

ERRORS AND ERROR ESTIMATION

Errors, precision and accuracy: why study them?

People in scientific and technological professions are regularly required to give quantitative answers. How long? How heavy? How loud? What force? What field? Their (and your) answers to such questions should include the value and an error. Measurements, values plotted on graphs, values determined from calculations: all should tell us how confident you are in the value. In the Physics Laboratories, you will acquire skills in analysing and determining errors. These skills will become automatic and will be valuable to you in almost any career related to science and technology. Error analysis is quite a sophisticated science. In the First Year Laboratory, we shall introduce only relatively simple techniques, but we expect you to use them in virtually all measurements and analysis.

What are errors?

Errors are a MEASURE OF THE LACK OF CERTAINTY IN A VALUE.

Example: The width of a piece of A4 paper is 210.0 ± 0.5 mm. I measured it with a ruler¹ divided in units of 1 mm and, taking care with measurements, I estimate that I can determine lengths to about half a division, including the alignments at both ends. Here the error reflects the limited resolution of the measuring device.

Example: An electronic balance is used to measure the weight of drops falling from an outlet. The balance measures accurately to 0.1 mg, but different drops have weights varying by much more than this. Most of the drops weigh between 132 and 139 mg. In this case we could write that the mass of a drop is (136 ± 4) mg. Here the error reflects the variation in the population or fluctuation in the value being measured.

Error has a technical meaning, which is not the same as the common use. If I say that the width of a sheet of A4 is 210 cm, that is a mistake or blunder, not an error in the scientific sense. Mistakes, such as reading the wrong value, pressing the wrong buttons on a calculator, or using the wrong formula, will give an answer that is wrong. Error estimates cannot account for blunders.

Learning about errors in the lab

The School of Physics First Year Teaching Laboratories are intended to be places of learning through supervised, self-directed experimentation. The demonstrators are there to help you learn. Assessment is secondary to learning. Therefore, do not be afraid of making a poor decision—it's a good way to learn. If you do, then your demonstrator will assist you in making a better decision.

Please avoid asking your demonstrator open ended questions like “How should I estimate the error”. That is a question you are being asked. Instead, try to ask questions such as “Would you agree that this data point is an outlier, and that I should reject it?”, to which the demonstrator can begin to answer by saying “yes” or “no”. However, do not hesitate in letting your demonstrator know if you are confused, or if you have not understood something.

Rounding values

During calculations, rounding of numbers during calculations should be avoided, as rounding approximations will accumulate. Carry one or two extra significant figures in all values through the calculations. Present rounded values for intermediate results, but use only non-rounded data for further processing. **PRESENT A ROUNDED VALUE FOR YOUR FINAL ANSWER.** Your final quoted errors should not have more than two significant figures.

Some important terms

Observed/calculated value A value, either observed or calculated from observations. e.g. the value obtained using a ruler to measure length, or the electronic balance to measure mass, or a calculation of the density based upon these.

True value The true value is a philosophically obscure term. According to one view of the world, there exists a true value for any measurable quantity and any attempt to measure the true value will give an observed value that includes inherent, and even unsuspected errors. More practically, an average of many repeated independent measurements is used to replace true value in the following definition.

Accuracy A measure of how close the observed value is to the true value. A numerical value of accuracy is given by:

$$\text{Accuracy} = 1 - \left(\frac{\text{observed value} - \text{true value}}{\text{true value}} \right) \times 100\%$$

Precision A measure of the detail of the value. This is often taken as the number of meaningful significant figures in the value.

Significant Figures Significant figures are defined in your textbook. Look carefully at the following numbers: 5.294, 3.750×10^7 , 0.0003593, 0.2740, 30.00. All have four significant figures. A simple measurement, especially with an automatic device, may return a value of many significant figures that include some non-meaningful figures. These non-meaningful significant figures are almost random, in that they will not be reproduced by repeated measurements. When you write down a value and do not put in errors explicitly, it will be assumed that the last digit is meaningful. Thus 5.294 implies $5.294 \pm \sim 0.0005$.

For example, I have just used a multimeter to measure the resistance between two points on my skin, and the meter read 564 k Ω —the first time. Try it yourself. Even for the same points on the skin, you will get a wide range of values, so the second or third digits are meaningless. Incidentally, notice that the resistance depends strongly on how hard you press and how sweaty you are, but does not vary so much with which two points you choose. Can you think why this could be?

Systematic and random errors

A systematic error is one that is reproduced on every simple repeat of the measurement. The error may be due to a calibration error, a zero error, a technique error due to the experimenter, or due to some other cause. A random error changes on every repeat of the measurement. Random errors are due to some fluctuation or instability in the observed phenomenon, the apparatus, the measuring instrument or the experimenter.

Independent and dependent errors

The diameter of a solid spherical object is 18.0 ± 0.2 mm. The volume, calculated from the usual formula, is 3.1 ± 0.1 cm³ (check this, including the error). These errors are dependent: each depends on the other. If I overestimate the diameter, I shall calculate a large value of the volume. If I measured a small volume, I would calculate a small diameter. Any measurements made with the same piece of equipment are dependent.

Suppose I measure the mass and find 13.0 ± 0.1 g. This is an **independent error**, because it comes from a different measurement, made with a different piece of equipment. There is a subtle point to make here: if the error is largely due to resolution error in the measurement technique, the variables mass measurement and diameter measurement will be **uncorrelated**: a plot of mass vs diameter will have no overall trend. If, on the other hand, the errors are due to population variation, then we expect them to be correlated: larger spheres will probably be more massive and a plot will have positive slope and thus **positive correlation**. Finally, if I found the mass by measuring the diameter, calculating the volume and multiplying by a value for the density, then the mass and size have inter-dependent errors.

Standard deviation
(σ_{n-1})

The standard deviation is a common measure of the random error of a large number of observations. For a very large number of observations, 68% lie within one standard deviation (σ) of the mean. Alternatively, one might prefer to define their use of the word "error" to mean two or three standard deviations. The sample standard deviation (σ_{n-1}) should be used. This quantity is calculated automatically on most scientific calculators when you use the ' $\sigma+$ ' key (see your calculator manual).

Absolute error

The error expressed in the same dimensions as the value. e.g. 43 ± 5 cm

Percentage error

The error expressed as a fraction of the value. The fraction is usually presented as a percentage. e.g. 43 cm \pm 12%

Error Estimation

We would like you to think about the measurements and to form some opinion as to how to estimate the error. There will possibly be several acceptable methods. There may be no “best” method. Sometimes “best” is a matter of opinion.

When attempting to estimate the error of a measurement, it is often important to determine whether the sources of error are systematic or random. A single measurement may have multiple error sources, and these may be mixed systematic and random errors.

To identify a random error, the measurement must be repeated a small number of times. If the observed value changes apparently randomly with each repeated measurement, then there is probably a random error. The random error is often quantified by the standard deviation of the measurements. Note that more measurements produce a more precise measure of the random error.

To detect a systematic error is more difficult. The method and apparatus should be carefully analysed. Assumptions should be checked. If possible, a measurement of the same quantity, but by a different method, may reveal the existence of a systematic error. A systematic error may be specific to the experimenter. Having the measurement repeated by a variety of experimenters would test this.

Error Processing

The processing of errors requires the use of some rules or formulae. The rules presented here are based on sound statistical theory, but we are primarily concerned with the applications rather than the statistical theory. It is more important that you learn to appreciate how, in practice, errors tend to behave when combined together. One question, for example, that we hope you will discover through practice, is this: **How large does one error have to be compared to other errors for that error to be considered a dominant error?**

An important decision must be made when errors are to be combined. You must assess whether different errors are **dependent** or **independent**. Dependent and independent errors combine in different ways. When values with errors that are dependent are combined, the errors accumulate in a simple linear way. If the errors are independent, then the randomness of the errors tends, somewhat, to cancel out each other and so they accumulate **in quadrature**, which means that their squares add, as shown in the examples below.

If the errors are dependent (based on the same source of error) they add linearly:

When	Example	Method
Adding or subtracting	$l = x + x$ $y = x_1 - x_2$ $p = 2(l + h)$, when the dominant error is a systematic error common to the measurement of l and h	Add absolute errors $\Delta y = \Delta x + \Delta x = 2\Delta x$ $\Delta y = \Delta x_1 + \Delta x_2$ $\Delta p = 2(\Delta l + \Delta h)$ <i>Note the plus sign</i>
Multiplying or dividing	$a = b \times c$ $v = l \times h \times b$, where the dominant error in each is a systematic error common to l, h, b	Add fractional/percentage errors $\frac{\Delta a}{a} = \frac{\Delta b}{b} + \frac{\Delta c}{c}$ $\frac{\Delta v}{v} = \frac{\Delta l}{l} + \frac{\Delta h}{h} + \frac{\Delta b}{b}$
Using formulae	$y = f(x)$	Use the differential method ¹ or test the extreme points. $\Delta y = \Delta x \times f'(x)$ or $\Delta y = \frac{f(x + \Delta x) - f(x - \Delta x)}{2}$

If the errors are independent (they have different, uncorrelated sources) they add in quadrature:

Operation	Example	Method
Adding or subtracting	$z = x + y$ $p = 2(l - h)$	Add absolute errors in quadrature. $\Delta z = \sqrt{(\Delta x)^2 + (\Delta y)^2}$ $\Delta p = 2\sqrt{(\Delta l)^2 + (\Delta h)^2}$
Multiplying or dividing	$z = x \times y$ $v = l \times h \times b$	Add fractional (percentage) errors in quadrature. $\frac{\Delta z}{z} = \sqrt{\left(\frac{\Delta x}{x}\right)^2 + \left(\frac{\Delta y}{y}\right)^2}$ $\frac{\Delta v}{v} = \sqrt{\left(\frac{\Delta l}{l}\right)^2 + \left(\frac{\Delta h}{h}\right)^2 + \left(\frac{\Delta b}{b}\right)^2}$
Using formulae	$z = f(x, y)$	Add partial derivative weighted errors ¹ in quadrature. $\Delta z = \sqrt{\left(\Delta x \times \frac{\partial f}{\partial x}\right)^2 + \left(\Delta y \times \frac{\partial f}{\partial y}\right)^2}$ or $\Delta z = \sqrt{\left(\frac{f(x + \Delta x, y) - f(x - \Delta x, y)}{2}\right)^2 + \left(\frac{f(x, y + \Delta y) - f(x, y - \Delta y)}{2}\right)^2}$

² Beware of the possibility of a turning point between $x + \Delta x$ and $x - \Delta x$, which will cause this method to fail. (e.g. $y = \sin(x)$, where $x = 90 \pm 10^\circ$)

Averages

When performing an experiment, a common way of obtaining a better estimate of something which you are trying to measure, is to take repeated measurements and calculate the average, or mean, of these measurements. Random errors in the measurements cause them all to be slightly different, so the mean may be thought of as an estimate of the “true value”.

For a set of n measurements $(x_1, x_2, x_3, \dots, x_n)$ the mean, \bar{x} , is given by

$$\bar{x} = \frac{\sum_{i=1}^n x_i}{n}$$

For example, supposing we measured the time it takes for a ball to fall through a height of 3 metres, and repeat it another six times. The results might look like this:

Time (s)	1.11	1.18	1.02	1.09	1.10	1.13	1.12
----------	------	------	------	------	------	------	------

The mean of this data is given by

$$\frac{1.11 + 1.18 + 1.02 + 1.09 + 1.10 + 1.13 + 1.12}{7} = 1.107 \text{ s}$$

Just as there is uncertainty in each of the measurements, there is also uncertainty in the mean. This uncertainty may be calculated in a variety of ways, which depend on how many measurements have been made.

The simplest estimate of the uncertainty in the mean comes from the range of the data. The range is simply the difference between the most extreme values in the set. In the above example the extreme (highest and lowest) values are 1.18 and 1.02, so the range is 0.16. The uncertainty in the mean, then, is given by the half of the range, that is

$$\Delta\bar{x} = \frac{x^+ - x^-}{2}$$

where x^+ and x^- are the highest and lowest values in the set, respectively. This method provides a good estimate when you have less than 10 measurements. So in our example, the uncertainty in the mean time is $0.16/2 = 0.08$, so that we can write the mean as (1.11 ± 0.08) s. Note that we have dropped the last significant figure in the mean, as the uncertainty is large enough to make it meaningless.

If you have taken a lot of measurements, then the uncertainty in the mean is given by the standard deviation, σ (mentioned earlier on page 18). Most calculators can calculate the standard deviation, but you can also use the formula:

$$\sigma = \sqrt{\frac{\sum (x_i - \bar{x})^2}{n}}$$

and the error in the mean is given by

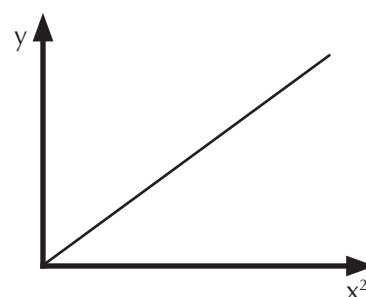
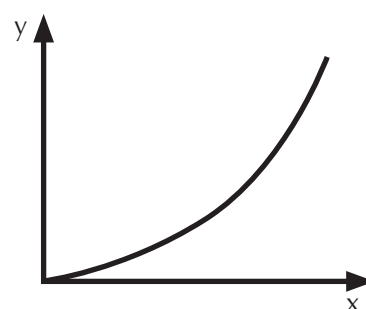
$$\Delta\bar{x} = \sigma_{n-1} = \frac{\sigma}{\sqrt{n-1}}$$

Generally, in First Year Lab, you won't be taking very large sets of measurements, and so will not be using this method. Use half the range for a quick simple error estimate, taking care to be sure that your extreme values are not outliers (unusually small or large values, usually the result of some one-off mistake). Outliers should generally be discarded when taking averages, and when graphing.

Graphing

Scientists and technologists very often organise their variables so that a particular theory becomes a straight line on the plot. The reason is that a straight line is very easy to recognize, and even small departures from it can be easily seen. This is harder with curves. For instance, suppose that a theory predicts $y = ax^2$. The curve in the upper graph at left looks a bit like $y = ax^2$, but it also looks a bit like $y = a(1 - \cos x)$, and a bit like $y = bx^{1.8}$. It's hard to tell.

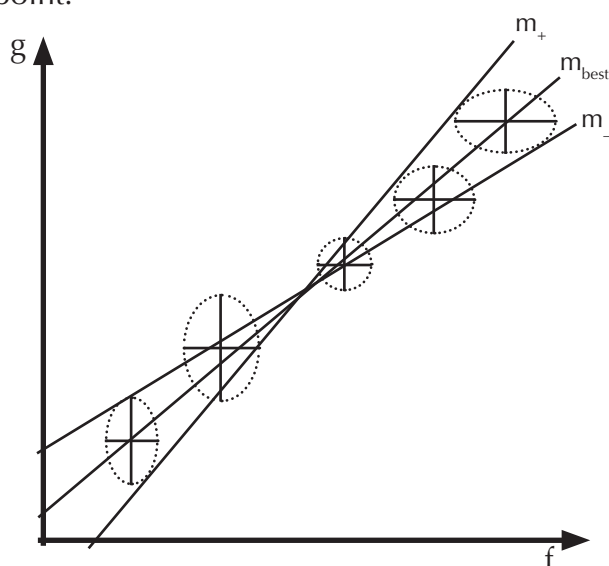
In the lower curve, y has been plotted against x^2 . Let's call the new variable $z \equiv x^2$. This graph does look rather like $y = az$, and it doesn't look like $y = a(1 - \cos \sqrt{z})$, unless z is very small, and it doesn't look at all like $y = bz^{0.9}$.



For the rest of this section, we shall therefore discuss only graphs in which the variables have been plotted to produce a straight line for the theory being investigated.

Graphing with Errors

When graphing, **plot error bars**. A very sharp pencil is good for getting the size just right. If error bars are less than about 1 mm, then do not try to show them: instead write the size of the error on the graph and also show your calculation of the error bar length. If both error bars (vertical and horizontal) are too small to plot, draw a circle around the experimental point.



Draw a line of *best fit* and lines of *worst fit*.

In many cases, your error bars will be the maximum probable error, so you should have a high probability that the true value lies within your error bar.

Sketch (or just imagine) an ellipse whose axes are the error bars—this is called *the error ellipse*². In this case, every line should pass through the error ellipse about every point. A worst fit line necessarily touches the edge of at least two points.

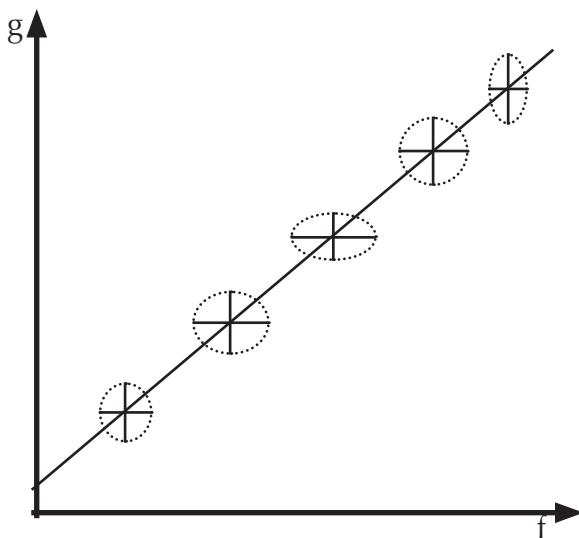
² We draw an ellipse, rather than a rectangle, because of the quadrature addition of independent errors.

The error in the gradient is

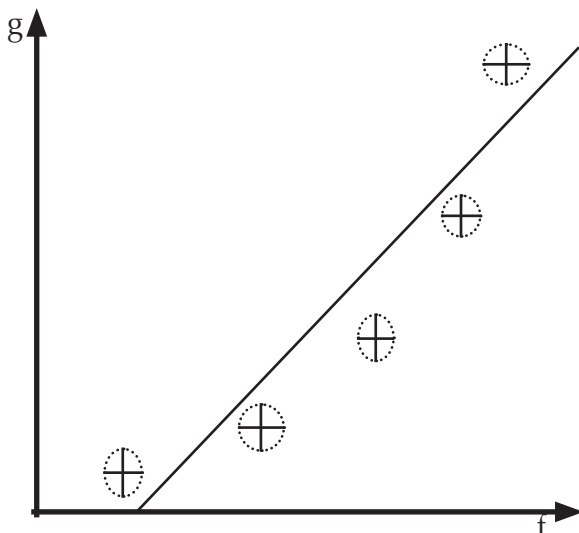
$$\Delta m = \frac{m_+ - m_-}{2}$$

To find the y-intercept, other lines of worst fit may have to be drawn. The worst fit that produces the greatest y-intercept, and the worst fit that produces the smallest y-intercept may not necessarily be the same as the worst fits used to find the extremes in gradient. The extremes in the y-intercept may be produced by a combination of rotating the fitted line and moving it without rotation.

In the case shown above, the deviation of the measured values from the fitted line are comparable in size to the error bars. This is a 'normal' case.



In the graph to the left, the error bars are large compared with the departure of the measured points from the fitted line. This suggests that the error estimates are too large: they should be re-examined.



In this graph, the error bars are small in comparison with the departure of the measured points from the fitted line. It is impossible to fit a straight line without rejecting a substantial fraction of the data as outliers. Such a result suggests either:

- (a) the error estimates are too small;
- (b) that the measurements were made carelessly;
- (c) that numerical blunders have been made in treating the data; or
- (d) that the relation is better described as non-linear, which means that the theory which gives a straight line in this plot is wrong or inappropriate here.

Further, the general shape of the points suggests that it would be a good idea to try a different plot, such as g vs f for $\ln g$ vs $\ln f$.

Automatic graphing routines

Most common software packages that graph data and fit lines or curves do not take errors into account. Many do not even plot the errors. Further, they give all points equal weight, even if there is a big variation in the error bars. Graphing by hand, as described above, you give the points with small error bars more importance (the statistical term is **weight**) because the line is more tightly constrained to pass through the smaller error bars. It is possible to include appropriate weighting factors (usually the reciprocal of the error) in automatic routines.

In the first year lab, you can use the excel template “Linear plot with Errors”. This will plot the line of best fit and also the maximum and minimum gradient lines that satisfy your data.

References

‘Experimental Methods. An Introduction to the Analysis and Presentation of Data.’ Les Kirkup, Wiley, (1994).

‘Data Reduction and Error Analysis for the Physical Sciences.’ Philip R. Bevington, McGraw Hill (1969).

‘Statistical Methods in Medical research.’ (3rd Edition) P. Armitage & G. Berry, Blackwell: Oxford, (1994).

‘Handling Experimental Data.’ Mike Pentz, Milo Shott and Francis Aprahamian, Open University Press (1988).

