Technical Intro to Error Propagation

Uncertainties in Measured Values: What do we mean by "Error"?

When getting quantitative information from a measurement, we are interested not just in the value we obtain, but in how sure we are that the value we have measured is correct. There are many factors that can produce a shift or an uncertainty in a measurement – we could have a meter stick with the end chipped off, our dials can only be read to a certain number of significant figures so the next digit is uncertain, or the setup conditions for our experiment can't be arranged precisely. In standard terminology these are referred to as "errors" – though this is a technical term that really means "uncertainties." There is no implication that there are any mistakes made in doing the experiment! Once we have a good estimate of how much uncertainty there is in our measurement, we estimate an *error bar* – a spread of values that says, "We expect the odds are 2:1 that the actual value is inside this range."

Errors like our meter stick being chipped off and thus too short are called *systematic errors*. They always shift the result in one specific direction and need special care to reduce them. Errors that arise from many small hard to control uncertainties (e.g. how well two fluids are mixed, how stable the temperature of the apparatus is, or whether the measurement is affected by building vibrations) are well studied mathematically and are referred to as *random error*, and these errors make the result bigger OR smaller in a RANDOM fashion. One way to get a handle on this random error is to repeat the experiment a number of times, preparing it as similarly as you can, and see how much variation there is. The statistical tools of *mean* (*average*) and *standard deviation* allow you to estimate both the average result and the error bar arising from random error.

Propagation of Error:

Often the measurement that we make is not the final answer we want. We may have to take a measured value as input in a calculation and do calculations with it. If there is uncertainty in the input numbers for our calculation, then there will be uncertainty in the output numbers as well – but they won't be the same uncertainty. The input and output numbers likely even have different units!

To figure out how an input uncertainty translates into an output uncertainty, we simply have to ask: if the input value changed, how would that affect the output? We can answer that question by doing the calculation – changing the input value a little and calculating the changes in output. But we can also use calculus: the derivative of a function (an output calculated from some input) tells you how that output changes if the input changes a little! Mathematically these two options are written as follows:

(1) $$\delta f = f(x+\delta x) - f(x)$$

(2) $$f(x+\delta x) = f(x) + \delta x \frac{df}{dx}$$ Therefore we can write: $$\delta f = \delta x \frac{df}{dx}$$

We use the lower case delta (δ) to represent a small change, since some of our inputs are changes already.

From Gen. Chem. II, you may have learned a form of error propagation based on fractional error:

$S_{a,rel} = \sqrt{(S_{c,rel})^2 + (S_{d,rel})^2}$ *with* $S_{x,rel} = \frac{S_x}{x}$ *and* S_x *the uncertainty in* x.

In these labs, we do<u>not</u> use this fractional error approach since this equation only holds under a few special circumstances. Instead, we are going to use a more sophisticated approach to error analysis based on your knowledge of calculus.

* General Form for Error Propagation:

Let δx be the known uncertainty in x and δy be the known uncertainty in y. A function of x and y, such as *f*(x,y), will have two parts to its uncertainty—one contribution from x information and another contribution from y information. The contributions to the uncertainty of *f* are found using the relationships below:

$$\delta f_x = \frac{df}{dx} * \delta x \text{ and } \delta f_y = \frac{df}{dy} * \delta y.$$

The total uncertainty in f, δf, can be found by using the relationship:

$$\delta f = \sqrt{(\delta f_x)^2 + (\delta f_y)^2}.$$

* Example 1: Average Velocity, <v>=Δx/Δt

If you know the uncertainty in Δx, δ(Δx), and the uncertainty in Δt, δ(Δt), then you can use the formula for average velocity, <v>=Δx/Δt, to find the uncertainty in the average velocity, δ(<v>).

$$\delta\langle v\rangle_{\Delta x} = \frac{d\langle v\rangle}{d(\Delta x)} * \delta(\Delta x) \quad \rightarrow \text{compute derivative} \rightarrow \quad \delta\langle v\rangle_{\Delta x} = \frac{1}{\Delta t} * \delta(\Delta x)$$

$$\delta\langle v\rangle_{\Delta t} = \frac{d\langle v\rangle}{d(\Delta t)} * \delta(\Delta t) \quad \rightarrow \text{compute derivative} \rightarrow \quad \delta\langle v\rangle_{\Delta t} = \frac{-\Delta x}{(\Delta t)^2} * \delta(\Delta t)$$

$$\delta(\langle v\rangle) = \sqrt{(\delta\langle v\rangle_{\Delta x})^2 + (\delta\langle v\rangle_{\Delta t})^2} \quad \rightarrow \quad \delta(\langle v\rangle) = \sqrt{(\frac{1}{\Delta t} * \delta(\Delta x))^2 + (\frac{-\Delta x}{(\Delta t)^2} * \delta(\Delta t))^2}$$

* Example 2: A sinusoidal function

Given a sinusoidal function, $a = \sin(\omega t) + 2$,and uncertainties in ω, δω, and in t, δt, the uncertainty in a, δa, can be calculated as follows:

$$\delta a_\omega = \frac{da}{d\omega} * \delta\omega \quad \rightarrow \text{compute derivative} \rightarrow \quad \delta a_\omega = t\cos(\omega t) * \delta\omega$$

$$\delta a_t = \frac{da}{dt} * \delta t \quad \rightarrow \text{compute derivative} \rightarrow \quad \delta a_t = \omega\cos(\omega t) * \delta t$$

$$\delta(a) = \sqrt{(\delta a_\omega)^2 + (\delta a_t)^2} \quad \rightarrow \quad \delta(a) = \sqrt{(tcos(\omega t) * \delta\omega)^2 + (\omega\cos(\omega t) * \delta t)^2}$$

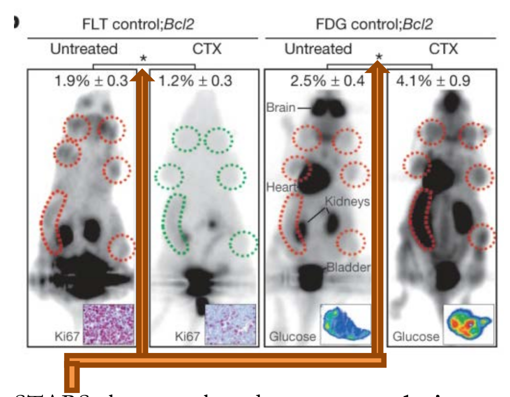
* So I can propagate error, but why do I care?

We often hear students express the following frustrations: "What's the big deal with all of this error analysis, anyway? Is it just busy work? Is there a more significant, scientific purpose? Why are there so many ways to determine error??!?" **What do you think?**

Here's what we think:

Error analysis is key to science and medicine. The key question in medical research and medical practice is, whether a treatment works! Does a new drug make a patient feel better or remove their disease better than another treatment? To answer this question requires comparing observations before and after, or comparing treated patients with so called "controls" in clinical trials. But how can we tell whether something is the same or different? This is where error analysis comes in as a crucial stepping stone!

We randomly chose an article on a medical topic (cancer therapy) from a recent issue of the prestigious journal Nature. In all of the four figures included in the article, we saw that the authors, in trying to show that their therapy works, compared mice that were treated with their therapy to "controls," i.e. untreated mice as shown in the sample image on the right. To highlight a statistically significant difference it is customary to put a "star" in the figure. In this figure you see two stars showing that the two images on the left are different, and the two images on the right are different. But how would you determine that the two x-ray images are different? Use error analysis and error propagation! In this example the input data are x-ray images, which have uncertainty from mouse to mouse and from x-ray exposure to x-ray exposure. The output, which is what the authors want to compare, are "tumor to background ratios."



STARS above are based on **error analysis.** They indicate that the differences (in this case between the behavior of treated and untreated mice) are significant. From JR Dörr et al. *"Synthetic lethal metabolic targeting of cellular senescence in cancer therapy"* Nature (2013).

Our take home message: Error analysis is the hidden backbone of scientific research. It may only show up as small stars in an otherwise glossy image, but without that star the authors could not draw any conclusion from these images other than "each mouse has a somewhat different number of tumors." Overall we counted 38 "stars" in the article that analyzed the effectiveness of one particular therapy! **We want you to become Stars of Error analysis! ☺**

There ARE a lot of ways to do error analysis. Part of what you are learning to do, as budding scientists and doctors, is to find a way to choose the method of error analysis that best matches the experimental design/protocol. There is no single 'formula' for error analysis, just as there is no single 'formula' for doing science! Here are some error analysis methods that you will encounter in this class:

- determine uncertainties in individual measurements and propagate error to find the output error (as described in this document); or,
- calculate the output for each input from the multiple trials and use the standard deviation of the output to establish the uncertainty of the output directly.
- (There ARE other methods, but these are the most commonly used....)