Error Analysis for Introductory Physics Labs*

M. Shane Burns
Professor of Physics
Colorado College

November 12, 2013

1 What is Uncertainty?

Every physical measurement has some uncertainty. Knowing the uncertainty is essential to testing physical theories. For example, suppose a new theory of matter predicts the mass of some new particle to be 137 MeV/c^2. You set out to experimentally measure the mass and you get 141 MeV/c^2. Do these numbers agree? It depends on the uncertainty. If the uncertainty in our measurement was ± 5 MeV/c^2 then our measurement is consistent with the theory, but if the uncertainty was 1 MeV/c^2 we would have to conclude that the theory is inconsistent with the measurement.

Experimental measurements never yield exact results. For example, suppose your physics instructor asked you to measure the length of a small table with a meter stick. You would carefully align one end of the meter stick with one edge of the table then look at the other edge to read off the length. You read off the length by looking at the marks on the meter stick and determining which one lines up with the edge of the table. Is there uncertainty in the measurement? Yes, because you can't read the scale more finely than about a millimeter. The smallest marks on a meter stick are usually one mm apart. The best you could probably do would be to say that the length of the table is in some range.

Suppose you make the measurement described above and find the length to be between 61.2 and 61.4 cm. Your physics instructor would like you to write a short lab report on your measurement. (I know, physics instructors ask you to do some weird things.) How would you report the length? You could say "the length of the table is between 61.2 and 61.4 cm." This would be correct, but physicists have a convention for specifying the results of a measurement; we specify a best estimate plus or minus an uncertainty. In this case, the best estimate of the length would be the value in the middle of the range or 61.3 cm. The uncertainty specifies how much bigger or smaller the length could be or in this case 0.1 cm. In this example, the range 61.2 to 61.4 cm should be specified as 61.3 ± 0.1cm.[1] In general,the result of a measurement of some quantity x is stated as

$$(\text{measured value of } x) = x_{\text{best}} \pm \delta x, \tag{1}$$

where x_{best} is the best estimate of x and δx is the uncertainty in the estimate. If the measurement of x is between x_{max} and x_{min}, then

$$x_{\text{best}} \approx \frac{x_{\text{max}} + x_{\text{min}}}{2}, \tag{2}$$

*Portions of this manual are adapted from "An Introduction to Error Analysis" by John R. Taylor.
[1]The $\pm$ symbol is read as "plus or minus".

and

$$\delta x \approx \frac{x_{\max} - x_{\min}}{2}. \tag{3}$$

2 Reporting Uncertainties and Significant Figures

A few basic rules for stating uncertainties are worth emphasizing. Uncertainty estimates are, after all, just estimates so they should not be stated with too much precision. In our table example, it wouldn't make any sense to quote our measured length as 61.3 ± 0.1239856 cm. It simply isn't possible to know the uncertainty to seven significant figures. Uncertainties are sometimes quoted to two significant figures, but one is usually sufficient. For our purposes we can use the following rule.

> **Rule for Stating Uncertainties:** In an introductory laboratory, experimental uncertainties should be rounded to one or two significant figure.

For example, suppose we make the following three measurements of a time interval: 2.33, 2.43, and 2.28 seconds. Using a calculator, we compute the uncertainty using Eq. 3. Our calculator reads 0.075, but according to the rule above we would round to one significant figure and report the uncertainty as 0.08 seconds.

Once the uncertainty has been rounded, we must also consider the number of significant figures to keep in the best estimate. A statement like

$$\text{measured speed} = 6056.78 \pm 3 \text{ m/s}$$

is obviously ridiculous. The uncertainty of 3 m/s means that the digit '6' in the fourth place of 6056.78 might really be as small as '3' or as large as '9'. Clearly the trailing digits '7', and '8' have no significance at all, and should be rounded off. The proper way to state the result is

$$\text{measured speed} = 6057 \pm 3 \text{ m/s}.$$

The general rule for stating the best estimate of a measured quantity is as follows.

> **Rule for Stating Measured Values** The least significant figure in any measured value should be of the same order of magnitude (in the same decimal position) as the uncertainty.

For the example of the three measured values of a time interval (2.33, 2.43, and 2.28 seconds) we have already used the first rule to determine that the uncertainty should be stated as 0.08 seconds. If we used a calculator to compute the average of the three values, it would read 2.34666667. Using the above rule the best estimate for the time interval is 2.35 and the final result with uncertainty should be reported as 2.35 ± 0.08 seconds.

For the sake of clarity, there are two other things to keep in mind when reporting measured values. The first is that since the uncertainty and the best estimate both have the same units it is clearer to write the result as 2.35 ± 0.08 seconds than 2.35 seconds ± 0.08 seconds. Second, we will often measure numbers that are stated in scientific notation. Suppose we measured a charge to be 1.61×10^{-19} coulombs with an uncertainty of 5×10^{-21} coulombs. The clearest way to report this is $(1.61 \pm .05) \times 10^{-19}$ coulombs rather than $1.61 \times 10^{-19} \pm 5 \times 10^{-21}$ coulombs.

3 Estimating Uncertainties

There are an unlimited number of ways to estimate uncertainty. In fact in complex scientific research it isn't unusual to estimate the uncertainty in a measurement in several different ways in order to cross check the estimates.

Reading Scales In the table example above we got our uncertainty estimate by simply estimating the accuracy to which we thought we could read the meter stick. This is a quick and simple way to get an estimate, but is somewhat subjective.

Digital Readouts Most modern lab equipment is digital. A modern voltmeter is a good example. Suppose we used a voltmeter to measure the voltage across a battery and the digital display read 1.23 volts. What is the uncertainty in this measurement. If we assume the voltage of the battery doesn't change in time and that the voltmeter is calibrated properly, the accuracy of the measurement is given by the accuracy of the display. The display would read the same if the actual voltage where anywhere in the range from 1.225 to 1.234 volts. Using the equation for the best estimate (Eq. 2), we get 1.230 volts and using equation for the uncertainty (Eq. 3) the uncertainty is 0.005 volts. Our best estimate of the battery voltage is then 1.230 ± 0.005 volts. In other words, when reading a digital display the uncertainty is 1/2 the least significant digit that can be displayed.

3.1 Systematic Uncertainties

We made two assumptions in this uncertainty estimate. One was that the battery's voltage didn't fluctuate during the measurement. If it did we would have to use some other technique to estimate the uncertainty like the repeated measurements technique described below.

The other assumption was that the voltmeter was properly calibrated. If it wasn't, then our best estimate wouldn't be correct. A calibration uncertainty like this is an example of a **systematic error**. Systematic errors are errors associated with a flaw in the equipment or in the design of the experiment. Systematic errors cannot be estimated by repeating the experiment with the same equipment. In the battery example, the best way to deal with the systematic error would be to recalibrate the voltmeter. If this isn't possible, then we would have to somehow make an estimate of the possible size of the systematic error and include it in our uncertainty estimate. Systematic errors are insidious because the experimenter usually doesn't know they are present. If they did they would correct the flaw before doing the experiment. The OPERA experiment's mistaken observation of faster than light neutrinos is good example of this kind of hidden error that can lead even good experimental physicists astray[2].

3.2 Random Errors

Many measurements involve uncertainties that can't be estimated by reading a scale or a digital display. For example, if I measure a time interval using a digital stopwatch, the main source of uncertainty isn't usually from the accuracy of stopwatch display, but is from the variability of my reaction time. A good way to estimate the uncertainty in this case would be to make repeated measurements. Errors that can be reliably estimated by repeating measurements are called **random errors**.

Let's say we wished to measure the time it takes a ball to fall from top of a building to the ground. We could measure the time and uncertainty by dropping the ball several times and time each fall with a stopwatch. Suppose we do this five times and get the following five times (in seconds)

2.3, 2.4, 2.6, 2.5, 2.4.

[2]See `http://en.wikipedia.org/wiki/Faster-than-light_neutrino_anomaly`.

We could just use the maximum and minimum values to estimate the best value for the time, but a better technique would be to take advantage of all the data and find the average or **mean** time,

$$t_{\text{best}} = \frac{2.3 + 2.2 + 2.6 + 2.5 + 2.4}{5} = 2.4 \text{ seconds.}$$

The range of measured values has a minimum of 2.2 seconds and a maximum of 2.6 seconds. By Eq. 3 this gives an uncertainty estimate of

$$\delta t \approx \frac{2.6 - 2.2}{2} = 0.2 \text{ seconds.}$$

The measured time is then 2.4 ± 0.2 seconds.

To summarize, if we make n measurements of some quantity x then

$$\boxed{x_{\text{best}} = \bar{x} = \frac{1}{N}\sum_{i=1}^{N} x_i} \tag{4}$$

and

$$\delta x \approx \frac{x_{\text{max}} - x_{\text{min}}}{2}, \tag{5}$$

where the bar above the x indicates the average or **mean**. Eq. 3 gives a quick and dirty estimate of the uncertainty, but it depends on the two extreme data values. This means that it is very sensitive to outlying values. Extreme deviations are rare, but if they do happen the uncertainty estimate of Eq. 3 will overestimate the uncertainty.

3.2.1 Standard Deviation

A much more robust and statistically significant estimate of uncertainty is the **standard deviation**[3]. It is the root mean square deviation of the data values from the mean and is denoted σ_x

$$\boxed{\sigma_x = \sqrt{\frac{1}{N-1}\sum_{i=1}^{N}(x_i - \bar{x})^2}}. \tag{6}$$

The standard deviation of the data for the ball dropping experiment described above is $\sigma_t = 0.1$ seconds (check this yourself). This is smaller than the estimate of Eq. 3 and doesn't cover the whole range of the data, but still gives an estimate of the dispersion of the data points. It's significance is that in most cases[4] if we did one more measurement there would be a 68% chance of this measurement lying in the range from $\bar{x} - \sigma_x$ and $\bar{x} + \sigma_x$. There is a 95% chance it would be within $\pm 2\sigma_x$ of the mean, and a 99.7% chance it would be within $\pm 3\sigma_x$ of $\bar{x}$.

[3]Technically Eq. 6 is called the **sample standard deviation** and is used when only a sample of the total population of possible measurements is used for the calculation. This is almost always the case for measurements made in a physics lab. See Appendix E of "An Introduction to Error Analysis" by John R. Taylor.

[4]By "in most cases" I mean that the data is drawn from a population that is normally distributed. See Chapter 5 of "An Introduction to Error Analysis" by John R. Taylor.

3.2.2 Standard Deviation of the Mean

The standard deviation is a quantitative measure of the uncertainty in a single measurement. In other words, if we measure a time for a single drop of the ball we are 68% certain that this time will be within one standard deviation of the best estimate time for the drop and almost positive (99.7% certain) it will be within three standard deviations of the best estimate time. What is the best estimate of the uncertainty in the value of the*mean*? Again, we can be statistically quantitative. If we made a second set of N measurements of the drop time we would find that 68% of the time, the mean of these measurements would be within one **standard deviation of the mean**,

$$\boxed{\sigma_{\bar{t}} = \sqrt{\frac{1}{N(N-1)}\sum_{i=1}^{N}(t-\bar{t})^2} = \frac{\sigma_t}{\sqrt{N}}}, \tag{7}$$

and 95% of the time within $2\sigma_{\bar{t}}$ etc. Note that the standard deviation and the standard deviation of the mean are related by a factor of $\sqrt{N}$. The distinction between $\sigma_{\bar{t}}$ and σ_t is a subtle but important one. The standard deviation, σ_t, is the *uncertainty in a single trial* and the standard deviation of the mean, $\sigma_{\bar{t}}$, is the *uncertainty in the mean of N trials*. In our introductory physics laboratories we typically make several measurements of some quantity, x, and average them to try to determine some x_{best}. In this case

$$x_{\text{best}} = \bar{x}, \tag{8}$$

and

$$\delta x = \sigma_{\bar{x}} = \frac{\sigma_x}{\sqrt{N}}. \tag{9}$$

Let's work through an example. Suppose we measure the time it takes the ball to fall from the roof to the ground five times and get the following five values for the time

$$2.35,\ 2.48,\ 2.46,\ 2.51,\ 2.62,$$

where all the times are in seconds. Taking the average we find that $t_{\text{best}} = 2.484$ seconds. The standard deviation, $\sigma_t = 0.097$. This means that if we made one more measurement, we would have a 68% of getting a time in the range from $t_{\text{best}} - \sigma_t = 2.387$ and $t_{\text{best}} + \sigma_t = 2.581$. The standard deviation of the mean $\sigma_{\bar{t}} = \frac{\sigma_t}{\sqrt{5}} = 0.043$. We would therefore report our result for the best estimate of the time as

$$t = 2.48 \pm 0.04 \text{ seconds}.$$

Notice that the standard deviation of the mean is inversely proportional to $\sqrt{N}$, this suggests that the larger the number of measurements we make the smaller the uncertainty in the average[5]. This means that if our experiment is dominated by random uncertainty, we can always get a more accurate result by making more measurements.

When we decide to make a measurement using repeated trials we have to decide how many trials to do. The bare minimum is two, but this doesn't give a very good estimate of the uncertainty. Taking at least 3 measurements allows us to determine if one of our measurements is in error. For example, if we made three measurements of a time interval and got 2.4, 6.5 and 2.6 seconds we could be quite sure that our second measurement was wrong (maybe we forgot to reset the stop watch). If we had only made two measurements we might not catch our error. The uncertainty in the mean goes down as we take more measurements, but since $\sigma_x \propto 1/\sqrt{N}$ there is a point of

[5]Although this sounds reasonable, it actually takes quite a bit of work to prove. See "An Introduction to Error Analysis" by John Taylor for the proof.

diminishing returns. The improvement in the uncertainty when we take five measurements instead of three is about the same as the improvement in going from five measurements to ten. In practice, five measurements is usually optimal.

4 Propagation of Uncertainty

Suppose we wanted to determine the perimeter and area of a rectangular table top from measurements of its width $w = 76.4 \pm 0.2$ cm and its length $\ell = 153.3 \pm 0.6$ cm. How would we determine the uncertainties in the perimeter and area? Let's compute the perimeter first. The perimeter $p = 2w + 2\ell = 2(w + \ell)$. The highest probable value for the perimeter is then

$$p_{\max} = 2(w + \delta w) + 2(\ell + \delta\ell) = 2(w + \ell) + 2(\delta w + \delta\ell),$$

and lowest probable value is

$$p_{\min} = 2(w - \delta w) + 2(\ell - \delta\ell) = 2(w + \ell) - 2(\delta w + \delta\ell).$$

We can use Eq. 3 to get the uncertainty in p

$$\delta p \approx \frac{p_{\max} - p_{\min}}{2} = 2(\delta w + \delta\ell).$$

For the example at hand, $\delta p = 1.6$ cm, so using the rounding rules of Section 2 gives $p = 459 \pm 2$ cm.

You can always get a crude estimate the uncertainty using this brute-force method of computing the maximum and minimum values and using Eq. 3, but it tends to overestimate random uncertainties. This is because to get the maximum uncertainty in the sum both ℓ and w must conspire to be overestimates. If the uncertainties are independent and random, then an underestimate in the measurement one of the variable, ℓ for example, is partially compensated for by an overestimate in the measurement of the other variable, w in this case. Statistical theory tells us that if some quantity $q = x + y$ and if x and y are normally distributed then the sum is also normally distributed with standard deviation

$$\sigma_q = \sqrt{\sigma_x^2 + \sigma_y^2}. \tag{10}$$

This is always less than the sum of the standard deviations of x and y. When we combine two numbers by squaring them, adding the squares, and taking the square root as in Eq. 10, the numbers are said to be **added in quadrature**. Eq. 10 would be the same if $q = x - y$ so in general we have the following rule.

> **Uncertainty in Sums and Differences:** If q is the sum or difference of several quantities $x, y, z...$
>
> $$q = x + y - z \cdots$$
>
> and if the uncertainties, $\delta x, \delta y, \delta z...$, are independent and random, then the uncertainty in q is the quadrature sum,
>
> $$\boxed{\delta q = \sqrt{\delta x^2 + \delta y^2 + \delta z^+ \cdots}}. \tag{11}$$

There is a similar rule for products and quotients.

Uncertainty in Products and Quotients If several quantities $w, x, y, z...$ are measured with small independent and random uncertainties $\delta w, \delta x, \delta y, \delta z...$ and

$$q = \frac{w \times x \times \cdots}{y \times z \times \cdots}$$

then the *fractional* uncertainty in q is the quadrature sum of the *factional* uncertainties in $w, x, y, z...$,

$$\boxed{\frac{\delta q}{|q|} = \sqrt{\left(\frac{\delta w}{w}\right)^2 + \left(\frac{\delta x}{x}\right)^2 + \left(\frac{\delta y}{y}\right)^2 + \left(\frac{\delta z}{z}\right)^2 + \cdots}}. \tag{12}$$

The area of the rectangle $A = w\ell = 11712.12$ cm^2 and the fractional uncertainty in A is

$$\frac{\delta A}{|A|} = \sqrt{\left(\frac{\delta w}{w}\right)^2 + \left(\frac{\delta \ell}{\ell}\right)^2} = \sqrt{\left(\frac{0.2}{76.4}\right)^2 + \left(\frac{0.6}{153.3}\right)^2} = 0.0047,$$

so $\delta A = 55$ cm^2. Using the rounding rules of Section 2 gives $A = 11710 \pm 55$ cm^2.

Sometimes the value we are interested in can't be written as a simple sum, difference, product, or quotient of the measured variables.

Uncertainty Propagation for Any Function Suppose that some physical quantity $q(x, y, z)$ is a function of three measured values x, y, and z. If δx, δy, and δz are the uncertainties in these measured values then statistical theory tells us that the uncertainty in q is

$$\boxed{\delta q = \sqrt{\left(\frac{\partial q}{\partial x}\delta x\right)^2 + \left(\frac{\partial q}{\partial y}\delta y\right)^2 + \left(\frac{\partial q}{\partial z}\delta z\right)^2}}. \tag{13}$$

When q is a complicated function of many variables, Eq. 13 can be formidable, but keep in mind that we only need to know δq to at most two significant figures. This means that if one of the terms under the radical in Eq. 13 is significantly larger than the others we can ignore all but the largest term.

5 Comparison of Two Values

Sometimes we wish to compare two measured values. This is often the case when we wish to compare a experimentally determined value with a theoretical value which also has some uncertainty. In this case, we have agreement if the ranges of the two measured values overlap. For example, suppose we determine the spring constant of a spring by measuring its displacement from equilibrium when a mass hangs from it and find $k_1 = 10 \pm 2$ N/m. Then we determine the spring constant a second time by measuring the period of the springs motion when a mass hangs from it and find $k_2 = 13 \pm 2$ N/m. In this case the range 8 to 12 N/m overlaps the range 11 to 14 N/m so the two values agree within experimental uncertainty. When the ranges are nowhere near overlapping, the values do not agree. If the two ranges are close but don't overlap then one can't make any definitive statement about whether or not the two values agree. In this case, the result of the measurement is inconclusive.

One convenient, but crude way to compare two quantities p and q is to take the difference $p - q$. In principle the difference should be consistent with zero. Using the error propitiation equation for

differences Eq. 11 gives $\delta(p-q) = \sqrt{\delta p^2 + \delta q^2}$. If the ratio of the absolute value of the difference and the uncertainty is less than about 2 we have acceptable agreement—if it is much greater than about 2.5 the two values *do not* agree. If the the value falls in the gray region between about 2 and 2.5, the experiment is inconclusive. Table 5 summarizes this crude rule-of-thumb for comparing two values.

$\frac{\lvert p-q\rvert}{\sqrt{\delta p^2+\delta q^2}} \lesssim 2$	acceptable agreement
$2 \lesssim \frac{\lvert p-q\rvert}{\sqrt{\delta p^2+\delta q^2}} \lesssim 2.5$	inconclusive
$\frac{\lvert p-q\rvert}{\sqrt{\delta p^2+\delta q^2}} \gg 2.5$	values don't agree

Table 1: Rule-of-thumb for comparing two values. This technique of comparing two values is related to Student's t-test.

The discussion above illustrates the importance of estimating uncertainties. One cannot make a comparison without an uncertainty estimate. *The uncertainty in a measured quantity is just as important as the value of the measured quantity.* Experimental physicists often spend as much or more time trying to estimate the uncertainty in a measured value as they do determining the measured value itself.

6 Graphical Representation of Data

Remember the old adage, "A picture is worth a thousand words." A graph is the best way to make your data understandable when one observed quantity is depends on another. For example, suppose you measure the distance a car has traveled as time goes on and get the values shown shown in Table 2.

time (seconds)	distance (meters)
0.24	5.76×10^{-2}
1.29	1.68×10^{0}
2.35	5.55×10^{0}
3.41	1.16×10^{1}
4.47	2.00×10^{1}
5.53	3.06×10^{1}
6.59	4.34×10^{1}
7.65	5.85×10^{1}

Table 2: Data for distance versus time measurement.

It is very difficult to see how the distance is changing with time just by looking at the numbers in the table. Figure 1 shows a graph of the data from Table 2. It's easy to see from a quick look at the graph that the car is accelerating. It's very difficult to see this kind of trend in the data in a table.

Take a close look at Figure 1. Notice that the figure has clearly labeled axes including units. You don't need to draw lines from data point to data points, in fact this is misleading in some cases. Try

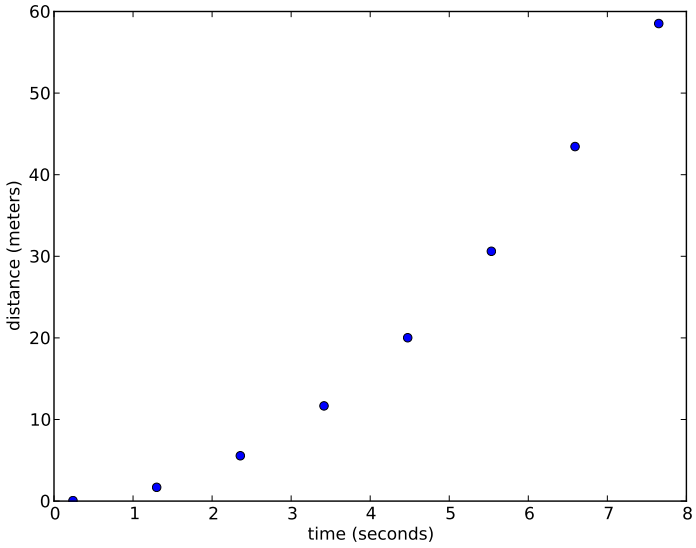

Figure 1: Graph of the data in Table 2.

to get into the habit of creating the plot as you take the data. Doing this allows you to recognize immediately when a particular measurement needs to repeated or when another data point should be taken to fill in a gap. Generating a final graph on a computer is okay, but it's quicker and easier to make a plot by hand on graph paper in the lab. Use a straight edge on graph paper and be careful to plot each point as accurately as possible.

You can represent uncertainties in your data by putting error bars on your graphs. Table 3 contains data from a radioactive decay experiment where $\dot{N}$ is the decay rate as a function of time t. There is an estimated uncertainty in the rate. Figure 2 shows a plot of these data with error bars to represent the uncertainty in $\dot{N}$.

t (seconds)	$\dot{N}$
0.00	16 ± 4
0.05	15 ± 3
0.10	10 ± 3
0.15	12 ± 3
0.20	9 ± 3
0.25	7 ± 2
0.30	3 ± 1
0.35	5 ± 2
0.40	4 ± 2
0.45	1 ± 1

Table 3: Rate of radioactive decay $\dot{N}$ as a function of time t.

The plot in Figure 2 is again computer generated. With a little practice, it's often quicker and

easier to do the plots in lab by hand. Microsoft Excel does an okay job of creating plots, but it's not really made for plotting scientific data. Ask your instructor if you'd like to learn about better software for generating scientific graphs.

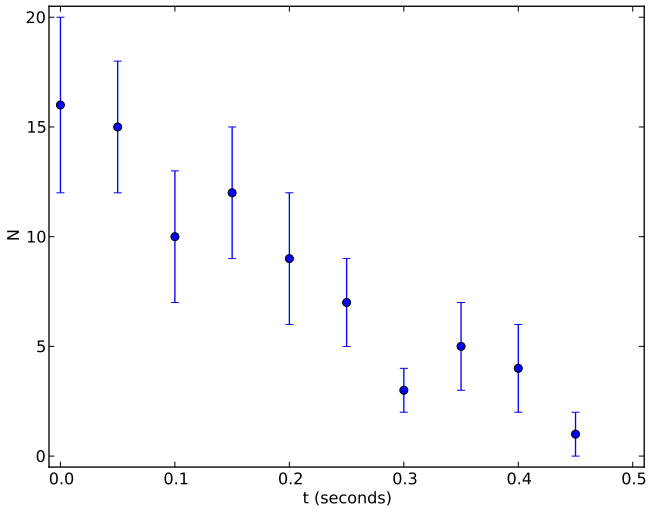

Figure 2: Graph of the data in Table 3.

6.1 Linear Fitting

Suppose you have a set of N data pairs (x_i, y_i), that are supposed to be linearly related so that

$$y = A + Bx. \tag{14}$$

The goal of a linear fit is to find the best fit parameters A and B. For a linear least squares fit, the equations for A and B are derived by finding A and B that minimizes

$$\chi^2 = \sum_{i=1}^{N} \frac{(y_i - A - Bx_i)^2}{\sigma_i^2}, \tag{15}$$

where σ_i is the uncertainty in y_i.[6] The value of χ^2 can be used to judge the quality of the fit. A fit is "good" if $\chi^2 \approx N - 2$. If $\chi^2 \ll N - 2$ it indicates that the uncertainties (σ_i) have been overestimated. If $\chi^2 \gg N - 2$ then either the uncertainties have been underestimated or y isn't really linearly related to x by Equ. 14.[7]

Given the equations for A and B, one can use standard error propagation to derive general equations for their uncertainties. However, the values one accepts for these uncertainties depends on the quality of the fit. In the following sections, I outline how to estimate the uncertainties in A and B given something other than a perfect fit. The next section deals with the case in which σ_i

[6] We will assume the error in x_i is negligible

[7] See Chapter 12 of *An Introduction to Error Analysis* by Taylor for an explanation.

is the same for all the data. The final section generalizes the technique to a weighted fit where the σ_i can be different. I state most of the results without proof. A more rigorous description can be found in Bevington's or Taylor's books on error analysis.

6.1.1 Uncertainties for an Unweighted Fit

In this section we assume that the y-uncertainty is the same for all the data. If we call this uncertainty σ, then

$$\chi^2 = \frac{1}{\sigma^2}\sum_{i=1}^{N}(y_i - A - Bx_i)^2. \tag{16}$$

The best estimates of A and B are those that minimize χ^2. Taking the partial derivatives of χ^2 with respect to A and B and setting them equal to zero gives

$$A = \frac{(\sum x_i^2)(\sum y_i) - (\sum x_i)(\sum x_i y_i)}{\Delta}, \tag{17}$$

$$B = \frac{N(\sum x_i y_i) - (\sum x_i)(\sum y_i)}{\Delta}, \tag{18}$$

where

$$\Delta = N(\sum x_i^2) - (\sum x_i)^2. \tag{19}$$

We already know the uncertainty in the y_i is σ, but can also use our fit to estimate what the uncertainties should be by looking at how much our data points deviate from the fit. The standard deviation from the fit gives us an estimated uncertainty

$$\sigma_{\text{est}} = \sqrt{\frac{1}{N-2}\sum_{i=1}^{N}(y_i - A - Bx_i)^2}. \tag{20}$$

By comparing this result with the equations for χ^2 (Equ. 16) it is relatively easy to show that

$$\sigma_{\text{est}}^2 = \frac{\chi^2}{N-2}\sigma^2. \tag{21}$$

At this point it is convenient to define the reduced chi-squared,

$$\tilde{\chi}^2 \equiv \frac{\chi^2}{N-2}. \tag{22}$$

Given this definition, $\sigma_{\text{est}}^2 = \tilde{\chi}^2\sigma^2$. Note that if we have a "good" fit, $\tilde{\chi}^2 = 1$ and $\sigma_{\text{est}} = \sigma$ as we would expect. In fact, this is what we mean by a good fit. If $\tilde{\chi}^2 \ll 1$ then $\sigma_{\text{est}} \ll \sigma$ and we have overestimated the errors, σ_i. If $\tilde{\chi}^2 \gg 1$ then $\sigma_{\text{est}} \gg \sigma$ and we have underestimated the error or the relation between x and y isn't really linear.

We can now use standard error propagation to find the uncertainties in A and B. It isn't too hard to show that

$$\sigma_A = \sqrt{\frac{\sum x_i^2}{\Delta}}\delta y,$$

$$\sigma_B = \sqrt{\frac{N}{\Delta}}\delta y,$$

where δy is the y-uncertainty. We now have a choice: do we let δy be our original uncertainty estimate, σ, or the uncertainty derived from the fit, $\sigma_{\rm est}$. If $\tilde{\chi}^2 = 1$ the choice is moot since they are equal. If the relation between x and y is truly linear, $\sigma_{\rm est}$ is a better estimate of the error that we actually have. Therefore the best estimates for σ_A and σ_B are

$$\sigma_A = \sigma_{\rm est}\sqrt{\frac{\sum x_i^2}{\Delta}} = \tilde{\chi}\sigma\sqrt{\frac{\sum x_i^2}{\Delta}}, \tag{23}$$

$$\sigma_B = \sigma_{\rm est}\sqrt{\frac{N}{\Delta}} = \tilde{\chi}\sigma\sqrt{\frac{N}{\Delta}}. \tag{24}$$

However, be very cautious if you find $\sigma_{\rm est} \gg \sigma$. Having a $\chi^2 > N - 2$ may mean the relation is not linear.

6.1.2 Uncertainties in a Weighted Fit

If the y uncertainties are not all equal then we can't factor them out of the sum in the expression for χ^2 as we did in Equ. 16. Instead we must use the more general expression for χ^2 given in Equ. 15. When we take the partial derivatives of Equ. 15 and set them equal to zero to minimize χ^2 we get the following equations for A and B:

$$A = \frac{(\sum w_i x_i^2)(\sum w_i y_i) - (\sum w_i x_i)(\sum w_i x_i y_i)}{\Delta}, \tag{25}$$

$$B = \frac{(\sum w_i)(\sum w_i x_i y_i) - (\sum w_i x_i)(\sum w_i y_i)}{\Delta}, \tag{26}$$

where

$$\Delta = (\sum w_i)(\sum w_i x_i^2) - (\sum w_i x_i)^2, \tag{27}$$

and $w_i = 1/\sigma_i^2$. Applying error propagation to the equations for A and B allows use to compute the uncertainties in A and B. We find that

$$\sigma_A = \sqrt{\frac{\sum w_i x_i^2}{\Delta}},$$

$$\sigma_B = \sqrt{\frac{\sum w_i}{\Delta}}.$$

However, to derive these equations we have assumed that the σ_i are accurate estimates of the uncertainty so that $\tilde{\chi}^2 = 1$. If this is not the case then the above equations will either underestimate or overestimate the uncertainties in A and B. If we are confident that the the relation between x and y is linear, then just as in the case in section 6.1.1, better estimates of the uncertainties are

$$\sigma_A = \tilde{\chi}\sqrt{\frac{\sum w_i x_i^2}{\Delta}}, \tag{28}$$

$$\sigma_B = \tilde{\chi}\sqrt{\frac{\sum w_i}{\Delta}}. \tag{29}$$

However, be cautious in using these equations. If $\tilde{\chi}^2 \gg 1$, it could mean that the relation between x and y is not linear. It could also mean your uncertainty estimates aren't reliable in which case it may be better to use an unweighted fit. However, if you are confident that the relation between x and y is linear, and that at least the relative size of the uncertainties are correct, then Equ. 28 and Equ. 29 give you the best estimates of the uncertainties in A and B.