Brief Introduction to Errors

"All measurements, however careful and scientific, are subject to some uncertainties".[1] When experimental measurements are made, they must be accompanied by reasonable estimates of these uncertainties, or **errors**. These may be systematic or random.

Systematic errors can arise from inaccuracies in the apparatus; does your instrument read zero when you are measuring nothing? These errors effect all the measurements the same way and are hard to detect. In the ideal case they can be measured directly by calibrating the apparatus, but since this is sometimes not feasible sensible estimates must be made.

Random errors arise since repeated measurements of the same quantity will not always give the same result. The spread in the measured values gives an indication of the random error.

Statistical Method

If only a single observation of a particular quantity can be carried out, then the estimate of the uncertainty in the observation, the error, must be made from the apparatus used to make the measurement. When, however, a quantity is measured many times it is possible to estimate the error using statistical methods. Note that such an approach can estimate only random errors; systematic errors cannot be obtained from statistics.

The Mean

The mean, $\bar{x}$, of a set of N measurements x_1, x_2, … x_N is

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

Standard Deviation and Standard Error

Whenever possible several measurements of the quantity should be made; for N measurements an estimate of the average uncertainty is given by the standard deviation, s,

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

In this case the standard error, s_m, on the mean, $\bar{x}$, is given by

$$s_m = \pm\frac{s}{\sqrt{N}}\,.$$

The significance of this is that if the same measurements are repeated a number of times 68% of results will lie within the range $\bar{x} \pm s_m$; this is called a 68% confidence level. A more realistic error estimate is the 95% confidence level. Provided that 10 or more observations have been carried out this is $2s_m$.

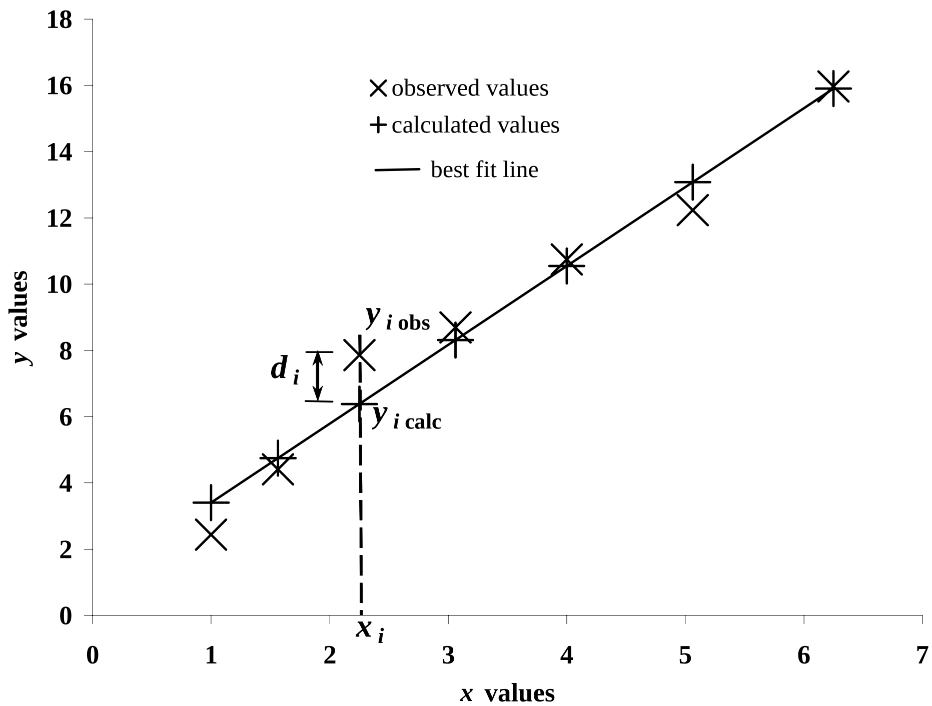
[1] JR Taylor *An Introduction to Error Analysis* second edition University Science Books:California (1997)

Graphs – the Method of Least Squares

When a value of y is linearly related to a quantity x then

$$y = a + bx$$

where a (the intercept) and b (the slope) are parameters to be determined from the measurements. In an experiment N sets of data (x_1, y_1), (x_2, y_2) … (x_N, y_N) are collected, by setting the x values and measuring the corresponding values of y. This should mean that the most significant error is in the y variable, with the error in x being much smaller. In practise y should be chosen as the variable with the dominant error.



The parameters a and b can then be estimated from the data by the method of least squares. Notice that at each value of x, say x_i, there are two y values, namely $y_{i\text{obs}}$ and $y_{i\text{cal}}$. The former is the experimentally measured value of y_i and in general, due to error, will not be exactly on the best fit line. The calculated value of y_i is $y_{i\text{cal}}$ and, by definition, must lie on the best fit line. In the method of least squares the differences $d_i = y_{i\text{cal}} - y_{i\text{obs}}$ are determined. The differences are then squared and added, and the sum is minimised by varying a and b. This method gives

$$b = \frac{N\sum x_i y_i - \sum x_i \sum y_i}{N\sum x_i^2 - \left(\sum x_i\right)^2} \qquad \text{and} \qquad a = \bar{y} - b\bar{x}$$

where

$$\bar{y} = \frac{\sum y_i}{N} \qquad \text{and} \qquad \bar{x} = \frac{\sum x_i}{N}.$$

Error Estimation for the Intercept and Slope

It can be shown that the standard errors on *a* and *b* are given by

$$s_a = \sqrt{\frac{\sum x_i^2 \sum d_i^2}{(N-2)\left[N\sum x_i^2 - \left(\sum x_i\right)^2\right]}} \quad \text{and} \quad s_b = \sqrt{\frac{N\sum d_i^2}{(N-2)\left[N\sum x_i^2 - \left(\sum x_i\right)^2\right]}}$$

$$\therefore\ s_a = s_b\sqrt{\frac{\sum x_i^2}{N}}$$

where

$$\sum d_i^2 = \sum\left(y_{i\text{obs}} - y_{i\text{cal}}\right)^2 .$$

Thus for a 68% confidence level the values for the intercept and slope and their errors are

$$a \pm s_a \quad \text{and} \quad b \pm s_b$$

or, for an approximate 95% confidence level,

$$a \pm 2s_a \quad \text{and} \quad b \pm 2s_b.$$

Combination of Errors

Most physical experiments involve more than a single variable being measured. In the majority of cases several variables are measured and these are then combined to give the final result. Thus the error on each variable must be estimated and these are then combined to give the total experimental error. Note that the errors on the different variables may be estimated in different ways. For example they may be physical, or from the standard error on *N* data points, or from the error on the slope or intercept of a graph.

In practice the equation for combining errors has to be broken down into those terms which are added or subtracted and those which are multiplied or divided.

Terms added or subtracted

For two variables the expression can be broken down into the form

$$u = ax \pm by$$

where x and y are the measured variables and a and b are constants. The errors in x and y are then estimated (by whatever method) as Δx and Δy and the error in u, Δu, is given by

$$\Delta u = \sqrt{(a\Delta x)^2 + (b\Delta y)^2} .$$

Thus the actual errors Δx and Δy are used.

Terms multiplied or divided

If $u = axy$ or $u = \frac{ax}{y}$

then $\frac{\Delta u}{u} = \sqrt{\left(\frac{\Delta x}{x}\right)^2 + \left(\frac{\Delta y}{y}\right)^2}$, i.e., $\Delta u = u\sqrt{\left(\frac{\Delta x}{x}\right)^2 + \left(\frac{\Delta y}{y}\right)^2}$.

Thus the fractional errors are used.

Variable raised to a power

If $u = ax^n$

then $\frac{\Delta u}{u} = n\frac{\Delta x}{x}$, i.e., the fractional error is increased by a factor equal to the power.