

Elementary Statistics of Measurement

Suppose that you measure a quantity, x , and repeat the measurement n times. All measurements have some associated error, often called “noise”, and hence you would expect that the n values that you measured will not all be the same (if they are, you should either attempt to measure to more significant figures or quit cheating). Assuming that you have been careful to make each measurement in the same way so that each should have the same uncertainty, then the most likely value of x , which we shall denote X (\bar{x} and $\langle x \rangle$ are also often used), is the mean:

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

The Excel spreadsheet function that does this is *average*. You can estimate the uncertainty of each individual measurement, $\sigma(x)$, from the root-mean scatter about the mean:

$$\sigma(x) = \sqrt{\frac{1}{(n-1)} \sum_{i=1}^n (x_i - X)^2}$$

The Excel function that does this is *stdev*. The uncertainty of the mean value, $\sigma(X)$, is:

$$\sigma(X) = \frac{\sigma(x)}{\sqrt{n}} = \sqrt{\frac{1}{n(n-1)} \sum_{i=1}^n (x_i - X)^2}$$

Note that $\sigma(x)$ is independent of the number of measurements, n , while the precision of the mean, $\sigma(X)$, improves as the square root of the number of measurements.

Suppose you want to calculate a quantity, y , that is some function of the measured quantity, x (i.e. $y = f(x)$). The best estimate of y is Y , where

$$Y = f(X),$$

and the uncertainty of Y is $\sigma(Y)$:

$$\sigma(Y) = \left| \frac{df(x)}{dx} \right|_{x=X} \sigma(X).$$

In some cases it may be more convenient to calculate a value of y for each observed x and calculate the mean y and its uncertainty from these values.

If a calculated quantity is a function of more than one measured quantity, $z = f(x, y)$, then the best estimate of z is Z , where

$$Z = f(X, Y),$$

and its uncertainty is $\sigma(Z)$:

$$\sigma(Z) = \sqrt{\left| \frac{\partial f(x, y)}{\partial x} \right|_{X, Y}^2 \sigma^2(X) + \left| \frac{\partial f(x, y)}{\partial y} \right|_{X, Y}^2 \sigma^2(Y)}$$

The above result can be simply generalized to the case of a calculated quantity that depends on an arbitrary number of measured quantities.