Quickie Statistics Summary

If N independent measurements are made of some quantity x and experimental conditions are not changed between measurements (the measurements are said to be drawn from the same *parent distribution* and constitute a *sample population*) then the best estimate of the "true" mean of x (the mean obtained as $N \to \infty$) is the mean of the sample population.

Mean (Bevington and Robinson, *Data Reduction and Error Analysis* (BR) p.9)

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

In this expression it is assumed that all N values x_n have the same uncertainty σ_n. In some instances this may not be true and σ_n will differ from measurement to measurement. If this is the case, the best estimate of the mean is the *weighted* mean (BR p.57)

$$\bar{x} = \frac{\sum_{1}^{N} w_n x_n}{\sum_{1}^{N} w_n} \qquad \text{where the weights } w_n = \frac{1}{\sigma_n^2}$$

Variances

For the case of equal uncertainties, an unbiased estimate of σ_n^2, the variance of an individual measurement of a sample of N measurements is (BR p.11 and p.54)

$$\sigma_n^2 = \frac{1}{N-1}\sum_{1}^{N}(x_n - \bar{x})^2,$$

and the variance in the mean of a sample of N measurements, $\sigma_{\bar{x}}^2$, is (BR p.54)

$$\sigma_{\bar{x}}^2 = \frac{1}{N}\sigma_n^2$$

This important result states that the uncertainty in the mean of N measurements decreases like $\frac{1}{\sqrt{N}}$. For the case in which the individual variances are not equal, the variance in the mean is given by (BR p. 57)

$$\frac{1}{\sigma_{\bar{x}}^2} = \sum_1^N \frac{1}{\sigma_n^2}$$

In general, the quantity reported for a measurement will be $\bar{x} \pm \sigma_{\bar{x}}$. If the parent distribution is a *normal* (Gaussian BR pp. 27-30) distribution, the most common case, then the true mean will be within the range $\bar{x} \pm \sigma_{\bar{x}}$ 68% of the time. Stated in other terms, if the set of *N* measurements is repeated, 68% of the time the mean obtained in the second set will lie within the range $\bar{x} \pm \sigma_{\bar{x}}$.

Propagation of Errors

When the quantity being measured, let's call it *u*, is some combination of <u>independent</u> quantities which we will call *x, y, z,*···, *u*=*u*(*x,y,z,*···), there is a simple general rule for calculating the uncertainty in *u*, σ_u, given the uncertainities $\sigma_x, \sigma_y, \sigma_z, \cdots$ in *x, y, z,*···. This is

$$\sigma_u^2 = \left(\frac{\partial u}{\partial x}\right)^2_{\bar{x},\bar{y},\bar{z},} \sigma_x^2 + \left(\frac{\partial u}{\partial y}\right)^2_{\bar{x},\bar{y},\bar{z},} \sigma_y^2 + \left(\frac{\partial u}{\partial z}\right)^2_{\bar{x},\bar{y},\bar{z},} \sigma_z^2 + \cdots$$

Two examples of this rule are of particular interest. The first is the situation in which *u* is the sum or difference of the quantities *x, y, z,*···, for example, *u* = *x* + *y* – *z*. The partial derivatives of *u* with respect to *x, y* and z are either +1 or –1 so the expression for the uncertainty of u reduces to (BR p.42)

$$\sigma_u^2 = \sigma_x^2 + \sigma_y^2 + \sigma_z^2$$

The square of the uncertainty of *u* is the sum of the squares of the uncertainties of *x, y,* and *z*.

The second example is that in which u can be expressed as the product and/or quotient of *x*, *y*, and *z*, for example, $u = \frac{xy}{z}$. It is a simple matter to show that the general expression reduces in this case to (BR p. 43)

$$\left(\frac{\sigma_u}{u}\right)^2 = \left(\frac{\sigma_x}{x}\right)^2 + \left(\frac{\sigma_y}{y}\right)^2 + \left(\frac{\sigma_z}{z}\right)^2$$

Thus the square of the *fractional* (or *relative*) uncertainly in *u*, $\frac{\sigma_u}{u}$, is the sum of the squares of the *fractional* (or *relative*) uncertainties in *x*, *y*, and *z*.

For both of the examples above, if the uncertainty in one of the quantities *x*, *y*, or *z* is several times that of the uncertainties in the other quantities it dominates the uncertainty in *u*. For example, if the uncertainty is *x* is twice that in *y* and *z*, $\sigma_x = 2\sigma_y = 2\sigma_z$, the uncertainty in *x* contributes 82% of the uncertainty in *u*.