Uncertainty in Measurement

Error analysis is the study and evaluation of uncertainty in measurement. In science the word “error” does not carry the usual connotations of “mistake”. “Error” in a scientific measurement means the inevitable uncertainty that attends all measurements. As such, errors are not mistakes; you cannot avoid them by being very careful. The best you can hope to do is to ensure that errors are as small as reasonably possible, and to have some reliable estimate of how large they are.

Errors in experimental measurements can be divided into two classes: (a) systematic errors and (b) random errors. It is possible to correct for errors of the first type if the source of the error is known. Random errors are indicated by fluctuation in successive measurements and lead to imprecise measurements. Systematic errors are reproduced in successive measurements, made under the same conditions.

Many systematic errors can be eliminated by the application of familiar corrections. For example, in the determination of atmospheric pressure using a mercury barometer, corrections must be applied to allow for the difference between the thermal expansion of mercury and the scale that is used to measure the mercury column height. This is required since the barometer only reads correctly if the mercury column and scale are at a temperature of 0 °C. In addition, it is necessary to correct for the capillary depression of the mercury and for the difference between the acceleration of gravity where the barometer is being used and the reference point (sea level, 45° latitude). In other cases where the theory has not been as well developed, it is necessary to determine corrections experimentally. A calibration of the scales of many instruments can best be obtained by making measurements on standard materials with well known properties. This procedure tends to eliminate systematic errors introduced by the instrument.

Systematic errors may not manifest themselves by fluctuations in measurements and cannot be eliminated by merely repeating the measurements. These errors are therefore especially serious and insidious, and can be avoided only by careful calibrations and consideration of all possible corrections. Sometimes systematic errors are indicated by the change in the measured value resulting from a change of experimental technique or when different values are obtained on different days.

Errors of the second type, random errors, are indicated by fluctuation in successive measurements. These random variations are due to a number of factors beyond the control of the observer. For example, if a barometer is read several times in succession, the values read from the instrument will be found to fluctuate about a mean value. Random errors are not necessarily of instrumental origin. There is sometimes an essential background noise superimposed on the signal being measured. If the fluctuations are in fact random, they can be treated by the methods of statistics.

To illustrate the distinction between random and systematic errors, let us consider some examples. Suppose first that we time a revolution of a steadily rotating turntable. One source of error will be our reaction time in starting and stopping the watch. If our reaction time was always exactly the same, these two delays would cancel one another. In practice, however, our reaction time will vary. We may delay in starting, and so underestimate the time; or we may delay in stopping, and overestimate the time. Since either possibility is equally likely, the sign of the effect is random. If we repeat the measurement several times, we will sometimes overestimate and sometimes underestimate. Thus our variable reaction time will show up as a variation of the answers found. By analyzing the spread in results statistically, we can get a very reliable estimate of this kind of error.

On the other hand, if our stopwatch is running consistently slow, then all our times will be underestimated, and no amount of repetition (with the same watch) will reveal this source of error. This kind of error is called systematic, because it always pushes our result in the same direction. Systematic errors cannot be discovered by the kind of statistical analysis we are contemplating here.

Estimation of Experimental Error

When it is not possible to repeat a measurement enough times for a statistical treatment, as is often the case in the physical chemistry laboratory, it is necessary to estimate the precision of a measurement. No fixed values can be given for the precision of various types of measurements because the precision depends upon the apparatus, the conditions under which it is used, and the care taken by the operator. Therefore it is necessary to develop an awareness of various sources of error in order to make reliable estimates.

In the measurement of weight in the laboratory, the precision may vary over a wide range. An ordinary analytical balance may be used to obtain weights to ±0.1 mg, but the precision will depend upon the sensitivity of the balance and the way in which it is used.

In the measurement of volume the precision will depend upon whether volumetric flasks, pipets, or burettes are used and on the size of the volume to be measured. The National Institute of Standards and Technology tolerances for volumetric equipment are given in textbooks on quantitative analysis.

The uncertainty in a measurement of temperature will be quite different if the temperature is measured with a good mercury-in-glass thermometer near room temperature or by use of a thermocouple at a high temperature. In calculating the percentage error in the temperature, it must be remembered that it is the uncertainty of the value which is used in the calculation that is significant. Thus an uncertainty of 1° at 25° C would cause not a 4% error in a calculation of molecular weight from the ideal-gal law, but a 0.3% error because the temperature used is in Kelvin not Celsius. In other types of experiments it is the change in temperature that is significant, rather than the absolute temperature, and so it is important to estimate the precision with which this difference can be measured.

Many measurements involve uncertainties that are much harder to estimate than those connected with locating points on a scale. For example, when we measure a time interval using a digital stopwatch, the main source of uncertainty is not the difficulty of reading the watch, but our own unknown reaction time in starting and stopping the watch. These kinds of uncertainty can sometimes be reliably estimated, if we can repeat the measurement several times. Suppose, for example, we time the flow time of a viscometer once and get an answer of 219.16 sec. From one measurement we can’t say much about the experimental uncertainty. But if we repeat and get a time of 220.25 sec, then we can immediately say that the uncertainty is probably of the order of 1 sec. If a sequence of four timings gives the results (in sec)

219.16

220.25

218.65

219.54

It is natural to assume that the best estimate of the flow time is the average, and it seems reasonable to assume that the correct flow time is somewhere between the lowest and the highest value.

best estimate = 219.40

probable range = 218.65 to 220.25

Later it will become apparent that the results should be rounded to 1 less significant figure and presented as

best estimate = 219.4

probable range = 218.6 to 220.2

The above form gives the best estimate and the range. A more compact form is more commonly used:

flow time = 219.40 ± 0.67 sec

In general, the result of any measurement of a quantity x is stated as

(measured value of x) = $x_{best} \pm \Delta x$

This statement means, first, that the experiment's best estimate for the quantity concerned is the number x_{best}, and, second, that he or she is reasonably confident the quantity lies somewhere between $x_{best} - \Delta x$ and $x_{best} + \Delta x$. The number Δx is called the uncertainty, or error, in the measurement of x. It is convenient to define the uncertainty Δx to be positive. Many scientists use a convention that the statement T = 26° without any qualification is presumed to mean that T is closer to 26° than to 25° or 27°; that is,

$T = 26°$

means

$25.5° \leq T \leq 26.5°$

In general you should not use this convention, but rather state the uncertainty with each value. However you should understand the convention, and know that it applies to any number stated without an uncertainty.

Influence of Experimental Errors on the Final Result

A final physical-chemical result is usually obtained by combining the results of different kinds of measurements. The accuracy of any final result is influenced by the accuracy of the measurements of the several quantities involved. It can happen that one of the measurements has

the preponderant effect in determining the accuracy of the final result. For example, the molecular weight of a sample can be determined by dissolving it and measuring the elevation of the boiling point of the solvent. The error in weighing the sample will contribute relatively little to the error in calculating molecular weight whereas error in determining the elevation of the solvent's boiling point is very important. How errors in the measurement of the independent variables (raw data) determine the error in the dependent variable is shown below. Just because a single measurement has a large absolute (or relative) error does not mean that this measurement controls the error in the final calculated result. Rather, an analysis of the propagation of errors is needed. Once the analysis indicates which measurement is most significant in determining the error in the dependent variable, improvement of that measurement will do the most towards improving the result.

The result of a calculation, called u, determined from a set of experimental values x,y,z, etc. constitutes a function dependent on these values (we will assume x,y,z, etc. are independent of each other). The differential change in the result u (called ∂ u) due to differential changes in the independent variables (∂ x, ∂ y, ∂ z) is given by the conventional expression for the differential of a function of several independent variables. Restricting the treatment to a basis of three independent variables (the extension to a larger number is obvious)

$$\partial u = \left(\frac{\partial u}{\partial x}\right)_{y,z} \cdot \partial x + \left(\frac{\partial u}{\partial y}\right)_{x,z} \cdot \partial y + \left(\frac{\partial u}{\partial z}\right)_{x,y} \cdot \partial z \qquad (1)$$

The partial derivatives are derived from the relation by means of which u itself is calculated. Equation (1) provides a simple basis for estimation of the possible range of uncertainty which must be assigned to the value calculated for u as a consequence of the known uncertainties in the experimental data. In this procedure, it is assumed that the accuracy of the measurements is reasonably good (of the order of a few percent or better), so that to an adequate degree of approximation Eq. (1) may be replaced by

$$\Delta u \approx \left|\left(\frac{\partial u}{\partial x}\right)_{y,z}\right| \cdot \Delta x + \left|\left(\frac{\partial u}{\partial y}\right)_{x,z}\right| \cdot \Delta y + \left|\left(\frac{\partial u}{\partial z}\right)_{x,y}\right| \cdot \Delta z \qquad (2)$$

In the above equation it is assumed that the errors Δx, Δy, Δz are random and independent of each other. Here Δu approximates the finite change u(x+Δx,y+Δy,z+Δz) - u(x,y,z) in the

calculated value of u, which results from the changes Δx, Δy, Δz in the independent variables away from the values x, y, z, for which u and the partial derivatives of u are evaluated. The absolute value is needed because the error contribution from one factor cannot cancel the contribution due to another factor.

Special Simple Cases

From this result, several special cases can be derived. These are also presented without proof. You could derive them yourself, if you were so inclined.

$u = x \pm y$ $\qquad \Delta u = \Delta x + \Delta y$ (3)

$u = x \cdot y$ $\qquad \frac{\Delta u}{|u|} = \frac{\Delta x}{|x|} \cdot + \frac{\Delta y}{|y|}$ (4)

$u = \frac{x}{y}$ $\qquad \frac{\Delta u}{|u|} = \frac{\Delta x}{|x|} \cdot + \frac{\Delta y}{|y|}$ (5)

$u = a \cdot x$ $\qquad \Delta u = |a| \cdot \Delta x$ (6)

where a has no error (it's a constant like π)

$u = x^n$ $\qquad \frac{\Delta u}{|u|} = |n| \frac{\Delta x}{|x|}$ (7)

The above equations make an assumption that all of the contributions to the error in u will be in the same direction. These equations give upper limits to the uncertainties. In real experiments with independent and random uncertainties in x, y, and z, some of the errors may have opposite signs and partially cancel. The equations can be modified to reflect this case. They are again presented without proof.

$$\partial u = \sqrt{\left(\frac{\partial u}{\partial x}\right)^2_{y,z} \cdot \partial x^2 + \left(\frac{\partial u}{\partial y}\right)^2_{x,z} \cdot \partial y^2 + \left(\frac{\partial u}{\partial z}\right)^2_{x,y} \cdot \partial z^2} \qquad (8)$$

The above specific examples are thus:

$u = x \pm y$ $\qquad \Delta u = \sqrt{(\Delta x)^2 + (\Delta y)^2}$ (9)

$u = x \cdot y$ $$\frac{\Delta u}{|u|} = \sqrt{\left(\frac{\Delta x}{|x|}\right)^2 + \left(\frac{\Delta y}{|y|}\right)^2} \quad (10)$$

$u = \frac{x}{y}$ $$\frac{\Delta u}{|u|} = \sqrt{\left(\frac{\Delta x}{|x|}\right)^2 + \left(\frac{\Delta y}{|y|}\right)^2} \quad (11)$$

$u = a \cdot x$ $\Delta u = |a| \cdot \Delta x$ (12)

where a has no error as before and

$u = x^n$ $$\frac{\Delta u}{|u|} = |n| \frac{\Delta x}{|x|} \quad (13)$$

Significant Figures

An excessive number of uncertain figures should not be retained. After the uncertainty estimate has been obtained, the value reported for the quantity should be rounded off so that it contains not more than one or two uncertain significant figures. If the figures to be dropped amount to more than half of the last figure retained, that figure should be rounded off upwards; if less than half, downwards. When the portion discarded is exactly one-half, it is customary to round off to the nearest even value.

Several basic rules for stating uncertainties are worth emphasizing. First, since the quantity Δx is an estimate of an uncertainty, it should obviously not be stated with too much precision. If we measure the acceleration of gravity g, it would be absurd to state a result like:

$$\text{measured } g = 9.82 \pm 0.02385 \text{ m/sec}^2.$$

It is inconceivable that the uncertainty in the measurement can be known to four significant figures. In high-precision work, uncertainties are sometimes stated with two significant figures, but for most cases we can state the following rule:

Rule for Stating Uncertainties

Experimental uncertainties should usually be rounded to one significant figure.

Thus, if some calculation yields the uncertainty $\Delta g = 0.02385$ m/sec^2, this answer should be rounded to $\Delta g = 0.02$ m/sec^2, and the conclusion should be rewritten as:

measured g : 9.82 ± 0.02 m/sec^2

There is only one significant exception to this rule. If the leading digit in the uncertainty Δx is a 1, then it may be better to keep two significant figures in Δx.

Once the uncertainty in a measurement has been estimated, one must also consider which are the significant figures in the measured value. A statement like:

measured speed = 6051.78 ± 30 m/sec

is obviously ridiculous. The uncertainty of 30 means that the digit 5 in the third place of 6051.78 might really be as small as 2 or as large as 8. Clearly the trailing digits 1, 7 and 8 have no significance at all, and should be rounded off. The general rule is

Rule for Stating Answers

The last significant figure in any stated answer should always be of the same order of magnitude (in the same decimal position) as the uncertainty.

For example, the answer 92.81 with an uncertainty of 0.3 should be rounded as:

92.8 ± 0.3

If its uncertainty is 3, then the same answer should be rounded as:

93 ± 3

and if the uncertainty is 30, then the answer should be:

90 ± 30

However, numbers to be used in calculations should generally be kept with one more significant figure that is finally justified. This will reduce the inaccuracies introduced by rounding the numbers. At the end of the calculation, the final answer should be rounded to remove this extra (and insignificant) figure. There is one more exception to the rules. If the leading digit in the uncertainty is small (a 1 or perhaps a 2), then it may be appropriate to retain one extra figure in the final answer. For example, an answer such as:

measured length = 27.6 ± 1.2 cm

is quite acceptable Note that the uncertainty in any measured quantity has the same dimensions as the measured quantity itself. It is therefore clearest and more economical to write the units after both the answer and the uncertainty. By the same token if a measured number is so large or small that it calls for scientific notation, then it is simpler and clearer to put the answer and uncertainty in the same form. For example:

$q = (1.61 \pm 0.05) \times 10^{-19}$coulombs

is much easier to read and understand that it would be in:

$q = 1.61 \times 10^{-19}$coulombs $\pm 5 \times 10^{-21}$coulombs

Example:

As an example, suppose that we measure g, the acceleration of gravity, using a simple pendulum. The period of such a pendulum is well-known to be $T = 2\pi\sqrt{L/g}$ where L is the length of the pendulum. Thus if L and T are measured, we can find g as

$$g = \frac{4\pi^2 \cdot L}{T^2} \tag{14}$$

This gives g as the product or quotient of three factors, $4\pi^2$, L, and T^2. If the various uncertainties are independent and random, the fractional uncertainty in our answer is just the quadratic sum of the fractional uncertainties in these factors. The factor $4\pi^2$ has no uncertainty, and the fractional uncertainty in T^2 is twice that in T;

$$\frac{\Delta\left(T^2\right)}{T} = 2 \cdot \frac{\Delta T}{T} \tag{15}$$

Thus the fractional uncertainty in our answer for g will be

$$\frac{\Delta g}{g} = \sqrt{\left(\frac{\Delta L}{L}\right)^2 + \left(2 \cdot \frac{\Delta T}{T}\right)^2} \tag{16}$$

Suppose we measure the period T for one value of the length L and get the result

$$L = 92.95 \pm 0.12 \text{ cm}$$

$$T = 1.936 \pm 0.004 \text{ sec}$$

Our best estimate for g is found from eq (14) as

$$g_{best} = \frac{4 \cdot \pi \times (92.95 cm)}{1.936 \cdot \sec^2} = 979 \cdot cm/\sec^2$$

To find our uncertainty in using eq (16), we need the fractional uncertainties in L and T. They are easily calculated as: $\frac{\Delta L}{L} = 0.001$ and $\frac{\Delta T}{T} = 0.002$

Substituting into eq (16), we find:

$$\frac{\Delta g}{g} = \sqrt{(.001)^2 + (0.004)^2} \approx 0.004$$

Therefore $\Delta g = (0.004) \times (979\ \text{cm/sec}^2) = 4\ \text{cm/sec}^2$

Thus our final answer, based on these measurements, is: $g = 979 \pm 4\ \text{cm/sec}^2$

Statistical Estimation of Error

The treatment of random errors is quite different from that of systematic errors. The statistical methods described in the following sections give a reliable estimate of the random uncertainties, and, as we shall see, provide a well-defined procedure for reducing them. On the other hand, systematic uncertainties are hard to evaluate and detect; the experienced scientist learns to anticipate the possible sources of systematic error, and to make sure that all systematic errors are much less than the required precision. Doing so will involve, for example, checking the meters against accepted standards, and correcting them or buying better ones if necessary.

Suppose we need to measure some quantity x, and have only random errors to worry about. We should be able to detect the random errors by repeating the measurement several times. We might, for example, make the measurement five times and find the results:

71, 72, 72, 73, 71

The first question is: Given the five measured values, what should we take for our best estimate x_{best} of the quantity x? It seems reasonable that our best estimate would be the average or mean x of the five values found: $x_{best} = \bar{x} = (71 + 72 + 72 + 73 + 71)/5 = 71.8$

where we have used the definition:

$$\bar{x} = \frac{\sum x_i}{N} \tag{17}$$

The standard deviation of x (written σ_x) characterizes the uncertainty of the individual measurements. σ_x is given by the formula:

$$\sigma_X = \sqrt{\frac{\sum (x - \bar{x})^2}{N - 1}} \tag{18}$$

The mean of $\bar{x}$ is a judicious combination of all N measurements, and there is every reason to think it will be more reliable than any one of the measurements considered separately. The uncertainty in the final answer $x_{best} = \bar{x}$ turns out to be the standard deviation σ_X divided by $\sqrt{N}$. This quantity is called the standard deviation of the mean (written $\sigma_{\bar{x}}$) or standard error.

Our final result can then be written

$$\text{value of x} : x_{best} \pm \Delta \bar{x}$$

where $x_{best} = \bar{x}$ and $\Delta \bar{x}$ is the standard deviation of the mean as defined from:

$$\Delta \bar{x} = \Delta \sigma_{\bar{X}} = \frac{\sigma_x}{\sqrt{N}}$$

Errors from Graphs

Frequently, you will be reporting a physical value which is itself calculated as the slope or intercept of a best (or linear least squares) fit to your data. How does one then estimate the error of this measurement given that there is no cut-and-dry formula for these quantities? Although the linear least squares formulas are somewhat complex, mathematicians have waded through all the formulas and derived equations for the errors in the slope and intercept of a linear least squares line fit. An example for this calculation is provided in Excel spreadsheet format on the class web site. Further, these formulas were obtained from:

http://mathworld.wolfram.com/LeastSquaresFitting.html

Note that we would never expect you to try to memorize these formulas; however, you do have to make sure you know how to properly use the tools we provide. Are you sure you can alter the Excel spreadsheet if you have more or less data than in the example? Can you translate the formula into MATLAB format if that is your chosen program for analyzing data?

References

1. John R. Taylor, "An Introduction to Error Analysis," University Science Books, Mill Valley, California, 1982.

2. Wilfrid J. Dixon and Frank J. Massey, Jr., "Introduction to Statistical Analysis," 3rd edition, McGraw Hill Book Company, New York 1969.

3. Philip R. Bevington, “Data Reduction and Error Analysis for the Physical Sciences,” McGraw Hill Book Company, New York, 1969.
4. Cuthbert Daniel and Fred S. Wood, “Fitting Equations to Data,” Wiley-Interscience, New York, 1971.