways in which searches involving simultaneous parameter shifts can be carried out. One which works well and is popular for its reliability is the Nelder-Mead downhill simplex method^{2 3}. The Nelder-Mead algorithm is not necessarily the most efficient in terms of speed, but it has the ability to expand and contract the size of the parameter perturbations used, enabling it to find narrow wells in the hypersurface. A good reason for recommending it also is that it has been turned into an Open-Document spreadsheet macro function by Robert Weaver⁴. This makes it possible to attack elaborate non-linear optimisation problems using the free Open Office⁵ software package.

Fitting Criteria

Since the only dynamic information available to the fitting algorithm is a single number (the Goof), the type and acceptability of the solution is determined by the way in which this number is calculated. If a least-squares fit is required, the correct choice for the Goof is Chi-squared⁶, i.e:

$$\chi^2 = \sum(w_i e_i^2)$$

where e_i is the i^{th} error or 'residual', and is usually defined as:

$$e_i = y_{i(\text{obs})} - y_{i(\text{calc})}$$

$y_{i(\text{obs})}$ is an observed value of the dependent variable, and $y_{i(\text{calc})}$ is the corresponding value calculated using the fitting function. The weight w_i is strictly analytically proportional to the reciprocal of the variance of the observation (i.e., $w_i = 1/\sigma_i^2$ or some multiple thereof, making χ^2 dimensionless), but may sometimes be set manually to exclude bad data or otherwise gerrymander the result.

For some types of problem however, the least-squares criterion is inappropriate. In particular, when adjusting a function to agree with an exact or reference dataset (such as when developing an approximation formula), the residuals will not be normally distributed, and so the only valid criterion is that which guarantees that the maximum error will lie within strict limits. In other words, we do not want to minimise the RMS error (which is what a least-squares fit does), we want to minimise the maximum error. Also we need to consider the definition of error (should it be absolute; or should it be proportionate so that the overall error can be expressed as a percentage or in parts-per-million, etc.), and perhaps consider the relative authority of the reference data in comparison to the fitting function. To explain these issues, let us start with a generalised composite error function "eta to the 2m":

$$\eta^{2m} = \sum(w_i e_i^{2m})$$

If $m=1$, and the weight and error are as defined above, then this is the same as χ^2 and a least-squares-type fit will result. If m is a much larger positive integer however, then large errors will have vastly greater effect than small ones and the parameter adjustment process will tend to crush the larger errors. The result will then be the desired minimisation of maximum error (minimisation of runout). The power $2m$ is limited only by the need to condition the problem so that the error

2 http://en.wikipedia.org/wiki/Nelder-Mead_method

3 **Numerical Recipes**. W H Press, B P Flannery, S A Teukolsky, W T Vetterling. CUP 1986. ISBN 0521 30811 9. [newer editions exist] Section 10.4: Downhill Simplex method in multidimensions.

4 <http://electronbunker.sasktelwebsite.net/>

5 <http://www.openoffice.org/>

6 See, for example: **Scientific Data Analysis**, D W Knight [www.g3ynh.info].

terms do not cause floating-point underflow or overflow, in which case η^{2m} will become discontinuous and the fitting process could become erratic. Note that the factor of 2 in the power is to ensure an even number, the point being to keep all summation terms positive and prevent error cancellation.

For the individual residuals; if we want to minimise the absolute error, then $y_{i(\text{obs})} - y_{i(\text{calc})}$ should be used as the definition. Otherwise, we might want to minimise the proportionate difference between $y_{i(\text{obs})}$ and $y_{i(\text{calc})}$, and that difference needs to be defined in a way which not only produces the composite error function required by the fitting algorithm, but also produces a rigorous measure of the goodness of the fitting process.

For a least-squares fit, the correct overall measure of goodness is 'reduced chi-squared', i.e.; χ^2/v , where v (Greek "nu") is the number of degrees of freedom in the data, i.e., the number of observations of finite weight minus the number of variable parameters used in the fit. Since v is a constant for a particular fitting run, it makes no difference to a search routine whether χ^2 is divided by v or not (provided that floating-point errors do not occur). For minimum runout problems however, the statistic required is often the worst-case error; and that depends on both the error in the fitting process and the error in the reference data.

In the most general case, the proportionate difference between two numbers is the absolute difference divided by the weighted average. The weighted average is:

$$(W_{(\text{obs})} Y_{(\text{obs})} + W_{(\text{calc})} Y_{(\text{calc})}) / (W_{(\text{obs})} + W_{(\text{calc})})$$

where the weights are proportional to the squares of the respective reciprocal uncertainties (i.e., $w \propto 1/\sigma^2$ for normally distributed errors). Note that when both weights are the same, this reverts to $(y_{(\text{obs})} + y_{(\text{calc})})/2$, which is the familiar straight average. The proportionate difference is thus (in general):

$$e_i = (Y_{i(\text{obs})} - Y_{i(\text{calc})}) (W_{i(\text{obs})} + W_{i(\text{calc})}) / (W_{i(\text{obs})} Y_{i(\text{obs})} + W_{i(\text{calc})} Y_{i(\text{calc})})$$

This might seem complicated, but if we import some knowledge about the data, it will revert to its most familiar form. The most common assumption is that one of the y values is absolutely authoritative, and the other has no authority by comparison. This situation applies (correctly) when we have an exact (or at least extremely accurate) way of calculating say $y_{i(\text{obs})}$. Such would be the case, for example, when trying to develop a simple fitting function to replace a much more complicated calculation procedure. In that case, $w_{i(\text{obs})}=1$ and $w_{i(\text{calc})}=0$, and we get:

$$e_i = (Y_{i(\text{obs})} - Y_{i(\text{calc})}) / Y_{i(\text{obs})}$$

OR

$$e_i = 1 - Y_{i(\text{calc})} / Y_{i(\text{obs})}$$

which can be multiplied by 100 to give the familiar formula for calculating the error in $y_{i(\text{calc})}$ as a percentage, or multiplied by 10^6 to give the error in ppM, etc.. If this type of error definition is put into η^{2m} (with $w_i=1$), then the search process (if successful, and presuming that m is large) will minimise the proportionate runout. Furthermore, the worst case runout can be determined as: $\text{Max}(|e_i|)$, and this is the statistic which will be required by people who intend to use the fitted formula.

Sometimes however, we might want to refer the proportionate error to the fitting function rather than to the supplied data, or we might want some intermediate form using relative weights. Every situation should be considered separately, but one which crops up frequently is that of wishing to fit data which are accurate but not precise to a function which is precise but not accurate. Such is the case, for example, when fitting data which are of theoretical origin (accurate), but which have been rounded to a fixed number of decimal places or exhibit significant floating-point machine error (and are therefore imprecise). The fitting process in this case is a type of smoothing; the proper objective being to find a curve which averages the noise in the data. The end result will be that the smoothing function has a better knowledge of the exact y -values than does the original data, and so the average used in obtaining the proportionate error (should we decide to fit the data in that way) should be weighted to the calculated values. Thus we get:

$$e_i = (y_{i(\text{obs})} - y_{i(\text{calc})}) / y_{i(\text{calc})}$$

i.e.,

$$e_i = (y_{i(\text{obs})} / y_{i(\text{calc})}) - 1$$

and the error function given to the fitting algorithm is:

$$\eta^{2m} = \sum(w_i e_i^{2m})$$

as it was before, but with the definition of e_i altered. For a fitting exercise in which all of the observations are taken to have the same uncertainty (absolute, or proportional, depending on how the residuals are defined), the weight of an observation, w_i , is usually set to 1 initially; but if we find that the dataset contains illegitimate errors (typographic errors, etc), or if the model breaks down in some regions, we can exclude individual observations from the fit by setting the corresponding w_i to zero. The problem of how to deal with data which do not all have the same uncertainty will be examined later.

Progressive weighting

When fitting data to minimise runout; it will often be found, after a first attempt, that the error increases or diminishes as the independent variable (x) is increased. Consider that we have a smooth non-linear fitting function:

$$y_{(\text{calc})} = f(x)$$

which has a reasonably large number of parameters to be adjusted to bring it into near coincidence with a smooth reference function:

$$y_{(\text{obs})} = g(x)$$

After fitting using a runout criterion, a plot of the error curve (e_i vs x_i) will usually show a series of undulations, but the peaks of the undulations will not necessarily be of the same height over the entire range of x . Instead, they will typically either increase or decrease in magnitude with increasing x , and if lines are drawn across the tops of the positive and negative peaks, it will be seen that the fitting function has developed a wedge-shaped pair of error boundaries. Now, if the problem is well-conditioned, it will be possible to trade peak height in one location for peak height in another, and so this is not the optimum fit. The solution is to apply progressive weighting. Recall that we have an error function:

$$\eta^{2m} = \sum(w_i e_i^{2m})$$

where the weights are usually set to 1. All we have to do to apply progressive weighting is use weighting coefficients defined thus:

$$w_i = x_i^u$$

For the first fitting run, we set $u=0$, so that all of the weights are unity. Then if we find that the error increases with x , we set u to a positive number (say 0.5), and run the fitting routine again. The fitting weight will then increase as x increases; the error for large x will decrease slightly, and the error for small x will increase slightly. With a little judicious adjustment of u between fitting runs, the error boundaries can be made parallel, and the maximum runout will then be less than it was before. Similarly, if the error decreases with increasing x , we make u negative; and the converse applies.

Sampling interval

When fitting experimental data to a mathematical model, the size of the dataset is limited by practical considerations. When comparing two smooth functions however, the dataset is potentially infinite, and we must select a finite number of points at which to make comparisons. There are no hard and fast rules on how to do that, but there is a simple test which will tell if the number of samples is insufficient. Always plot a graph of the residuals (e_i vs. x_i). The plotting software will draw straight lines between the points. If straight sections are visible in the peaks of the residual

function, and especially if it obvious that a peak lies between two points, then the maximum runout will not be reported correctly if that is the highest (positive or negative) peak. The solution is to increase the number of data.

Since the size of the dataset can become enormous in some cases, the problem can often be made more manageable by noting that it is not necessary to use a constant sampling interval when generating argument values. Say we have a reference function:

$$y_{(\text{obs})} = g(x)$$

The obvious thing to do is create a spreadsheet column for x and then populate it with x_i values using the 'Fill down' tool. Frequently however, we will find that the residual function has much more detail for low values of x than it has for high values, especially if the fitting function is an asymptotic form. The solution is to create a $\log(x)$ column (with a constant interval between values), and calculate x_i values from it using $x=10^{\log(x)}$.

In general; the interval generator can be any function which bunches the data where the residual function is changing rapidly and spreads it out where nothing much is happening. In this way the total number of samples can be reduced to a few hundred (say), rather than the thousands it might have taken to do the job well using a linear interval generator.

Dimensionless Goof

From the nature of the foregoing discussion, it might appear that the process of deciding how to go about fitting a given set of data is somewhat vaguely defined. Certainly, it can seem like that when trying to deal with unusual problems; but in fact there is an underlying logic which, if applied specifically, will indicate what should be done. It is a matter of addressing two issues:

- Are the errors normally distributed?
- Is the Goof (effectively) dimensionless?

If the errors are normally distributed (such as when fitting experimental data) then the data should be subjected to a least-squares fit. In that case, the Goof is:

$$\chi^2 = \sum(w_i e_i^2)$$

and if the residuals are defined as

$$e_i = y_{i(\text{obs})} - y_{i(\text{calc})}$$

then the weights must be defined as proportional to reciprocal variances.

We can re-write the expression for chi-squared as follows.

$$\chi^2 = k \sum[w_{i(\text{fit})} (e_i/\sigma_i)^2]$$

Now we have separated the fitting weight into three parts, i.e.;

$$w_i = w_{i(\text{fit})} k / \sigma_i^2$$

$w_{i(\text{fit})}$ is a fitting flag, which can be set to 1 if we want to include the i^{th} observation, or set to 0 if it is suspected that a logging mistake or a breakdown of the model prevents this particular observation from being fitted. The number of observations of finite weight (used for calculating reduced chi-squared) is then simply $\sum w_{i(\text{fit})}$. k is a global constant, which should be 1 if the standard deviations of the observations have been correctly scaled, but we may have to re-scale the standard deviations after fitting in order to get $k=1$.

Now notice that both e_i and σ_i have the same dimensions. Hence the effect of all of the decisions which were made in constructing the composite error function is to make the Goof dimensionless.

Now consider the problem of fitting data which do not have normally distributed errors. We can tell whether or not the errors are normally distributed by asking the question: 'Will the apparent standard deviation of fit change if the data are sampled in a different way?' In the case where one smooth function is being fitted to another (for example); the degree of agreement between the two functions will depend on the chosen range of x -values. Hence the concept of standard deviation

becomes meaningless, because the differences are not random and depend on choices made by the investigator. Hence we must first try to make the dataset comprehensive; by taking plenty of samples overall, and particularly in regions where the disagreement is greatest. This has the effect of maximising the composite error (however we might choose to define it). Then, still unable to define a standard deviation, we must use a runout criterion. Thus we define the Goof as:

$$\eta^{2m} = \Sigma(w_i e_i^{2m})$$

Now the weighting coefficients become empirical parameters, initially set to 1, but later used to adjust the final error boundaries. They cannot be defined as reciprocal variances because the errors are not random. Finally, we need to decide on a definition for the residuals. In most cases involving runout minimisation, the error is defined as a proportion, in which case the residuals are dimensionless. Indeed, it is not possible to define a meaningful single-valued statistic for the accuracy of the fitting function unless the error is expressed as a proportion (% , ppM, etc.). Thus, unless we intend to supply users of the function with a comprehensive table of differences; the residuals must be defined according to a criterion which makes the Goof dimensionless.

Using the preceding logic, we can, if so desired, use a runout criterion for fitting data with normally distributed (or quasi-normally distributed) errors. The typical reason for wanting to do that is to fit the data using the minimum possible number of variable parameters; i.e., to keep the fitting formula as simple as possible but still get a good fit.

A commonly encountered situation is that of wanting to produce a formula to reproduce theoretical data given as a table. The numbers in the table will have rounding errors; i.e., the errors will not be strictly normally distributed, but there is a relationship between rounding error and apparent standard deviation which gives important information about the fitting function. If the errors are randomly distributed, then the most probable error in an observation is σ_i . If the errors are due to rounding, then the most probable error is half the worst-case rounding error. Hence, if the fitting function has sufficient variable parameters to smooth the data, the estimated 'standard deviation' of fit will be very close to half the worst-case rounding error. Thus, for example, if we have data rounded to 4 decimal places, we should get

$$\sigma_{fit} = \sqrt{[(1/v) \Sigma(w_{i(fit)} e_i^2)]} = 0.000025$$

(where $w_{i(fit)}$ is a fitting flag, 0 or 1). If the ESD is much smaller than that, then there are too many parameters and the empirical function is fitting the noise. If the ESD is much larger than that, then the function does not fit the data.

Nevertheless, despite the diagnostic value of σ_{fit} , we still want to fit the data using a runout criterion. The reason is that the true errors have a 'brick-wall' tolerance, i.e., there should be no errors significantly outside the maximum rounding error, and a runout criterion will enforce that condition. Hence the success criterion is different from the fitting criterion. The Goof used by the fitting algorithm is:

$$\eta^{2m} = \Sigma(w_{i(fit)} e_i^{2m})$$

but note that all of the residuals have the same uncertainty (all of the numbers are rounded to the same number of decimal places). Hence, the residuals can all be notionally divided by an arbitrary number having the same dimensions, and the only effect will be to scale the Goof. Scaling the Goof (multiplying the whole thing by a constant) has no effect on the fit (provided that the operation does not lead to floating-point errors). Hence the Goof is effectively dimensionless.

Now, having spotted that the trick in fitting data properly is to make the Goof either actually or effectively dimensionless, we can use this requirement as a test of correctness. It is impossible to envisage all of the situations which might arise; but take, for example, the problem of fitting accurate tabulated data which do not have a constant number of decimal places. In that case, the most probable error differs between observations. The solution is to divide each residual by its most probable error. Thus, if we call the most probable error $e_{i(mp)}$, then we have:

for 2 decimal places: $e_{i(\text{mp})} = 0.0025$

for 3 decimal places: $e_{i(\text{mp})} = 0.00025$

for 4 decimal places: $e_{i(\text{mp})} = 0.000025$

etc. The fitting criterion is then:

$$\eta^{2m} = \Sigma[w_{i(\text{fit})} (e_i / e_{i(\text{mp})})^{2m}]$$

which is dimensionless.

The success criterion is reduced chi-squared:

$$\chi^2/v = (1/v) \Sigma[w_{i(\text{fit})} (e_i / e_{i(\text{mp})})^2] \approx 1$$

That this statistic should come out to be approximately 1 may not be immediately obvious, but note that an error divided by the most probable error should, on average, be 1. For a statistically representative dataset, the number of degrees of freedom (v) should tend towards the number of observations. Hence, if the number of observations is n ; the sum of n quantities which should be 1 on average, divided by a number which tends towards n , should tend towards 1 in the limit of large n if the data reduction has been carried out correctly.

* * *

A demonstration of the Nelder-Mead downhill simplex search method, using some of the fitting optimisation techniques discussed in this article is given in the accompanying Open Document spreadsheet: **Nelder_demo.ods** .

