Experimental Uncertainties (Errors)

Some references for the following discussion:

[SQU85] G.L. Squires, Practical Physics, 3rd edition, Cambridge University Press, Cambridge, 1985, p. 7-54.
[TAY82] J.R. Taylor, An Introduction to Error Analysis, Oxford University Press, Oxford, 1982.
[BRA83] S. Brandt, Statistical and Computational Methods in Data Analysis, North-Holland, Amsterdam, 1983.
[BEV69] P.R. Bevington, Data Reduction and Error Analysis for the Physical Sciences, McGraw-Hill, 1969.

Errors

Assume you measure the resistance of a wire at two different temperatures and the two results are different. Does that mean that you have observed a temperature dependence of resistance? The answer is probably yes, provided that the uncertainty (or 'error') of the measurement is smaller than the difference between the two results. If you don't know the errors of the two measurements, there is no way you can tell. Thus, a measurement of a physical quantity is meaningless if it is not accompanied by an estimate of its uncertainty.

Systematic errors

Systematic errors always have the same sign. For instance, the effect of temperature on steel measuring tape, or a faulty calibration of a voltmeter will lead to systematic errors. If one is aware of a systematic error one can usually take it into account by applying a correction to the data. There is no recipe to find and deal with systematic errors. Using care, ingenuity and imagination, the experimenter must make an attempt to identify potential sources of systematic errors. Additional measurements in which one varies parameters that should *not* have an effect are often used to find or to set an upper limit on possible systematic errors.

Random Errors

Random errors arise from fluctuations of the result of an experiment when it is repeated over and over. An example is the depth of a river that fluctuates because of surface disturbances, or the decay rate of a radioactive sample, which fluctuates because of counting statistics. The fluctuations are random in sign, and we write

$$x_{exact} = x_i \pm \delta x_i \tag{1}$$

where x_i is the observed quantity and δx_i is its departure from the true (but unknown) value. The average $|\delta x_i|$ characterizes the width of the distribution of x_i around the true value, x_{exact}.

Error Propagation

Assume that you have experimentally determined some parameters x, y, and their errors δx, δy, etc., but the result of interest is a *function* $f(x,y,\ldots)$ of these parameters. Let us further assume that the errors δx, δy, etc. are independent of each other, i.e., knowing one error does not tell you anything about any other error. Then the error of the function f is given by

$$\delta f = \sqrt{\left(\frac{\partial f}{\partial x}\right)^2 \delta x^2 + \left(\frac{\partial f}{\partial y}\right)^2 \delta y^2 + \cdots}\,. \tag{2}$$

To illustrate this, let us calculate two examples,

$$f = x \pm y; \qquad \text{then} \qquad \delta f = \sqrt{\delta x^2 + \delta y^2}\,, \tag{3}$$

and

$$f = \frac{x}{y}; \qquad \text{then} \qquad \frac{\delta f}{f} = \sqrt{\left(\frac{\delta x}{x}\right)^2 + \left(\frac{\delta y}{y}\right)^2} \tag{4}$$

The quantity $\delta x/x$ is also called the "relative" error of x.

The Weighted Mean

Assume that a given physical quantity has been measured a number of times with the results $x_i \forall \delta x_i$, where i=1...N. We can combine these observations into a single value by calculating the weighted mean $\langle x\rangle$ and its error $\delta\langle x\rangle$,

$$\langle x\rangle = \frac{\sum \frac{x_i}{\delta x_i^2}}{\sum \frac{1}{\delta x_i^2}} \qquad \delta\langle x\rangle = \sqrt{\frac{1}{\sum \frac{1}{\delta x_i^2}}} \tag{5}$$

The sum extends from 1 to N. Note, that measurements with a large δx contribute less to $\langle x\rangle$.

Estimate of Random Error by Repeating Measurements

If the error δx of a measurement is not yet known it can be determined experimentally by repeating a given measurement many times. Let the results be x_i, where i=1...N. The best estimate of the measured quantity is then the mean

$$\langle x\rangle = \frac{1}{N}\sum x_i \tag{6}$$

which follows from eq.5 when all δx_i's are taken to be the same, or $\delta x_1 = \delta x_2 = \ldots = \delta x_N = s$. The best estimate of the error of a *single* measurement is then

$$s = \sqrt{\frac{1}{N-1}\sum\left(x_i - \langle x\rangle\right)^2} \tag{7}$$

while the best estimate of the error of the average <*x*> is given by

$$\delta\langle x\rangle = \frac{s}{\sqrt{N}} \quad . \tag{8}$$

Probability Distribution

The exact outcome of a measurement is uncertain, but the probability *p(x)dx* to obtain a value between *x* and *x+dx* is usually known. This probability peaks at (or near) the true result x_{exact} and falls off on either side. The mean of the distribution is

$$\mu = \int_{-\infty}^{\infty} x\, g(x)\, dx \tag{9}$$

and the width parameter, also called *variance*, or standard deviation, σ, is given by

$$\sigma^2 = \int_{-\infty}^{\infty} (x-\mu)^2\; g(x)\, dx \tag{10}$$

Often, the following mathematical expression is used to describe the probability distribution

$$g(x) = \frac{1}{\sigma\sqrt{2\pi}}\; e^{-\frac{(\mu-x)^2}{2\sigma^2}} \quad . \tag{11}$$

This is called "normal", or "Gaussian" distribution. Not all experimental errors are of this form: one assumes a normal distribution because it is reasonable and easy to use mathematically. By inserting eq.11 into eqs.9 and 10, it is easy to show that the use of μ and σ in the three equations is consistent. The factor in front of the exponential function in eq.11 ensures that the probability to measure *anything* is unity:

$$\int_{-\infty}^{\infty} g(x)\, dx = 1 \quad . \tag{12}$$

Fig.1 shows the Gaussian distribution. For simplicity, we have substituted $y=(\mu-x)/\sigma$; thus, y=1 represents an observation that is 1 standard deviation away from the mean. We see that g(0)=0.399, *g*(1)=0.242, and that the full width at half the maximum value g(0) (FWHM) is larger than twice the standard deviation (one finds FWHM=2.345σ). In Fig.2 the area of *g(y)* for an interval centered around the mean is plotted. From this figure one can learn for instance that the probability for an observation to be within 1 standard deviation of the mean is about 68%. The probability for a 2σ interval is about 96%.

$$g(y) := \frac{1}{\sqrt{2\cdot\pi}}\cdot\exp\left[-\frac{(y)^2}{2}\right]$$

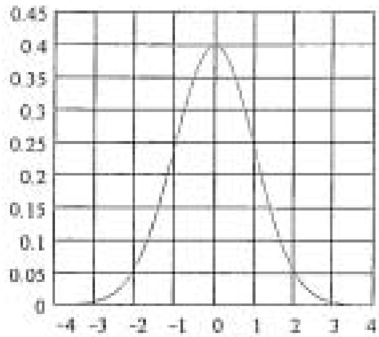

Fig.1

$$y := \frac{\mu - x}{\sigma}$$

$$\frac{1}{\sqrt{2\cdot\pi}}\cdot\int_{-yy}^{yy}\exp\left[-\frac{(y)^2}{2}\right]dy$$

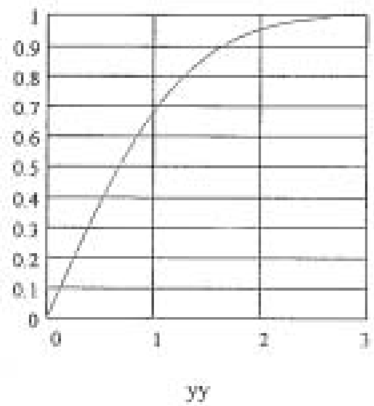

Fig.2

Rejection of data

Anomalous measurements can occur. When they can be traced to some systematic disturbance or to a mistake, such measurements can be corrected or rejected. If no explanation for the anomaly can be found, you are *not allowed to discard the observation*. Rather, this must be the stimulus for a more extensive study of your apparatus or the experimental procedure.