APPENDIX A 3 September 2023

ACCURACY, PRECISION, ERRORS, UNCERTAINTY, ETC.

Part of making and reporting a measurement is deciding how accurate it might be. Finding that a distance is 10.0000 cm ± .0001 cm could tell you something very different from 10 cm ± 1 cm. In common speech, the words accuracy and precision are often used interchangeably. However, many scientists like to make a distinction between the meanings of the two words. Accuracy refers to the relationship between a measured quantity and the real value of that quantity. The accuracy of a single measurement can be defined as the difference between the measured value and the true value of the quantity. Since in most cases you don't know the true value (if you did, you wouldn't be bothering to measure it!), you seldom know the true accuracy of your answer. Exceptions to this occur primarily when you are testing an apparatus or new measurement method, and in teaching labs like this one. Since here we often do know the true value, or have measured the same quantity two different ways, whenever you have this opportunity you should always compare the achieved ***accuracy*** (the difference between the true value and your measured value) with your independently estimated error as described below. If the difference is comparable to your estimated uncertainty, then your measurement is ***consistent*** with the accepted value. If it is five time larger, you owe some discussion in your notebook of why this might be the case.

More vocabulary: The word error in standard English is the same as mistake. In science however, it means the estimated ***uncertainty***, or ***probable accuracy*** of a result (i.e. the amount by which the measured value is expected to differ from the true value). The point is that in most experiments we do not know the true value of the quantity we are measuring, and therefore cannot determine the actual error in our result. However, it is still possible to make an estimate of the uncertainty (or the "probable error") in the measurement based on what we know about the properties of the measuring, instruments, etc.

Precision: The maximum possible accuracy of a measurement. If you measure a 1 cm long object with a ruler, you won't be able to determine the length to better than a few tenths of a mm. If you measure it with a micrometer, you can get a value that is *precise* to a couple of μm. If you copy down a 6-digit voltmeter display in your notebook with only three significant figures (very often the right thing to do!), its precision is +/- 1 digit in the third place. There are many things that will make a number less accurate than its precision, but it can never be more accurate.

Measurement uncertainties can be divided into two distinct classes: ***random*** or ***statistical*** errors, and ***systematic*** errors. Systematic errors are things like the accuracy of a voltmeter calibration or perhaps that you made all your length measurements with a metal tape measure that had expanded because you are in a much warmer room than the one where the tape was constructed. Systematic errors can be quite difficult to estimate, since you have to understand everything about how your measurement system works.

Somewhat counter-intuitively, the random error is usually easier to estimate. It is due to some combination of the limited precision to which a quantity can be read from a ruler or meter scale, and intrinsic "noise" on the measurement. For example, if a radioactive source that gives an average of one count per second is counted for exactly 100 seconds, you will find that you don't always get exactly 100 counts even if your count perfectly accurately (no ***mistakes***). About one-third of the time, you will get fewer than 90 or more than 110 counts, and occasionally (about 0.5% of the time) you will get fewer than 70 or more than 130 counts. If you make a plot of the distribution of a large

number of 100-second counts, you will get a curve called a "Poisson distribution." Unless the number of counts is very small, this curve will be very close to a gaussian or "bell curve". Most random errors follow this kind of distribution. The expected size of the uncertainty in a measurement is described by the width of this curve. The limits that contain 2/3 of the measurements (±10 in our example) are called the "1-sigma" uncertainty. If the errors follow the bell curve, then 95% of the results will be within ±2σ, and 99.5% within ±3σ.

You can often estimate the random error in a measurement empirically. If you can make a few *independent* measurements of some quantity, you can obtain an estimate of the precision of each individual measurement. (The "independent" part is important: if you measure a length with a meter stick, and on the first try estimate 113.3 mm, you are likely to write down 113.3 on subsequent measurements as well, even if you can really only estimate to ±0.1 or 0.2 mm. One way around this is to have different people make each measurement, and write them down without looking at each other's answers. Or by yourself, you could start from a random point on the ruler each time and estimate the readings at both ends, then do all the subtractions afterwards.)

The following example illustrates several of these ideas. In this example the resistance of a known 1000 ± 0.01 Ω resistor is determined by measuring V and I for several different voltage settings. The results are given in the table on the following page. The average value of R in this example is 1002.4 Ω, so our final result has an accuracy of 2.4 Ω. The expected error of any individual measurement of R can be determined by calculating the standard deviation of the distribution:

$$\sigma \equiv \frac{1}{\sqrt{N-1}}\left[\sum_{i=1}^{N}(x_i-\bar{x})^2\right]^{1/2}$$

where $\bar{x}$ is the average value of x and where N is the number of measurements. In this example the standard deviation is 5.6 Ω. We could take this as an estimate of the uncertainty or probable error, since any individual measurement has a reasonable probability of being in error by at least that amount. It should be emphasized, however, that the actual error in a measurement can be much larger than the standard deviation if there are systematic errors (for example errors in the calibration of some meter) that affect all the measurements the same way.

Data for a 1000 ± 0.01 kΩ Resistor

V[a] (volts)	I[b] (mA)	R[c] (Ω)
1.000	0.99	1010
2.000	1.99	1005
3.000	3.00	1000
4.000	4.02	995
5.000	4.99	1002

Average = 1002.4 ± 2.5 Ω (usually better: 1002 ± 3Ω)
"Mean Standard Deviation" = 5.6 Ω (expected accuracy of a single measurement)
"Standard Deviation of the Mean" = 2.5 Ω (expected accuracy of average of 5 measurements)
Actual accuracy = 2.4 Ω

[a]Measured with digital voltmeter.
[b]Measured with Simpson VOM.
[c]Calculated from R = V/I.

Propagation of Errors

In many experiments, our desired result Q is determined from a mathematical formula that uses two or more separately measured quantities: $Q = f(x_i, \ldots x_n)$, where $x_1, \ldots x_n$ are measured values, and f is the mathematical function. If each of the x_i were to change by an amount δx_i, then to first order Q will change by

$$\Delta Q = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i} \Delta x_i. \qquad (1)$$

We have estimated the uncertainties in the n measured quantities, and want to calculate the uncertainty in Q. We know the expected magnitude of Δx_i, but expect it is equally likely to be positive or negative, so its average value would be zero. We usually try to estimate (or assume we know) the quantity $\sigma_{x_i} = \left\langle \Delta x_i^2 \right\rangle^{1/2}$, or the square root of the average of Δx_i^2 (the "root mean square" or r.m.s. value of the expected error). The problem of figuring out the uncertainty in the result, given the formula and the uncertainties in the numbers going into it, is called "error propagation."

Since $\langle \Delta Q \rangle$, the average or "expected" value of ΔQ is also zero, we need to calculate the expected value of ΔQ^2:

$$\left\langle \Delta Q^2 \right\rangle = \left\langle \left(\sum_{i=1}^{n} \frac{\partial f}{\partial x_i} \Delta x_i \right)^2 \right\rangle. \qquad (2)$$

Taking the square will produce terms of the form $coeff \cdot \left(\Delta x_i \Delta x_j \right)^2$. For $i \neq j$ we generally assume the expected value is zero, since if Δx_i is positive, Δx_j should be equally likely to be positive or negative. This assumes that the errors in Δx_i and Δx_j are ***independent***. If this is not true, you must keep these terms! The expected value of the terms with $i = j$ are $\left\langle \Delta x_i^2 \right\rangle$, which is just $\sigma_{x_i}{}^2$.

There are just two cases that in combination will cover 98% of the error propagation problems you will run into. We give these results here with the recommendation that you memorize them, although all can easily be derived from equation (2) above. A and B are two measured (or calculated) values:

For $Q = A + B$ **or** $Q = A - B$: $\sigma_Q \equiv \left\langle \Delta Q^2 \right\rangle^{1/2} = \sqrt{\sigma_A^2 + \sigma_B^2}$. (Note errors <u>add</u> even for $Q = A - B$.)

For $Q = A \cdot B$ or $Q = A/B$: $\dfrac{\sigma_Q}{Q} = \sqrt{\left(\dfrac{\sigma_A}{A} \right)^2 + \left(\dfrac{\sigma_B}{B} \right)^2}$.

For independent errors, it makes some sense that the errors add as the square root of the sum of their squares, or "in quadrature": the errors might add, or they might have opposite signs, and at least partially cancel. So on the average, we could expect them to add "at right angles". Beyond that, the two formulae above are easily remembered as "for addition or subtraction, add absolute errors, for multiplication or division, add percentage errors." (Where "add" means "add in quadrature".) One other occasionally useful result is for $Q = A^n$, $\dfrac{\sigma_Q}{Q} = n \dfrac{\sigma_A}{A}$. Note that in the case of the resistor measurements tabulated above, we made several measurements to estimate the error in a single measurement by looking at the repeatability, or scatter. But then we might as well report the average of these measurements, since propagating the individual error estimate through the simple formula for calculating an average shows that the uncertainty is reduced by a factor of sqrt(N). The technical terms for these quantities are *mean standard deviation* and *standard deviation of the mean*.

For a mixed case like $Q = (A + B)/(C + D)$, you first add absolute errors for the numerator and denominator, then convert these to % errors, and add them to get the error in Q. So error propagation becomes largely an exercise in converting back and forth from absolute to percentages.
You do have to be careful of correlated errors. This actually happens most often when it's really the same quantity that shows up in more than one place. In that case, the errors are perfectly correlated. Take the case of $Q = 2A$, which could also be written $Q = A + A$. If you use the formulae given above for independent errors, you'll get different answers for σ_Q! If you use equation (2) and keep the cross term, they'll come out the same. A more subtle example comes up in experiment 5: R_4 in the Wheatstone Bridge consists of R_E (0.5% accuracy) in parallel with R_H (10% accuracy), so that:

$$R_4 = \frac{R_E R_H}{R_E + R_H}. \tag{3}$$

You can calculate the errors in the numerator and denominator separately using the independent error formulae, but then you can't combine them with the independent errors formula because they contain the same variables, so these errors aren't independent.

To do such cases exactly, it's usually easiest (and always safest) to go back to equation (2), which gives:

$$\frac{\sigma_{R_4}}{R_4} = \left(\left(\frac{R_4}{R_E} \right)^2 \left(\frac{\sigma_{R_E}}{R_E} \right)^2 + \left(\frac{R_4}{R_H} \right)^2 \left(\frac{\sigma_{R_H}}{R_H} \right)^2 \right)^{1/2}. \tag{4}$$

But you can often save a huge amount of effort by looking at the magnitude of the numbers and making approximations. In this case, typically R_E = 1 kΩ and R_H = 200 kΩ. Although the fractional error in R_H is large, its error doesn't contribute much to the total error in R_4 since it is multiplied by the square of a factor ($R_4/R_H \approx 1/200$) that is small compared to 1. You could have told this without bothering to derive equation (4): looking at equation (3). $R_H >> R_E$, so the denominator $\approx R_H$, and this will approximately cancel the R_H in the numerator. So $R_4 \approx R_E$ and has the same uncertainty as R_E, or 0.5%. This you can all do in your head! For measurements where it's clear that the uncertainty will not contribute significantly to the error in the final result (usually true for measurements you make with a six-digit multimeter, for instance), you don't need to write down any ± for the number.

****

It is usually not required to make detailed calculations of the uncertainties in your measurements, and the calculations are often long and time consuming to do exactly. But it is always important for experimenters to have an approximate idea of the uncertainties in their results. With suitable approximations, by ignoring variables that make insignificant contributions and using the two simple independent-error results, you can do most of the error estimation in your head and just put down ± σ, where this is a rough estimate of what you expect to be within ~2/3 of the time. *Note that you seldom need to actually calculate square roots of sums of squares: you can just use the larger number as the result if they are significantly different, or half again the value if they are about equal.* The usual convention is to convert x to the same units as the result and give the absolute error. One significant figure is usually quite adequate for errors, and in some cases just being sure to round your result to an appropriate number of significant figures is enough and you don't bother to write down the ± σ.

More discussion of errors and detailed derivations can be found in *Data Reduction & Error Analysis for the Physical Sciences*, 3rd Edition by Bevington & Robinson (QA278 B48 2003), chapters 1-5.