Accuracy and Errors in Experimental Engineering*

[These notes summarise some basic statistical ideas. If you have not done A-level statistics you will probably need to do some background reading – see, for example C. Chatfield, *Statistics for Technology*, Chapman and Hall, 1978]

Any experiment which produces numerical results should always be accompanied by an estimate of the accuracy of those results. If the experiment is repeated by someone else using similar specimens and apparatus, what sort of agreement is likely between the two sets of results? Variation in experimental results comes from several sources.

1. Genuine variation in the quantities being determined, due to the unavoidable variation in the physical properties of real materials.
2. Instrumental inaccuracies. These may be systematic (due to a faulty zero or scale constant) or random (due to friction or "noise").
3. Estimates made when fractional parts of scale divisions are read.
4. Numerical rounding-off during calculations.

In a well-designed experiment the "experimental error" (due to sources 2, 3 and 4) should be small relative to any genuine variation (source 1). The various measurements made during the experiment should have roughly the same degree of accuracy. This implies that you should know the accuracy of each of the instruments you use.

As far as possible each instrument in the Department's laboratories is checked regularly to make sure that it conforms to the limits of accuracy set for it. These limits will usually be found marked on the instrument itself and are set out in various British Standards which may be referred to either in the laboratory or the Library. Reference to a British Standard will only indicate the <u>maximum</u> error that an instrument may have. If more detailed knowledge of the performance of a particular instrument is required it must be <u>calibrated</u> against a high-accuracy standard. Individual values read from the scale of the instrument may then be adjusted in accordance with the calibration.

1. <u>The "external" or "a priori" estimation of experimental error</u>

An "external" or "a priori" estimate of the possible error in an experiment is an estimate we make (or could make) <u>before</u> we actually do the experiment. It is constructed by estimating the magnitudes of the individual errors associated with sources 2 to 4 in the list given above, and carrying these errors through the numerical processing of the experimental data to give the corresponding error in the result.

* The term "error" used in these notes refers to the intrinsic inaccuracy of any experimental procedure. It is not the same as a "mistake". Mistakes may happen when an instrument is read (by using the "wrong" voltage scale, reading a micrometer as 10.53 mm instead of 9.53 mm, etc.) when a value is transcribed from a laboratory notebook to a report or during a calculation. The only way to guard against mistakes is to check each phase of the work as soon as it is completed. Once you have left the laboratory, or even changed the setting of your apparatus, it may be too late.

An error in some quantity x can be expressed as an absolute error ε or as a relative (or proportional) error r, where $r = \varepsilon / x$. Numerical processing involves:

(a) Addition and subtraction: if two quantities x_1 and x_2 may be in error by as much as $\pm \varepsilon_1$ and $\pm \varepsilon_2$ then the absolute error in either the sum or difference of x_1 and x_2 may be as much as $\pm (\varepsilon_1 + \varepsilon_2)$. Note that subtracting two nearly equal quantities can produce a large value for the relative error, $\pm (\varepsilon_1 + \varepsilon_2)/(x_1 - x_2)$, in the result.

(b) Multiplication and division: if two quantities x_1 and x_2 may be in error by as much as $r_1 x_1$ and $r_2 x_2$ then the relative error in the product or quotient of x_1 and x_2 may be as much as $\pm (r_1 + r_2)$, provided r_1 and r_2 are small.

Note that absolute errors must be used when dealing with sums and differences: relative errors must be used when dealing with products and quotients.

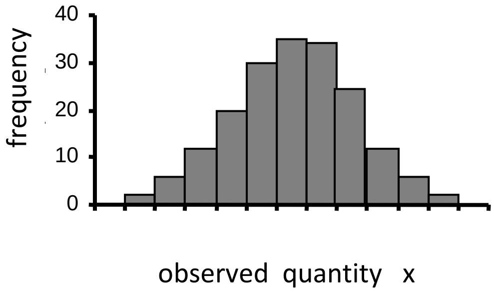

Fig. 1

The above ideas are usually adequate. However, it may be argued that if x_1 and x_2 are independent it is unlikely that the largest errors in both quantities will occur simultaneously. It may be shown that if the probable absolute errors in x_1 and x_2 are $\pm \varepsilon_1$ and $\pm \varepsilon_2$ respectively then the probable absolute error in the sum or difference of x_1 and x_2 is $\pm \sqrt{\varepsilon_1^2 + \varepsilon_2^2}$. Similarly, if the probable relative errors in x_1 and x_2 are $\pm r_1$ and $\pm r_2$ then the probable relative error in the product or quotient of x_1 and x_2 is $\pm \sqrt{r_1^2 + r_2^2}$. (A precise definition of probable error is given in Section 3.)

2. The "internal" or "a posteriori" estimation of experimental error

An "internal" or "a posteriori" estimate of the accuracy of an experiment is obtained by carrying out the experiment a number of times to obtain a sample and examining the "spread" of the sample. If we measure a quantity x a total of n times we obtain a set of n values x_i which forms a distribution. This distribution includes both the genuine variation in x and the experimental error. We can present the information graphically as a histogram (see Fig. 1), dividing the range of the values x_i into a series of equal intervals Δx and drawing a series of rectangles of width Δx and height equal to the number (or proportion) of values lying in that interval. The mean of the sample x is defined as $\bar{x} = \Sigma x_i / n$. The variance s^2, which gives us a measure of the spread of the sample, is defined as $s^2 = \Sigma(x - \bar{x})^2/n$. The quantity s is known as the standard deviation of the sample distribution. Systematic errors in the experiment are likely to shift the value of $\bar{x}$ without affecting the value of s.

The quantity x which the experiment is designed to measure will also have a distribution. This can be displayed as a continuous curve $\phi(x)$ in which the area under the curve between two ordinates x_1 and x_2 represents the probability of a single experimental result lying between x_1 and x_2 (see Fig. 2). The total area under the curve, $\int_{-\infty}^{+\infty} \phi(x)dx$, is equal to unity. The mean of the distribution is $\mu = \int_{-\infty}^{+\infty} x\phi(x)dx$ and the variance is $\sigma^2 = \int_{-\infty}^{+\infty} (x-\mu)^2 \phi(x)\, dx$.

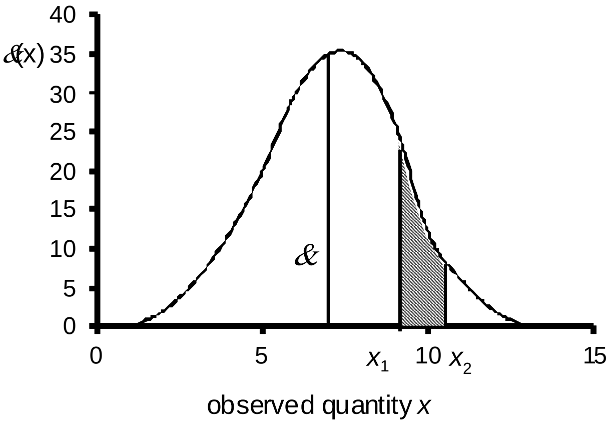

Fig. 2

The purpose of an experiment is normally to estimate the values of μ and σ from the experimental values of $\bar{x}$ and s. If we have grounds for believing that the experimental errors are "small" then the best estimate of μ is equal to $\bar{x}$ and the best estimate of σ is given by the expression $\sqrt{\{n/(n-1)\}}\, s$.

The value of $\bar{x}$ determined from the n experiments also has a distribution, which has mean μ and standard deviation $\sigma/\sqrt{n}$. This reduction in the standard deviation is a mathematical expression of our intuitive belief that the mean $\bar{x}$ of a set of experimental values is likely to be a better estimate of the true value of μ than the value x_1 obtained from a single experiment.

3. Statements about errors: the normal distribution

If no information is available about the shape of a distribution it is usually assumed to be normal. The shape of the normal distribution is given in Fig. 3.

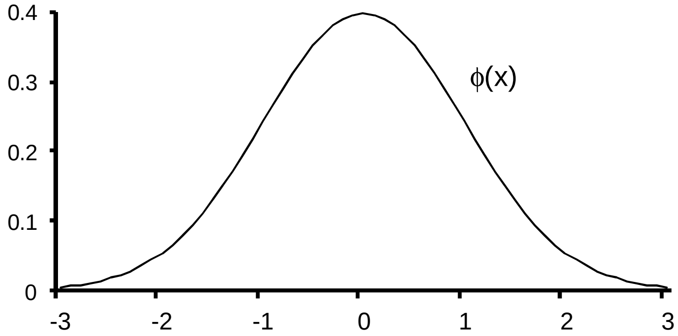

Fig. 3

If a variable x has a normal distribution with mean μ and standard deviation σ then:

50% of the measurements of x will lie within the range $\mu \pm 0.67\sigma$
68% of the measurements of x will lie within the range $\mu \pm \sigma$
95% of the measurements of x will lie within the range $\mu \pm 2\sigma$
99.73% of the measurements of x will lie within the range $\mu \pm 3\sigma$

The above percentages are often expressed as confidence limits. For example, if x_1 is an experimental value of a variable x with unknown mean μ and known standard deviation σ we can say "with 95% confidence" that x_1 lies in the range $x_1 - 2\sigma$ to $x_1 + 2\sigma$. The quantity 0.67σ is sometimes referred to as the probable error - see section 1.

If the sample size n is small in your experiment then the sample variance is not normally distributed but instead is distributed according to Student's-t distribution (see text books and the Mathematics Data Book for details of this distribution). It follows that the confidence intervals just quoted are too optimistic and we actually have a broader spread. Thus ± one standard deviation will not cover as much as 68% of the measurements.

A useful approximation in the laboratory is to re-estimate the variance σ^2 as $s^2n/(n - 3)$ and then use the normal distribution probabilities. This works well down to a sample of $n = 5$.

4. Plotting graphs

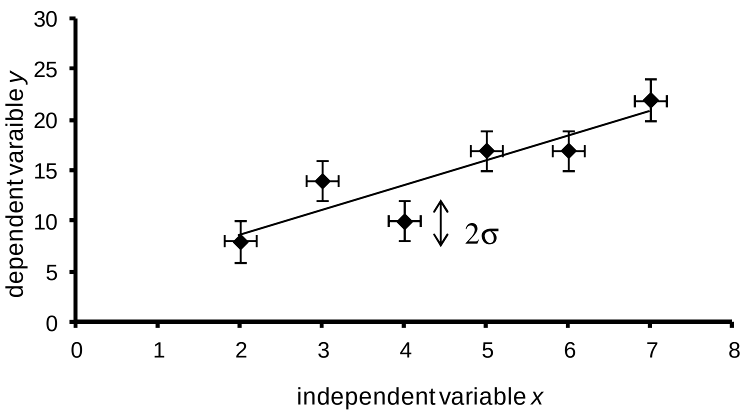

Fig. 4

In many experiments you will be asked to find the relationship of some quantity y to another quantity x. Plotting your results as precise points in the (x, y) plane leads to the problem of finding the "best" straight line (or higher-order polynomial) fit to the points. The method of least squares, described in the mathematics course, provides a solution to this problem. However it is better to include external estimates of the experimental errors at the plotting stage by plotting your results as rectangles, as shown in Fig. 4. (In this figure σ is the standard deviation of the estimated distribution of experimental error.) If you can draw a straight line through at least 2/3 of the rectangles then your experiment is consistent with the hypothesis that y is a linear function of x.

Some programmable calculators have a built-in routine for finding the best straight-line fit to a set of points. However, the hypothesis that one variable in your experiment is linearly dependent on another is something that should come from engineering science, not from the experiment itself. Plot the points first, to see if the relationship is a linear one. Then use your calculator, if you want to, to find the numerical values of the parameters of the line.

5. <u>Simple examples showing estimation of errors</u>

(a) The constant acceleration a of a slowly moving object is to be found by determining the time t taken to traverse a measured distance s. The equation of motion that applies is:

$$s = \frac{1}{2}at^2 . \quad \text{Rearranging, } a = \frac{2s}{t^2} .$$

The time is measured with a stopwatch, the distance with a metre ruler. The measured values and their errors are:-

$s = 2 \pm 0.005$ m.	This is 0.25 %.
$t = 4.2 \pm 0.2$ s.	This is 4.8 %.

What is the acceleration and its estimated error?

The relative errors in a and t^2 may be added to give the relative error in a. The relative error in t^2 is twice the relative error in t.

Hence relative error in a is: $0.25\% + 2\times 4.8\%$ = 9.8 %.

The size of the relative error in the time measurement, plus the factor of 2, causes that term to dominate. The ¼ per cent error due to the distance measurement is clearly negligible compared to the 9.6% error due to the time measurement, so the result (the acceleration) would most sensibly be written:

$$a = 0.23 \pm 0.02 \text{ m s}^{-2}.$$

(b) An inductor is formed by winding N_1 turns of wire in the form of a cylindrical coil of length l and diameter d_1. A second circular coil with N_2 turns and of smaller diameter d_2 is placed coaxially within the first, at its centre. The mutual inductance between the two windings is to be determined from measurements on the two coils, using the formula:-

$$M = \frac{\pi\mu_0}{4}\frac{N_1 N_2 d_2^2}{\sqrt{d_1^2 + l^2}} \text{ H}$$

The numbers of turns were counted and are exact. Lengths were measured with a metre rule; the measured values and their errors (percentages quoted in parentheses) are:-

N_1 = 200 turns (exact)	$d_1 = 20 \pm 0.5$ mm (2.5 %)
N_2 = 20 turns (exact)	$d_2 = 10 \pm 0.5$ mm (5 %)
$\mu_0 = 4\pi \times 10^{-7}$ H m^{-1}	$l = 50 \pm 0.5$ mm (1 %)

The relative error in the numerator due to d_2^2 is: $2 \times 5 = 10\%$. Relative errors in d_1^2 and l^2 are 5% and 2% respectively, so the absolute values are: $d_1^2 = 400 \pm 20$ mm^2 and $l^2 = 2500 \pm 50$ mm^2, and the value of the term inside the square root is: 2900 ± 70 mm^2.

The relative error in the denominator is: $\frac{1}{2} \times 70/2900$, or 1.2%, and the overall relative error is 10 + 1.2 = 11.2 %. It can be seen that the most serious error here arises from poor accuracy in measurement of d_2. The result might therefore be written:

$$M = 7.3 \pm 0.8 \times 10^{-6} \text{ H}$$

October 2014
CR, RSC, DMH, CAH