Unit I

Error Propagation: an Introduction to the Monte Carlo Method

Unit I

Error Propagation: an Introduction to the Monte Carlo Method

This activity guide borrows a trick learned from Prof. William Titus of Carleton, that is, the use of random numbers to help one do error analysis. In fact, this is *not* how error analysis is done in the real world; one rather uses calculus and the chain rule. Prof. Titus designed this technique for use with students who felt unsure about their calculus skills, and could develop some intuition for how errors propagated from the uncertainties in the original measurements through a calculation to a final, more useful result. We easily have the calculus skill to understand the traditional methods of error propagation, and rather will primarily use this error example as a way of letting us understand the use of random numbers in computer simulations. The nice feature of looking at the error case is that it is easy to understand, and easy to simulate using a spreadsheet. Later in the course we will return to the Monte Carlo method, and you will do some simple programming, the more traditional vehicle for Monte Carlo methods. If you feel uncomfortable with either the Macintosh, or with Excel, at this point you should take the time to run through the Macintosh Basics program, or the Excel introduction--ask me about either of these and I will get you started.

Here is the basic problem. Imagine that we have are trying to measure the ratio of the two fundamental constants e (electron charge) and m (electron mass). The classic way of doing this is to get the electron moving by accelerating it through an electric potential difference of V, and then let it move in a circle in a magnetic field B. Many of you have done this experiment in Modern Physics. The actual measurements you do are of B (or perhaps the current through some coils), V, and the radius r of the resulting electron orbit. There are uncertainties associated with each one of these measurements. What sort of error can one quote in the final result for the ratio e/m? How does this depend on the size of the individual errors in B, V, and r?

Guidebook Entry I.1: Finite Differences and Calculus

First, convince yourself that the correct formula for the ratio e/m is in fact

$$\frac{e}{m} = \frac{2V}{r^2 B^2}.$$

Assume that the electron is very non-relativistic, and moving perpendicular to B at all times. You will need to recall a few intro physics formulae:

$$\vec{F}_B = q\vec{v} \times \vec{B}$$

$$a_{centripetal} = v^2/r$$

$$U_{electric} = qV$$

$$E_{total} = U + \tfrac{1}{2} mv^2$$

Let's now put some real numbers into action. Guess a reasonable V and B you might achieve in the laboratory. Look up e and m, and calculate (*to at least 3 significant figures)* what r must be. Use these values now as "ideal" measurements.

Make an Excel spreadsheet with columns for V, r, B and e/m. Enter your ideal values for V, r, and B in the first row, and the formula for e/m. See that you get the right number.

In the next row down, copy the same values and formula, except increase V by 1% over its original value. By what percentage does e/m change? What direction does e/m change? Repeat this for r and B, and fill out the result chart below.

	e/m percent change	direction of change
V:		
r:		
B:		

The traditional method of error propagation relies on calculus. That is, we approximate the finite variations of real errors by the infinitesimal ones of calculus. To write this in equation form, if we have an uncertainty in some measured value x which is Δx, a mathematical function of that variable, say f(x) ends up having some uncertainty of its own, Δf, given by

$$\Delta f \approx \frac{df}{dx}\,\Delta x\ .$$

This should be clear from the definition of the derivative; if it is not, or if you are the least unsure, ask your instructor for help.

Use this expression to find the absolute uncertainty in e/m as a function of an uncertainty in V.

We, however, looked at *relative* errors, that is things like Δf/f, not just Δf. Translate your result above to a relative error in e/m as a function of a relative error in V.

Does this agree with what you found when you did a 1% variation in V? If not, check with your instructor.

Now repeat this for relative errors in r and B. Do your results agree with what you found by 1% variations? Be sure to look at direction of changes, and whether the variations are close to, or exactly the same as your predictions.

What would happen if you made 1% changes in *each* variable at the *same time* in a way that would decrease e/m? To be more specific, you had decreased V, and increased r, and B, each one by 1%? You may want to check by making new entries in your spreadsheet.

Now you have a reasonable idea of how an uncertainty or error in a measured value affects, or is propagated to, a final, more interesting number. What we haven't done yet is see what happens if several of the original numbers are varying all at the same time. Before we can do this, we need to understand a bit better what we mean by uncertainty.

Guidebook Entry I.2: What is an Uncertainty Anyway?

We often refer to uncertainties as "errors." That is, it is the difference between the actual result, and the one we got. However, in the real world, we often don't know the real value; that is why we bothered to make the measurement in the first place! In that case, we need to know how much trust we can place in that number. (If we knew the actual error, we could just subtract that off, and get the exact value!) Given that we don't know the actual value, how can we possibly determine an uncertainty?

We need first to distinguish three types of "errors." The first is truly an error, and not an uncertainty. That is, we simply blundered--read the ruler wrong, connected the meter to the wrong terminals, looked at the wrong scale, or what have you. This is sometimes politely referred to as "human error;" in fact, it is just a screw-up. The second type of error is called systematic, and comes from defective measuring instruments that give wrong results of a repeatable type. For example, if our meter stick is actually only 0.95 m long, it will always give us length values that are about 5% too big. Similar problems can occur from miscalibrated meters, parallax in reading a scale, a timer that runs too fast or too slow....

The third type of error is the one we are most concerned with--random, or statistical errors. These are errors that occur in a way that varies from one measurement to the next in a way we at least appear unable to control. No matter how hard we try, we simply cannot use a stopwatch to time the fall of an object to an accuracy of a millisecond--our results will fluctuate about some average value.

This provides us a way of measuring an uncertainty: repeat the measurement a number of times. Our best estimate of the correct value is the average measurement, and the uncertainty is approximated by the spread of the values about the average. This spread is most commonly measured by the standard deviation. The standard deviation is chosen because it has nice properties when the values are distributed according to a Gaussian distribution ("bell-shaped curve"), and also because the propagation techniques turn out to be simple. The formula for the standard deviation of N different y values is

$$\text{Standard Deviation} = \sqrt{\frac{\sum \left(y - y_{average}\right)^2}{N-1}}$$

Excel allows you to generate what are known as random numbers (or more properly pseudo-random numbers). Entering the function RAND() in an equation introduces a random number at that point that lies somewhere between 0 and 1. All values over that range are equally likely. A new number is calculated each time the sheet is recalculated (which means every time a cell is altered, in general, or it can be forced by a command-= keystroke combination).

What should be the average value of a large set of random numbers distributed evenly over the range from 0 to 1?

Check your prediction by finding the average (you may use the AVERAGE(range) function) of 100 random numbers in an Excel sheet. Did you get what you expected? Try another set or two of 100 by recalculating.

What do you predict should be the standard deviation of those 100 random numbers if they are evenly distributed? Be as quantitative as you can be in your answer.

Use the STDEV(range) function to check your prediction.

How can you generate a set of random numbers that range from 0 to 20? Do it to check your answer.

How can you generate a set of random numbers that range from 9 to 10? Do it to check your answer.

How can you generate a set of random numbers that range from -1 to 1? Do it to check your answer.

Use the RAND function to create a set of 400 numbers that would model a set of measurements with an average value of 5.2 and a standard deviation of 0.5. Write your formula below.

Enter this in a sheet, and check your result. How well does it agree? Recalculate a few times.

Now we have a mechanism by which we can simulate random errors (RAND function), and way to measure how big those errors are, in original data or as propagated (STDEV function). Let's now apply this to a case where we have errors in two different measured values as the same time. We'll use our original example of

$$\frac{e}{m} = \frac{2V}{r^2 B^2} \,.$$

Guidebook Entry I.3: A Real Example with Two Sources of Error

Imagine that we know B exactly somehow, but there are 2% (measured by standard deviation) random errors in both V and r. How big do you expect the relative uncertainty in e/m to be as a result of these uncertainties in V and r?

Create a spreadsheet that models this with 100 different values for V and r. Check to make sure you have reasonable 2% uncertainties in V and r. Write the calculated averages and standard deviations here.

Now, calculate the standard deviation in e/m. How large a percent uncertainty is this?

Explain in a short paragraph why you didn't simply get a final uncertainty of 6% = 4%+2% from the original uncertainties. (If you didn't understand where the 6% came from, ask about that first!)

It is not too difficult to extrapolate our formula on page 2 to the case with several measured quantities (let's assume only two, and call them x and y) that have errors. There we would say that in a similar linear approximation, variations Δx and Δy propagate to a variation Δf given by

$$\Delta f \approx \frac{\partial f}{\partial x}\Delta x + \frac{\partial f}{\partial y}\Delta y \ .$$

In the case of only one variable, we were done at this point, since finding the standard deviation only meant squaring each side, averaging, and taking the square root, which left the basic relationship linear. However, here it is more complicated, and it depends on assuming any individual error in x is unrelated to the corresponding error in y. In other words, if for a particular measurement we know x is actually below average, that gives us no hint as to whether y is high or low relative to its average. So, squaring and averaging now gives:

$$\left\langle (\Delta f)^2 \right\rangle \approx \left\langle \left(\frac{\partial f}{\partial x}\Delta x + \frac{\partial f}{\partial y}\Delta y \right)^2 \right\rangle$$

$$= \left\langle \left(\frac{\partial f}{\partial x}\Delta x \right)^2 \right\rangle + \left\langle \left(\frac{\partial f}{\partial y}\Delta y \right)^2 \right\rangle + 2\left\langle \frac{\partial f}{\partial x}\frac{\partial f}{\partial y}\Delta x \Delta y \right\rangle$$

Recall that this approximation (really a first term in a Taylor's series) uses the value of the derivative at our average values, where Δx = Δy = Δf = 0, and so is just a constant which comes our of the averages. This means that the cross term in the last expression becomes

$$2\left\langle \frac{\partial f}{\partial x}\frac{\partial f}{\partial y}\Delta x \Delta y \right\rangle = 2\frac{\partial f}{\partial x}\frac{\partial f}{\partial y}\langle \Delta x \Delta y \rangle = 0$$

because Δx and Δy are uncorrelated. We can now write the general error propagation formula in terms of standard deviations σ_x, σ_y, and σ_f, instead of Δx, Δy, and Δf in the form

$$\sigma_f^2 \approx \left(\frac{\partial f}{\partial x}\right)^2 \sigma_x^2 + \left(\frac{\partial f}{\partial y}\right)^2 \sigma_y^2$$

You should be able to convince yourself that for simple products and quotients, this is easily applied to the *relative* errors through "adding in quadrature" like perpendicular vectors.

$$\frac{\sigma_f}{f} = \sqrt{\left(\frac{\sigma_x}{x}\right)^2 + \left(\frac{\sigma_y}{y}\right)^2}$$

You can now test your Monte Carlo result with this analytical result above, using for x and y the variables V and r^2. Remember that the relative error for r^2 is double that for r. Do the two propagated errors agree?

(You will do the sum and difference formulae for homework.) Do your results above agree with this formula?