Princeton University
Physics Department

Physics 103/105 Lab

Reference Packet

ERROR ANALYSIS INTERLUDES

Fall, 2003

NOTE TO STUDENTS:

This material closely parallels that found in John R. Taylor's An Introduction to Error Analysis, 2nd Edition (University Science Books, Sausalito, California; 1997). Taylor's text is a standard on the subject, and is well worth acquiring for your long-term use.

The Physics Department has in the past also provided Errors Without Tears, by Professor Daniel Marlow, as a handout in Physics 103/105 lab. It covers much the same material as this packet; copies are available on request.

Material from Matt Trawick (Physics 103, Fall 2002)
Format revised Summer 2003 (S. Smith)

"Table of Contents" for Error Analysis Interludes

ERROR ANALYSIS INTERLUDE #1

What is uncertainty, and who cares?

In 1919, Albert Einstein's general theory of relativity faced a crucial test. On May 29, the moon would come between the Earth and the Sun in a solar eclipse, making it possible to see stars very close in the sky to the Sun's position without being blinded by the glare from the Sun itself. Einstein predicted that as light from a distant star passed by the sun, it would be deflected by the sun's immense gravitational field by a miniscule angle of 1.75 arc seconds (3600 arc seconds = 1 degree). A calculation based on Newtonian physics predicted an even smaller deflection of 0.875 arc seconds.

Observations led by Arthur Eddington at three different telescopes measured the deflection: the resulting deflections were 0.86, 1.61, and 1.98 arc seconds. None of these three measurements agrees exactly with either prediction. They don't even agree with each other! How could three measurements of the same thing all be so different?

All measurements, however careful and precise, are subject to some degree of "uncertainty." (The scale in a doctor's office is probably more accurate than a cheap bathroom scale from Kmart, but neither one is perfect.) However, by knowing the amount of uncertainty in a measurement, it is often possible to make definite statements even from indefinite results. For example, Eddington reported that the middle of his three measurements had an uncertainty of 0.30 arc seconds. In other words, although his best estimate from that measurement was 1.61 arc seconds, the actual value might have been anywhere between 1.31 and 1.91 arc seconds. Einstein's prediction of 1.75 is clearly consistent with that measurement. More importantly, the Newtonian prediction of 0.875 is clearly not. Notice that the uncertainty of a measurement is just as important as the measurement itself. Were it not for the uncertainty, it would be impossible to differentiate between two competing theories. Based on his analysis of his uncertainty, Eddington concluded that the Newtonian paradigm was inconsistent with his new measurements. The general theory of relativity had passed its first experimental test, and Albert Einstein became an instant celebrity.[1]

We seldom get to test theories as important as Einstein's general theory of relativity. But professional scientists and engineers routinely use measurements and their uncertainties to differentiate between competing theories large and small. You might be surprised how often such situations arise even in everyday life. What if your car is supposed to get 25 miles per gallon, and you measure its fuel economy to be only 20 miles per gallon? Only by knowing the uncertainty of your measurement can you determine whether there's really something wrong with your car.

The uncertainty in a measurement is often called the "error" of a measurement. "Error analysis" means figuring out how big the uncertainty is. It is important to remember that in this sense the "error" of a measurement is not a mistake. All measurements, no matter how well done, have some degree of error or uncertainty.

[1] Although relativity has since passed many other experimental tests, it is not clear whether Eddington's conclusion in 1919 was really justifiable based on his data. Error analysis is tricky business, and even the pros get it wrong sometimes. See, eg. Jim Holt, Lingua Franca, Volume 11, No. 2—March 2001.

Although "uncertainty" is clearly the best word to use, the words "error" and "uncertainty" are used interchangeably.

The "error" of a measurement is not just the difference between the measured value and the "right" answer. In the case of Eddington's measurements, he couldn't possibly calculate that, because nobody knew for sure what the "right" answer was. In the case of your gas-guzzling car, you might wish to calculate the difference or "discrepancy" between the book value of the fuel economy and your own measurement. Stated in absolute terms, the difference between the two is 20 mpg – 25 mpg, a discrepancy of -5 mpg. You may also wish to state the discrepancy in percentage terms:

$$\text{percentage discrepancy} = (20 \text{ mpg} - 25 \text{ mpg}) / 25 \text{ mpg} \times 100\% = -20\%.$$

This is *not* the uncertainty in your measurement. The uncertainty might be more or less than -20%, depending how carefully you made your measurements.

You can only determine the uncertainty of a measurement by examining in detail exactly how the measurement was made. In all likelihood, you would have calculated the fuel economy by measuring both the distance you drove, and the amount of fuel you used to get there. Each of these measurements alone has some uncertainty, and these are combined mathematically to calculate the uncertainty in the final quantity. It is this kind of calculation that error analysis is all about.

How to Report Uncertainties

Let's suppose that you want to measure the length of a fish you have caught. (This is a challenging measurement, since the fish is wet, slippery, and considerably less interested in having its length measured than you are.) Bracing yourself against your wobbly canoe, you place a ruler alongside the reluctant creature and judge its length to be about 27.5 cm. (It's clearly between 27 cm and 28 cm.) After a little consideration you decide that you are *pretty sure* that the length is between 27.2 and 27.8 centimeters. Satisfied, you record the length in your notebook as $x = 27.5 \pm 0.3$ cm. (The symbol "±" is read "plus or minus.") The value before the ± sign, 27.5 cm, is called the "best value" of your measurement x. The value after the ± sign, 0.3 cm, is the uncertainty in x, and is written δx, where δ is the lower-case Greek letter "delta." (You may also see the uncertainty written as Δx or σ_x in some texts.)

In the example above, we said that you were "pretty sure" the length of the fish was between 27.2 and 27.8 centimeters. Just how sure is pretty sure? Scientists typically report their uncertainties based on a "confidence level" of 68%. In this case, "pretty sure" means that you are 68% sure that the actual length of the fish lies somewhere in your reported range.[2] Said another way, if you caught and measured lots of fish, the actual lengths of about two-thirds of them would lie within your reported range.

If the 68% confidence level seems a bit unwieldy, try asking yourself the following: *would you bet a dollar that the fish is between 27.2 and 27.8 centimeters*? In general, you might be willing to make such bets as long as you have a pretty good chance of winning, and as long as losing isn't going to destroy you. With that in mind, you will want to estimate your uncertainty so that the actual length of the fish has a pretty good chance of being within your stated range. If somebody with more patience than you

[2] 68% seems like an odd number to choose for a confidence level. We'll see in a later error analysis diversion where this comes from.

manages to persuade all of your fish to hold still for laser interferometry measurements, it's okay if about a third of those measurements fall outside your stated range. In general, if you are just willing to bet a dollar that the actual value falls within your error range, your error estimate is probably about right.

Of course, you could simply overstate your uncertainty, and report all of your fish measurements as ± 10 centimeters. That ought to be safe enough, right? Indeed, if you were trying to sell the fish, and you wanted to guarantee that each was within a certain range, it might be a good idea to make that range as broad as possible, to avoid the hassle of giving lots of refunds to angry customers. (Manufacturers often set "tolerances" this way, with lots of room to spare!) But for scientific purposes, your notebook will be most valuable if you record your honest best guess of your uncertainty, without any extra padding.

> **Rule for confidence level of errors:** When you write a quantity as $x \pm \delta x$, you mean that you are about 68% sure the actual value of x falls between $x - \delta x$ and $x + \delta x$.

Absolute and Relative Uncertainties

In reporting your uncertainties, it is sometimes important to know the size of your uncertainty relative to the measurement itself. For example, measuring the distance x from the tip of your nose to the stake in the ground at the South Pole to within one meter would require a fantastically precise measurement... unless you were standing right next to the South Pole. What matters here is the ratio $\delta x/x$, which is defined as the "relative uncertainty" of the measurement. If you were standing 5 meters from the South Pole, the relative uncertainty of your measurement would be

$$\delta x/x = 1 \text{ m} / 5 \text{ m} = 0.2, \text{ or } 20\%.$$

But if you were a continent away from the South Pole, the relative uncertainty might be

$$\delta x/x = 1 \text{ m} / 5{,}000{,}000 \text{ m} = 0.0000002, \text{ or } 0.00002\%.$$

(Amazingly, the Global Positioning System (GPS), which uses satellites with very accurate atomic clocks, makes such precise measurements routine. We live in extraordinary times!)

The relative uncertainty is also called the "fractional uncertainty" or "percentage uncertainty." (It is also sometimes called the “precision” of a measurement, but this is potentially confusing because the word has another more specific meaning, as we’ll see later.) To avoid confusion with the relative uncertainty, the quantity δx is sometimes called the "absolute uncertainty." It's always a good idea to have in the back of your mind roughly what your relative uncertainty is for a given experiment. Are you doing a 1% measurement or a 10% measurement?

Rule for relative and absolute uncertainties: the *relative uncertainty* of a measurement of x is defined as $\delta x/x$, and is usually expressed as a percentage. The *absolute uncertainty* is δx. *Uncertainty* alone can mean either one.

Significant Figures

Scientists often use a kind of shorthand for writing uncertainties: significant figures. Writing a mass as 3.26 grams implies that you know the mass is not 3.25 or 3.27 grams, but 3.26± 0.005 grams. In this case your measurement has three "significant figures" or "significant digits:" the 3, the 2, and the 6. Written in kilograms, 0.00326 kg also has three significant digits; the leading zeroes are merely place markers. If the uncertainty of this measurement was 0.00005 grams, then you would write five significant digits: 3.2600 grams.

Unfortunately, this kind of shorthand can lead to some problems. First, if an object is "250 pounds," it is unclear whether the final zero is significant. (Does this mean 250 ± 5 pounds or 250 ± 0.5 pound?) The confusion can be resolved by writing 2.50×10^2 pounds, but that's a little awkward. Second, this shorthand does not always allow you to state your uncertainty as precisely as you'd like. Suppose you have measured the width of a beetle's antenna as 0.30 ± 0.02 mm. Writing "0.3 mm" implies 0.30 ± 0.05 mm, which overstates your error, but writing "0.30 mm" implies "0.300 ± 0.005 mm, which understates your error. Simply put, significant figures are not a very precise way to report your uncertainties.

In Physics 103, you will be expected to report all of your uncertainties explicitly. But we also expect you to be sensible about significant figures. If you are calculating the circumference C of a wheel 3.0 centimeters in diameter, *do not under any circumstances, write that* $C = \pi d = 9.424778$ *centimeters.* You can't possibly know the circumference to that kind of precision, no matter what your calculator says!

We're not going to fret over whether you write 9.4 or 9.42 as part of some intermediate calculation. In fact, feel free to keep an extra digit or two in your intermediate calculations, and then drop them later on. But please, resist the temptation to write down all of the digits produced by your calculator. It just makes you look silly!

In general, you will only know your uncertainty δx to one or two decimal places at most, and that is what you should write. For example, you should write the uncertainty of the fish's length as "δx = 0.3 cm" (not "δx = 0.31853 cm"). Report your results so that the last significant digit of your best value is in the same position as the last significant digit of the uncertainty. For example, write "27.5 ± 0.3 cm" (not "27.5127 ± 0.3 cm"). Again, the extra digits just look silly.

Rule for significant figures: Don't be writing ridiculous numbers of meaningless digits, or you'll look like a doofus.

Different Kinds of Errors

Your calculator is not the only instrument that can give you more digits than are really useful. For example, you may be wearing a digital watch right now that gives you the time of day right down to the second. But we instinctively know that unless you were very careful in setting your watch (and did so recently) it will only be accurate to a few minutes. The distinction here is that while a digital watch is very "precise", it may or may not be very "accurate." A measurement that is “precise” has the ability to distinguish between two values that may be close together. A measurement that is “accurate” is close to the correct value.

As another example, the speedometer in your car may tell you your speed with a precision of about ± 1 mile per hour; that is, you can distinguish the position of the needle to ± 1 mph. However, the speedometer may not be as accurate as it is precise. What if your tires are the wrong size, or the speedometer's electronics are not properly calibrated? Beware of this kind of "false precision," and don't believe everything you read!

Another word that you may see is the "reproducibility" of a measurement. Even if your measuring device is very precise and accurate, your measurement may still vary from measurement to measurement. For instance, if you were to listen to your heart beating, you might count exactly 41 beats in 30 seconds, a rate of 82 beats per minute. This single measurement is very accurate and very precise (probably ± 1 beat per minute), but if you tried the same measurement 5 days or even 5 minutes later, you might get a very different answer. This is an example of a measurement with a high precision and a high accuracy, but a low reproducibility.

The lack of reproducibility of your own heart rate is an example of a random error; the result of the measurement is likely to change from measurement to measurement. It may sometimes be a little higher or a little lower than normal, but the results will usually cluster around the same place over time. Similarly, if you try to estimate tenths of millimeters from a standard ruler, you might by chance guess either a little high or a little low. This lack of precision is also a random error, unless you are holding the ruler at an angle and reading the result too low every time.

Holding a ruler at an angle and reading the result too low every time would be an example of a "systematic error:" one that does not cause your measurement to change from time to time. Similarly, if the tires on your car are a little too big, the accuracy of your speedometer reading will be affected by the same amount every time you measure your speed. These systematic errors are often hard to detect, and sometimes the only way to discover them is to compare your measuring device to another one that is known to be more accurate. (In the case of the speedometer, you could drive past a police speed trap, and see if they chase you.)

Vocabulary Review

We've introduced many new words and phrases in this first section. Take a quick look at these, and be sure you can come up with a reasonable definition of what they mean. Many of these are different words for the same thing.

uncertainty	*percentage uncertainty*	*false precision*
error	*absolute uncertainty*	*reproducibility*
discrepancy	*best value*	*accuracy*

error analysis	*confidence level*	*random error*
relative uncertainty	*precision*	*systematic error*
fractional uncertainty		

As you might guess, words like "precision," and "accuracy" that have been defined in this section are not always used exactly as they have been described here. Although many authors and scientists are real sticklers for these distinctions, others (including your lab instructor!) may not be so careful. Try to use these words as "precisely" as you can. When you hear others using these words, be prepared to ask them what they really mean to say.

ERROR ANALYSIS INTERLUDE #2

What Is Error Propagation?

As professional scientists and engineers, you will almost never get to measure the final result of an experiment directly. Instead, you will deduce the final result from measurements of related quantities. For example, you might deduce the mass of a black hole by measuring its effects on nearby stars. Or you might measure the optical properties of a chemical solution to calculate what fractions of the reactants have combined to form a new compound.

In each case, the uncertainty of your result is determined by the uncertainties of your original measurements. Just how these measured uncertainties "propagate" through your calculations to your final results is no easy matter, and is much of what error analysis is all about. In this unit, we begin to study error propagation.

Unit Conversion

Recall that in the last unit, you had measured the length of a fish to be $x = 27.5 \pm 0.3$ cm, a relative uncertainty of $\delta x/x = 0.3\text{ cm}/27.5\text{ cm} \approx 1\%$. Let's suppose you wanted to report this in millimeters instead of centimeters: how would you write that? Of course, you would write $x = 275 \pm 3$ mm. Think for a second about how you did that: you multiplied both your result and your uncertainty by the same factor of 10. Has the relative uncertainty of your measurement changed? Of course not; it's still $\delta x/x = 3\text{ cm}/275\text{ cm} \approx 1\%$.

> **Rule for converting units:** Convert your uncertainty in the same way as your best value. The relative uncertainty stays the same.

Easy Propagation of Single Errors

Let's suppose that you plan to drop a water balloon out of your window, and you want to know how fast it will be going when it splats on the sidewalk below. Trying it out, you measure the time of the fall to be about 1 second, and you calculate that that

$$\begin{aligned} v &= gt \\ &= \left(10\text{ m/sec}^2\right)\left(1\text{ sec}\right) \\ &= 10\text{ m/sec}. \end{aligned}$$

(For convenience, we'll assume that g, the acceleration due to gravity, is exactly 10 m/sec.)

Suddenly driven by a compulsion you can't quite explain, you decide to calculate your uncertainty in this measurement. Although you timed the fall carefully with your digital stopwatch, you guess that you may have been off by about 0.1 seconds, a relative uncertainty of about 10%. In the one extreme, this leads to a result of

$$\begin{aligned} v &= gt \\ &= \left(10 \text{ m/sec}^2\right)\left(0.9 \text{ sec}\right) \\ &= 9 \text{ m/sec}. \end{aligned}$$

In the other extreme, the result becomes

$$\begin{aligned} v &= gt \\ &= \left(10 \text{ m/sec}^2\right)\left(1.1 \text{ sec}\right) \\ &= 11 \text{ m/sec}. \end{aligned}$$

Based on your uncertainty in the time of the fall t, you conclude that the velocity right before impact was between 9 m/sec and 11 m/sec, so you state your result as $v = 10 \pm 1 \text{ m/sec}$. Congratulations, you've just done your first error propagation!

> **Rule for easy propagation of single errors:** if $f = f(x)$, you can always calculate $f(x+\delta x)$ and $f(x-\delta x)$ to see how big your error is.

Multiplying by a Constant

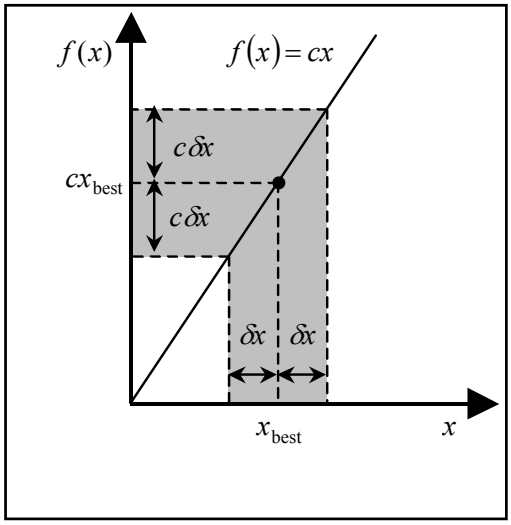
Let's look at that last example again. You actually calculated

$$\begin{aligned} v &= gt \\ &= \left(10 \text{ m/sec}^2\right)\left(10 \pm 1 \text{ sec}\right) \\ &= 10 \pm 1 \text{ m/sec}. \end{aligned}$$

The reader with a flair for observation will notice that the absolute uncertainty in the time δt = 0.1 sec, multiplied through by the constant g = 10 m/sec, gives the absolute uncertainty in the velocity of δv = 1 m/sec. Furthermore, the relative uncertainty of the result is exactly 10%, the same as the relative uncertainty of the original measurement. We can generalize this as a rule for multiplying by a constant.

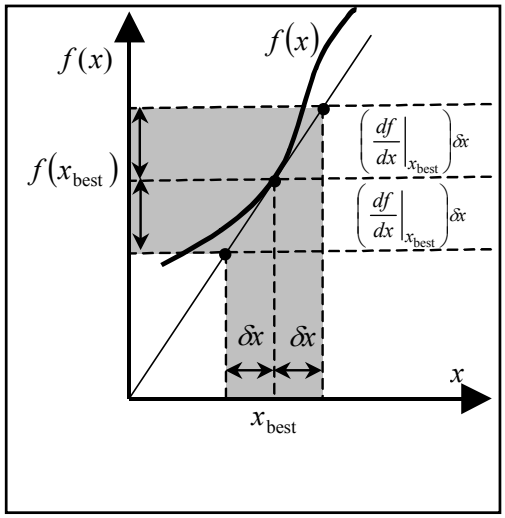
> **Rule for multiplying by a constant:** if $f(x) = cx$, where c is a constant, then $\delta f = c\,\delta x$. The relative uncertainty is unchanged: $\delta f / f = \delta x / x$.

A look at the first graph at the right will show why this is true. The graph shows a situation where $f(x) = cx$, where c is a precisely known constant. The slope of the line is c, and so the uncertainty is $\delta f = \left(\frac{df}{dx}\right)\delta x = c\,\delta x$.



General rule for single small errors

The special case of multiplying a measurement by a constant shows us how we can deduce a general rule for propagating the uncertainty in any single measured variable, provided the relative uncertainty in that variable is small. A look at the second graph at the right will show how this is so.



The second graph shows an arbitrary function $f(x)$ that is not a straight line. If the uncertainty δx is small, then we can approximate the function $f(x)$ near the best value x_{best} as a straight line through the point $(x_{\text{best}}, f(x_{\text{best}}))$. The slope of this line is the value of $\frac{df}{dx}$ evaluated at x_{best}, which we denote by $\left(\frac{df}{dx}\Big|_{x_{\text{best}}}\right)$. We can see from the graph that $\delta f \approx \left(\frac{df}{dx}\Big|_{x_{\text{best}}}\right)\delta x$.

Rule for propagation of single small uncertainties: if $f = f(x)$ then $\delta f \approx \left(\frac{df}{dx}\Big|_{x_{\text{best}}}\right)\delta x$ for small values of δx.

Example: a power law.

Problem: Let's suppose you want to measure the height of a five-story building by dropping a baseball off of the roof and timing its fall. Using your digital watch, you measure the fall to be $t = 2.0 \pm 0.1$ seconds. Your best value of the building height h is

$$h = (1/2)gt^2$$
$$h = (1/2)(10\ \text{m/sec}^2)(2\ \text{sec})^2$$
$$h = 20\ \text{m/sec}.$$

What is the uncertainty of this measurement?

Solution: First, we'll find δh using the rule for propagation of single small uncertainties:

$$\delta h \approx \left(\left. \frac{dh}{dt} \right|_{t_{\text{best}}} \right) \delta t$$

$$\delta h \approx \left(g t_{\text{best}} \right) \delta t$$

$$\delta h \approx \left(10 m / \sec^2 \right) \left(20 \sec \right) \left(0.1 \sec \right)$$

$$\delta h \approx 2 \text{ meters.}$$

Answer: The measurement is $h = 20 \pm 2$ meters.

Let's compare this to what happens if we calculate our error using the rule for easy propagation of single errors, based on our highest and lowest estimates of t.

Lowest case	Highest case
$h = (1/2) g t^2$	$h = (1/2) g t^2$
$h = (1/2)\left(10 \text{ m} / \sec^2\right)(1.9 \sec)^2$	$h = (1/2)\left(10 \text{ m} / \sec^2\right)(2.1 \sec)^2$
$h = 18.05$ m	$h = 22.05$ m

Here, we have kept a few extra digits from our calculation to make a point. Our rule for propagation of single small uncertainties tells us that h is between 18 and 20 meters, but the rule for easy propagation of single errors tells us that h is between 18.05 and 22.05 meters. Remember, the rule for propagation of single small uncertainties is only an approximation for small errors. In this case, the approximation is close enough.

ERROR ANALYSIS INTERLUDE #3

Adding or Subtracting Two Quantities

Standing in the middle of a soccer field, you and your friend suddenly wonder whether the field you have been practicing on is the regulation length of 115 yards. You decide to pace it off, each walking from the middle to opposite ends. You each measure a distance of 50 ± 10 yards, or ± 20%. (Pacing is a pretty crude measurement, after all.) Adroitly adding 50 yards plus 50 yards, you calculate the your best guess of the length of the field to be 100 yards. Now the hard part: what is your uncertainty in this measurement? Is it:

a) 10 yards?
b) 20 yards?
c) Something between 10 and 20 yards?

If you think about it carefully, you can eliminate answer (a) as a possibility. After all, 10 yards is the uncertainty you would expect if you had measured your part as you did, but somehow managed to know the other part exactly. Surely, having your friend pace it off instead should add something to the uncertainty in the resulting measurement.

Answer (b) is a little more tempting. After all, let's consider the two worst-case scenarios: If both distances are on the high side of your range, the result is a field of $60 \text{ yards} + 60 \text{ yards} = 120 \text{ yards}$. If both distances are on the low side of your range, the result is an area of $40 \text{ yards} + 40 \text{ yards} = 80 \text{ yards}$. That's a measurement of $100 \pm 20 \text{ yards}$, right?

Not so fast. It turns out that answer (b) is an overestimate of the error, because it is also possible that the two errors will tend to cancel each other. The table to the right is a summary of some possible discrepancies between the actual lengths and the measured values of 50 yards. Here, we assume that there are only three possibilities: the actual length of the distance you measured might be 10 yards greater (+10), or 10 yards less (-10) than your estimate of 50 yards, or the actual value might be exactly 50 yards (0). The same three possibilities are shown for your friend's side of the field. (Of course, for each measurement there are actually an infinite number of possibilities, but to list them all would make the table rather long and tedious, and difficult to post on the web.)

Your Discrepancy	Friend's Discrepancy	Total Discrepancy
+10	+10	+20
+10	0	+10
+10	-10	0
0	+10	+10
0	0	0
0	-10	-10
-10	+10	0
-10	0	-10
-10	-10	-20

Of the 9 possibilities listed, only 2 of them lead to an error as high as 20 yards. It appears to be very unlikely that the uncertainty in the total length of the field could be as much as 20 yards.

Let's be a little more quantitative about it, remembering the "Rule for confidence level of errors" from Error Analysis Interlude #1. If your measurement of the field is

50 ± 10 yards, this means that you estimate that there is a 68% chance that the actual length is between 40 and 60 yards. There is a 16% chance that the length is 60 yards or greater, and a 16% chance that the length is 40 yards or less. We can now ask ourselves: what are the chances that you and your friend's measurements are both 60 yards or greater? Go ahead. Take a minute and do the calculation. Don't read ahead.... Try it yourself first.... Don't look....

The probability that both measurements would be over 60 yards is $(0.16)(0.16) = 0.026$ or 2.6%. The probability that both measurements would be less than 40 yards is also 2.6%. So if you did state your measurement of the length of the field as 100 ± 20 yards, you would be giving the confidence interval of 95% instead of the usual 68%. Clearly, the 68% confidence interval (the one we usually use) must be less than that. So the uncertainty of this measurement must be somewhere between 10 and 20 yards, answer (c).

So what is the uncertainty of your measurement? We can't prove it here, but it turns out that the best way to combine those two independent errors is by adding in *quadrature*: square each one, add them together, and take the square root. In this case, $\delta l = \sqrt{(10 \text{ yards})^2 + (10 \text{ yards})^2} = 14$ yards. (Just like finding the hypotenuse of a right triangle, $c = \sqrt{a^2 + b^2}$.) This means that if you had to bet a dollar one way or the other, you would bet that the field was less than the regulation 115 yards. But remember, your error estimate is only your 68% confidence level; you sure wouldn't want to bet your life on it.

Rule for adding and subtracting quantities with independent errors: if $f = x + y + z + \ldots - a - b - c - \ldots$ then $\delta f = \sqrt{(\delta x)^2 + (\delta y)^2 + (\delta z)^2 + \ldots + (\delta a)^2 + (\delta b)^2 + (\delta c)^2 + \ldots}$.

Dependent Errors

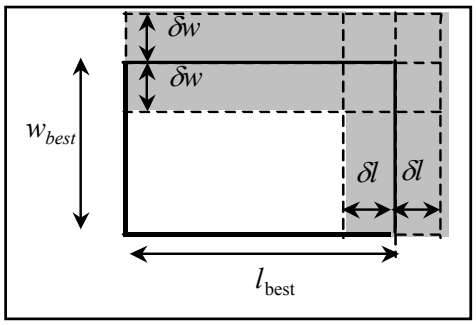
Adding in quadrature is the rule we use when we think the two errors are independent of each other. In the example above, it's quite likely that your measurement is too low and your friend's is too high, or vice versa. However, if the two quantities are not independent of each other, for example if the same person paces off both halves of the field with a systematic error in stride length that effects both halves equally, then adding in quadrature isn't the right thing to do. In that case, the errors would simply add together.

Rule for adding and subtracting quantities with dependent errors: if $f = x + y + z + \ldots - a - b - c - \ldots$ then $\delta f = \delta x + \delta y + \delta z + \ldots + \delta a + \delta b + \delta c + \ldots$.

Products and Quotients

Having just measured the length of the field, you and your doubtlessly geeky friend are just getting warmed up! You decide to measure the width of it and calculate its area too. Your friend quickly paces it off and reports the answer: "70 plus or minus 10 yards. Let's see, 70 times 100, that's 7000 square yards. Now what about the error?"

You can't use the rule for adding and subtracting quantities here. (If you try it, you'll see that the units won't even work out. You'll end up with 7000 square yards plus or minus 17 yards, which is nonsense.)



The figure to the right shows the situation graphically. The best value for the area is $A_{\text{best}} = l_{\text{best}} \times w_{\text{best}}$, but it might be as large as $A_{\text{best}} + \delta A = (l_{\text{best}} + \delta l)(w_{\text{best}} + \delta w)$, or as small as $A_{\text{best}} - \delta A = (l_{\text{best}} - \delta l)(w_{\text{best}} - \delta w)$. The uncertainty in the area, shown by the shaded regions, is roughly $\delta A = w\delta l + l\delta w$. The relative uncertainty would be

$$\delta A / A = (w\delta l + l\delta w) / A$$
$$\delta A / A = \delta l / l + \delta w / w.$$

Notice that for the product of l and w, it's the relative uncertainties that add together, not the absolute uncertainties.

But as in the case of addition and subtraction, simply adding the relative uncertainties $\delta A / A = \delta l / l + \delta w / w$ actually overstates the error, because there is only a small chance that δw and δl will both be very large or both be very small. Again, the best estimate of the error is obtained by adding the two relative uncertainties in quadrature, $\delta A / A = \sqrt{(\delta l / l)^2 + (\delta w / w)^2}$. This is the general rule for combining uncertainties in products, and it works for quotients too.

Rule for products and quotients with independent errors: if $f = \dfrac{x \times y \times \ldots}{a \times b \times \ldots}$, then

$$\frac{\delta f}{f} = \sqrt{\left(\frac{\delta x}{x}\right)^2 + \left(\frac{\delta y}{y}\right)^2 + \left(\frac{\delta a}{a}\right)^2 + \left(\frac{\delta b}{b}\right)^2 + \ldots .}$$

General Rules for Multiple Errors

In the previous error unit, we wrote a rule for small errors in $f(x)$: $\delta f \approx \left(\left. \dfrac{df}{dx} \right|_{x_{\text{best}}} \right) \delta x$. This rule can be combined with the principle of adding independent errors in quadrature, leading to a very general rule for combining multiple independent errors:

Rule for multiple independent errors: if $f = f(x, y, z, \ldots)$ then

$$\delta f \approx \sqrt{\left(\left. \frac{\partial f}{\partial x} \right|_{x_{\text{best}}} \delta x \right)^2 + \left(\left. \frac{\partial f}{\partial y} \right|_{y_{\text{best}}} \delta y \right)^2 + \left(\left. \frac{\partial f}{\partial z} \right|_{z_{\text{best}}} \delta z \right)^2 + \ldots .}$$

(The symbol $\frac{\partial f}{\partial x}$ is the partial derivative. It means to hold the rest of the stuff constant, and take the derivative of just that one variable.) In fact, both the rule for addition and subtraction and the rule for multiplication and division can be derived from this one general rule.

Of course, multiple errors might not always be independent of each other. In the case of dependent errors, we would not add the terms in quadrature.

General rule for multiple dependent errors: if $f = f(x, y, z, \ldots)$ then

$$\delta f \approx \left(\left.\frac{\partial f}{\partial x}\right|_{x_{\text{best}}} \delta x \right) + \left(\left.\frac{\partial f}{\partial y}\right|_{y_{\text{best}}} \delta y \right) + \left(\left.\frac{\partial f}{\partial z}\right|_{z_{\text{best}}} \delta z \right) + \ldots.$$

Don't be intimidated if these general rules for multiple errors look big, hairy, and unfriendly to you. In fact, they are merely extensions of two principles you have seen before: the rule for propagating single small uncertainties, $\delta f \approx \left(\left.\frac{df}{dx}\right|_{x_{\text{best}}} \right) \delta x$, and the general principle of combining uncertainties by adding them in quadrature when they are independent.

Example: Combining Two Errors.

Problem: In bright sunlight, the total solar power P incident on a circular solar panel is given by $P = C\pi R^2 \cos(\theta)$, where the constant $C = 0.1367$ Watt/cm^2, R is the solar panel's radius and θ is the angle at which it faces the sun. Suppose that you have measured $R = 5.0 \pm 0.1$ cm and $\theta = 30 \pm 1$ degrees, and calculated $P = 9.30$ Watts. *What is your uncertainty in this result?*

Solution: We'll find δP using the general rule for multiple independent errors. Notice that we have to convert $\delta\theta$ to radians for our calculation:

$$\delta P \approx \sqrt{\left(\left.\frac{\partial P}{\partial R}\right|_{R_{\text{best}}} \delta R \right)^2 + \left(\left.\frac{\partial P}{\partial \theta}\right|_{\theta_{\text{best}}} \delta\theta \right)^2}$$

$$\delta P \approx \sqrt{\left(2C\pi R_{\text{best}} \cos(\theta_{\text{best}})\delta R\right)^2 + \left(C\pi R_{\text{best}}^2 \sin(\theta_{\text{best}})\delta\theta\right)^2}$$

$$\delta P \approx \sqrt{\begin{array}{l}\left((2)(0.1367\ \text{W/cm}^2)(\pi)(5.0\ \text{cm})(\cos 30^\circ)(0.1\ \text{cm})\right)^2 \\ \quad + \left((0.1367\ \text{W/cm}^2)(\pi)(5.0\ \text{cm})^2(\sin 30^\circ)(0.0175\ \text{rad})\right)^2\end{array}}$$

$$\delta P \approx \sqrt{(0.37\ \text{W})^2 + (0.093\ \text{W})^2}$$

$$\delta P \approx 0.38 \text{ Watts}$$

Notice that although the relative uncertainty in θ is larger than that for R (3%, compared to 2%), the contribution to δP from $\delta\theta$ (0.093 W) is significantly smaller than that from

δR (0.37 W). This is due to both the magnifying effect of squaring R, and the relatively shallow slope of $\cos\theta$ at $\theta = 30^{o}$.

Ignoring Small Errors

If you look again at the example problem above, you will see that the effect of the uncertainty $\delta\theta$ disappears almost entirely when it is added in quadrature to the larger uncertainty due to δR. Let's look at another example. Suppose you are multiplying two independent quantities together: $f = xy$. Let's suppose your relative uncertainty in x is 4%, and your uncertainty in y is 1%. Look what happens when you calculate the relative uncertainty in f:

$$\frac{\delta f}{f} = \sqrt{(4\%)^2 + (1\%)^2} = 4.12\% \approx 4\%.$$

Notice that the smaller error makes almost no difference in the final calculation of the uncertainty. Adding in quadrature makes it possible to ignore the small one completely. It will often be the case that your measurements will be dominated by a single source of error, and all the others can be safely swept under the rug.

Rule for Ignoring small errors: Feel free to ignore small independent errors. They only add in quadrature, so what the heck?

Although you have been given many "rules" to keep track of, you will find as you use them that they become more a matter of common sense than anything else. (Don't dread long calculations either; in many cases, the uncertainty in your final result will be dominated by the error in a single variable, and other uncertainties can be safely ignored.) You will quickly get a feel for estimating the effects of large errors and ignoring small ones. In time, you error analysis will become quite natural. In the same way that a good auto mechanic knows which parts of an old car are most likely to cause trouble, you will soon get a feel for which laboratory measurements are your largest sources of uncertainty, how to estimate their effects, and even how best to minimize them.

ERROR ANALYSIS INTERLUDE #4

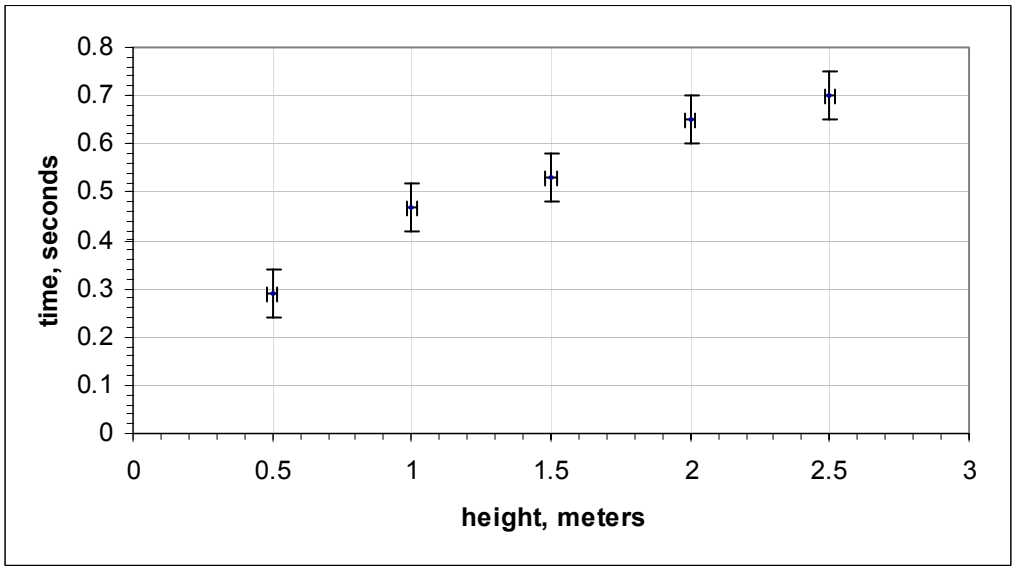
Plotting Data With Error Bars

So far in physics 103, whenever you have made a graph of your experimental data, you have represented your data points with little round dots or circles on the graph. The size of the dots was either whatever size the computer program happened to give you, or whatever size was convenient to make with your pencil in your notebook. And although the dots or circles were actually a few millimeters in diameter on the paper, we thought of them as representing infinitesimally small points, at their centers.

But if you stop and think about it, there's something inherently wrong (or at least pretty misleading) about treating your data as infinitesimally small points. In fact, every

measurement you make is subject to some uncertainty. While the center of a dot is a good way to denote your best guess of a measurement's value $(x_{\text{best}}, y_{\text{best}})$, it leaves out the equally important information about your measurements' uncertainties, δx and δy. The convention used by scientists and engineers to convey this additional information in a graph is to use *error bars* around the data points.

In the picture below, a student has dropped a marble from several different heights *h* and plotted the time *t* of the fall as a function of *h*. Both *h* and *t* are somewhat uncertain for each data point, so there are two error bars for every point, representing δh and δt. (Sometimes you'll see the two error bars represented as an ellipse or a circle.) Often, if the uncertainty of one of the coordinates is either very small or simply not known, a single error bar may be used. In this case, since the relative uncertainty in *h* is apparently much smaller than the relative uncertainty in *t*, the student might well have chosen to skip the error bars showing δh and shown only δt.

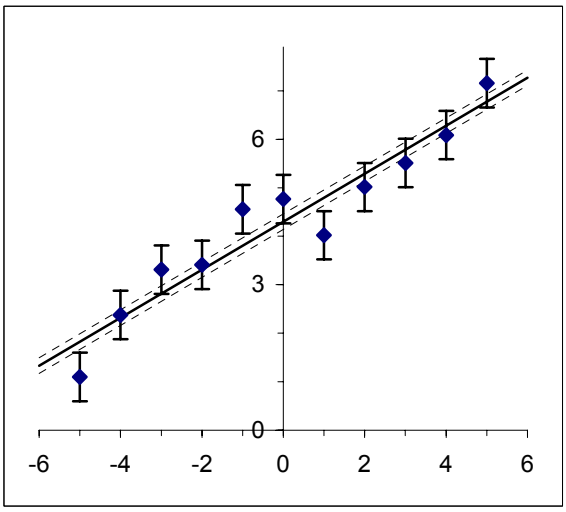


(You can easily add error bars to your own Excel graphs. Double click on one of your data points on the graph to open the "Format Data Series" window, then click on the "X Error Bars" or "Y Error Bars" tab.)

Fitting Data To A Straight Line: The Eyeball Method

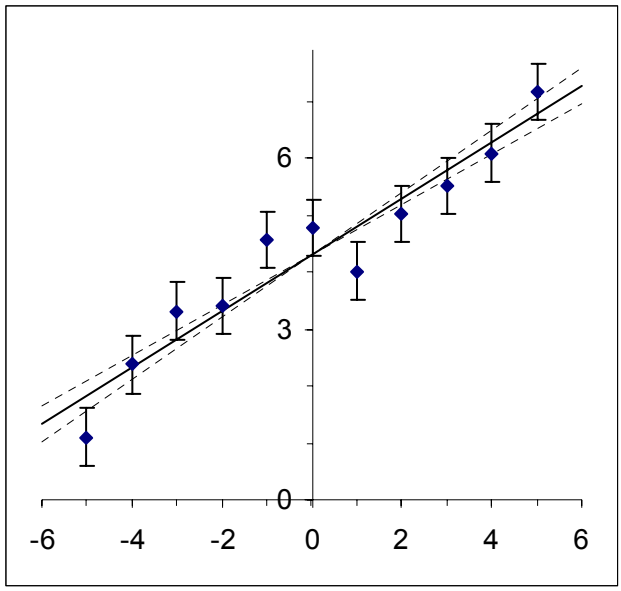
The next graph shows a set of experimental data points that closely resembles a straight line. Of course, the points do not fall exactly on a line, in part because the measurement that produced them is subject to some uncertainty. Here, we assume that any uncertainty in the *x* values is much smaller than the uncertainty in the *y* values, so we include only error bars for δy in the graph.

Let's suppose that you want to find the equation for the line that "best" fits this apparently linear data; that is, you suspect that the data is of the form $y(x) = A_0 + A_1 x$, and you wish to find values of A_0 and A_1 that are *most likely* to be the actual, true values, based on the available data. If you had to guess at where the line would be based on your intuition and common sense, you would probably draw the line right about where the solid line is in the figure. You could then read the *y*-intercept ($A_0 \approx 4.3$) and slope ($A_1 \approx 0.50$) of that line from the graph, and you'd have your answer. This is what we call the eyeball method.



Now you need to know the uncertainty in A_0 and A_1. After all, somebody else might have drawn the "best fit" line in a slightly different place. Imagine holding the slope of the line constant, and sliding the line up and down along the *y*-axis. A reasonable person might slide the "best fit" line as high or as low as the two dotted lines in the figure. This uncertainty in the location of the line leads to an uncertainty of $\delta A_0 \approx 0.2$. Similarly, you can imagine holding the *y*-intercept fixed at $A_0 = 0.50$ and changing the slope to find the uncertainty in A_1, as is suggested in the following figure.

The error bars included in the graph are a valuable tool for determining δA_0 and δA_1. Do you remember the rule of the 68% confidence limit? If these data points really describe a linear relationship between *x* and *y*, and if the error bars have been drawn correctly to indicate the usual 68% confidence limit, then the error bar on each point should have about a 68% chance of overlapping the line. Stated another way, the best fit line should hit about two-thirds of the error bars in the graph. Any fewer than that, and the line is probably not in the right place. Sure enough, we see that the solid straight line in the figure hits 8 out of the 11 error bars, which is close enough to two-thirds that we may suspect we're on the right track.

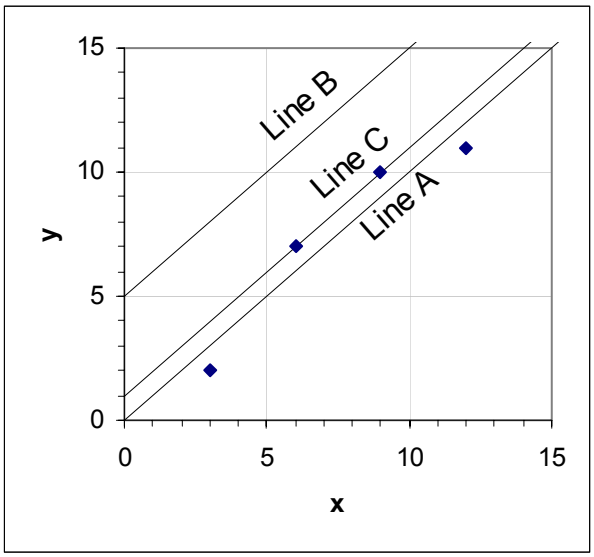


The bottom line here is that the uncertainties in the y values in each of the data points leads to an uncertainty in the "best fit" parameters A_0 and A_1.

Fitting Data: The Method Of "Least Squares"

Computers don't have eyeballs or common sense, so they pretty much stink at the "eyeball method" described above. Instead, a computer program uses a specific mathematical algorithm to find the best fit parameters. The algorithm you will learn about here is called the "least squares" fit. Using a process like this to fit data to a line is called a "linear regression."

Suppose you have four data points to fit to a straight line, as shown in the figure to the right. Intuitively, you want your line to be "close to" where the points are, so line A drawn in the figure seems like a good first start.



Is minimizing the distance between the line and the points exactly what we want the computer to do? The vertical distance between line A and each of the points in the figure is 1 unit, and the sum of these distances is 4. Clearly if you raised the best fit line to line B, you would increase that total distance, and make the fit worse. On the other hand, line C is clearly less likely than line A to be the actual line approximated by the data, and yet the total distance between the points and line C is also exactly 4. Apparently, minimizing the total vertical distance, or differences, between the points and the line is not quite right.

Instead, we can hit on a reasonable algorithm if we try to minimize the sum of the squares of these differences. For line A, each difference is 1 unit, so each difference squared is also 1, and the sum is 4. For line C, the difference between the lines and the two missed points is 2 units; square it and get 4; the sum of the squares is 8. By this algorithm, line A is clearly the better fit than line C. Apparently we have just stumbled on a reasonable way to find a best-fit straight line: find the line that minimizes the sum of the squares of the differences between the line and all the data points. This is the method we call "least squares fitting."

Finding the "least squares" line need not be a matter of trial and error. (Although a computer is infinitely patient for that kind of tedium, you may not be.) Hopefully, you can see right away what has to be done.

Imagine a line $y(x) = A_0 + A_1 x$ on the same graph as your data, and a bunch of data points (x_i, y_i), where $i = 1,2,3 \ldots N$ up to however many data points you have. Write down the difference between each of the data points and the line:

$$\text{difference} = y_i - (A_0 + A_1 x_i)$$

Square the differences and sum them:

$$\text{Sum of differences squared} = \sum_{i=1}^{N} \left(y_i - \left(A_0 + A_1 x_i \right) \right)^2$$

To find the values of A_0 and A_1 for which the sum is a minimum, take the derivatives and set them to zero:

$$\frac{\partial}{\partial A_0} \left[\sum_{i=1}^{N} \left(y_i - \left(A_0 + A_1 x_i \right) \right)^2 \right] = 0$$

$$\frac{\partial}{\partial A_1} \left[\sum_{i=1}^{N} \left(y_i - \left(A_0 + A_1 x_i \right) \right)^2 \right] = 0$$

The result is a set of two equations and two unknowns, which either you or a computer can solve. The result, shown here for completeness only, is

$$A_0 = \frac{\left(\sum x_i^2 \right) \left(\sum y_i \right) - \left(\sum x_i \right) \left(\sum x_i y_i \right)}{N \left(\sum x_i^2 \right) - \left(\sum x_i \right)^2}$$

and

$$A_1 = \frac{N \left(\sum x_i y_i \right) - \left(\sum x_i \right) \left(\sum y_i \right)}{N \left(\sum x_i^2 \right) - \left(\sum x_i \right)^2}$$

Life is too short to bother memorizing these equations, in fact, with a computer at hand, life is too short to even bother ever using them yourself.

Although you've seen that the least squares method is a reasonable way to find the best fit straight line, we haven't proved that it's the best way. ("Why not minimize the sum of differences to the fourth power," you might ask.) The real reason has to do with the mathematical form of what is called the *normal distribution* (more popularly called the bell curve), which you will learn about in the next error analysis interlude. You can make the argument for the exponent of 2 as mathematically rigorous as you like, but we won't do it here.

The method of least squares works for fitting other functions besides just straight lines. For fitting a quadratic, you'll have three parameters: A_0, A_1, and A_2. For non-polynomial fits, the mathematics becomes a bit more hairy, but least squares fitting is still the basic recipe you'll need.

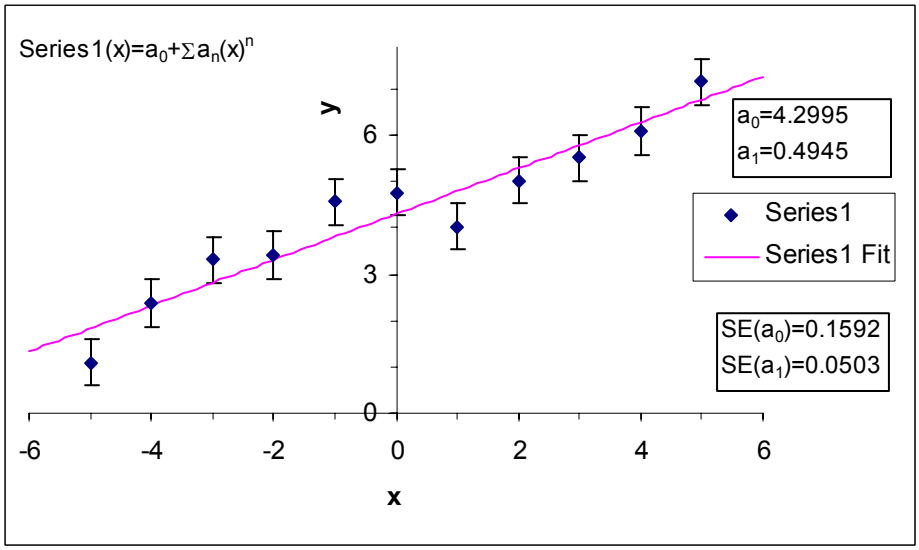
Uncertainty In The Best Fit Straight Lines.

Having just written down a formula for how to find the parameters A_0 and A_1 that will best fit your data to a straight line, we now ask, "can we calculate the uncertainty in these parameters?"

The answer is that you can, and you already know how to do it, at least in principle. The formula that we wrote down for A_0, for example, is just a big, messy formula of lots of variables, $A_0(x_1, y_1, x_2, y_2, \ldots)$. Some of those variables (maybe all of them) have uncertainties associated with them. (Remember those error bars on the *y* coordinates?) In principle, you could use the rule for the propagation of multiple independent errors that we learned about in Error Analysis Interlude #3,

$$\delta f \approx \sqrt{\left(\left.\frac{\partial f}{\partial x}\right|_{x_{\text{best}}} \delta x\right)^2 + \left(\left.\frac{\partial f}{\partial y}\right|_{y_{\text{best}}} \delta y\right)^2 + \left(\left.\frac{\partial f}{\partial z}\right|_{z_{\text{best}}} \delta z\right)^2 + \ldots}$$

and just crank it through. Mercifully, you won't have to. Whenever you ask a program like WPTools to find the best values of A_0 and A_1, it can find the uncertainty in them too. That's what WPTools has done in the figure on the below, where it tells you "SE(a_0)=0.1529." The "SE" stands for *standard error*, which you'll learn about in the next error analysis interlude.



But now you may be wondering just what WPTools is actually doing, since you never told it anything about the uncertainties in all of your data. How does it know? The answer is that it estimates the uncertainty of the data points based on how far away they are from the best-fit line or best-fit curve it finds. If the data are very close to the line, the uncertainties in their values must be small. If the data are very scattered, their uncertainties must be large. WPTools uses the scatter in the data (more precisely, their deviation from the best-fit line or curve) to determine their uncertainties, and then uses those uncertainties to calculate the uncertainty in the fitting parameters A_0, A_1, A_2, and so on.

The good news is that the random differences between the data and the best-fit curve are usually an excellent measurement of the amount of random errors in the data. The bad news is that this totally ignores any non-random, or systematic errors in the data. These will have to be factored in later.

Example: combining with systematic error.

Problem: Amir and Belinda have just taken a video of a cart moving along an air track at a constant velocity, and they want to know its speed. They copied their data of x position vs. time into Excel, and had WPTools fit their data to a straight line. WPTools tells them the slope of the *x*-position *vs.* time graph is "0.71", and the standard error in that slope is "0.031" Amir and Belinda write down the speed in their notebooks: 0.71 ± 0.03 m/sec, or 4%. Are they done?

Solution: They're not done yet. The linear regression routine in WPTools has given them the uncertainty in the slope based only on the random errors in the data. It has no way to determine the uncertainty in the speed due to any systematic errors in the data. Amir and Belinda will have to account for those themselves.

The random errors are things that would vary from point-to-point. The biggest random error is the clicking of the points themselves, which at best have a precision of only one pixel. The systematic errors are the ones that are the same for all of the (x,t) data points. Perhaps the clock is running a bit slowly on their computer, affecting all of the time measurements? Perhaps the camera was tilted slightly so that there was also some motion of the cart along the y-axis in addition to the x-axis? Or perhaps the scale factor converting pixels to meters is slightly imprecise? Of these, the scale factor is by far the largest, and the others may be safely ignored.

Amir recalls that to scale their movie, they measured a half meter stick to be 65 pixels across. Belinda figures they knew could only have measured that distance to within 2 pixels, a relative uncertainty of $2/65 = 3\%$. Alone, this would yield a 3% uncertainty in their speed calculations. Combining their 3% systematic error to their 4% random error from the curve fitting, Amir and Belinda find their real error in their velocity is $\sqrt{(3\%)^2 + (4\%)^2} = 5\%$.

If you wish, you can look at Amir and Belinda's problem in another way. The fitting routine really only told them the slope of their line in pixels per second. If they hadn't input their scaled data, the fitting routine would tell them their speed was 92 ± 4 pixels/second, or 4%. To convert to meters, they divide this by their scale factor, 65 ± 2 pixels/0.5 meter, or 3%. To combine these errors, they use the rule for multiplying and dividing and combine the relative uncertainties in quadrature, as before.

Don't Believe Everything You Read: A Cautionary Note.

In general, we know better than to try to stick a square peg in a round hole. But computer software generally isn't that bright. WPTools will happily fit a straight line to curved data, and report back to you the best fit parameters with their uncertainties as if everything was just fine.

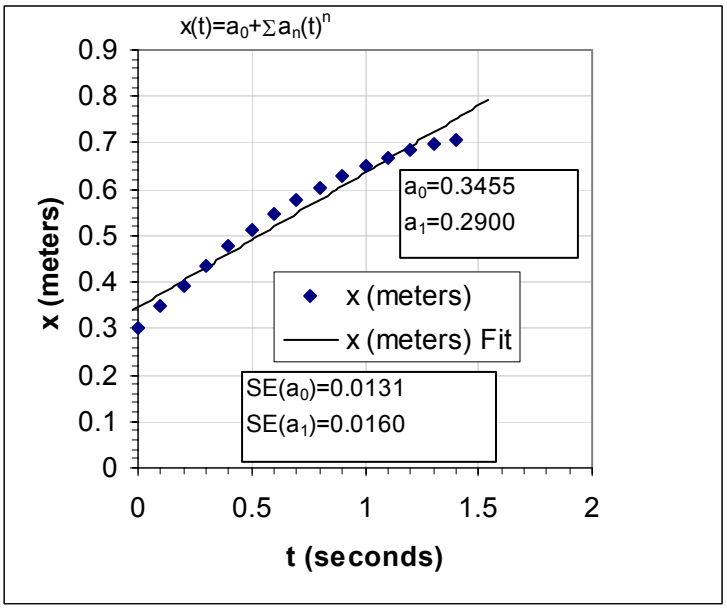
Don't believe everything you read! After the computer fits your data, look at it and see if it makes sense. Your data points should be randomly distributed above and below the fitted curve. If all the data in the middle of the range is above the straight line, and all of the data on the ends are below it, that would be your big clue that your data isn't really a straight line. Something else is going on, and you'd better understand it or you're going to get screwy results.

Example: Misuse of a least squares fit

Problem: Albert and Barb have are analyzing a collision between two carts. To find the velocity of one of the carts after the collision, they plot the x_position *vs.* time, and fit the data to a straight line, shown in the graph below, where the collision has happened at time $t = 0$. Is the slope of the straight line fit (0.29 m/sec) really the best value for the velocity immediately after the collision?

Solution: Clearly, data after the collision is not exactly a straight line. Apparently, their cart does not have a constant velocity as they had guessed. Reading the slope off of the straight line fit will not give the best estimate of the speed immediately after the collision.

The students could try to fit a smaller portion of their data right after the collision, during which time the velocity would be more constant, though with fewer data points they would have a larger uncertainty in their value. Bettter yet, they could model their cart as undergoing constant acceleration. They could then fit their position data to a quadratic equation $x(t) = x_0 + v_0 t + \frac{1}{2} a t^2$, and find the best fit of the initial velocity v_0 that way. Or they could oil the cart's wheels and try the experiment again.

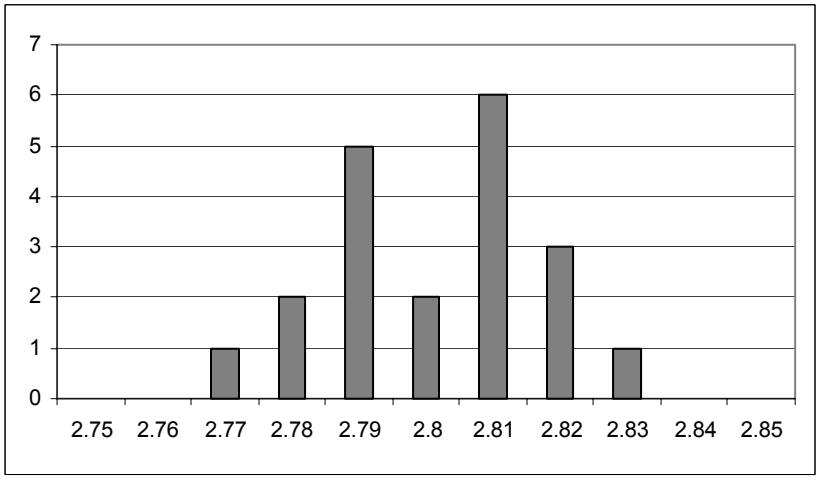


ERROR ANALYSIS INTERLUDE #5

Repeated measurements and distributions

Many of the experimental measurements we make are subject to random errors: our results are slightly different every time we make the measurement. For example, suppose that we are measuring *g* by timing the period of a pendulum with a handheld stopwatch. We might measure the following times, in seconds: {2.79, 2.81, 2.83, 2.77, 2.80, 2.81, 2.81, 2.79, 2.82, 2.82, 2.81, 2.82, 2.79, 2.81, 2.81, 2.78, 2.79, 2.78, 2.79, 2.80}. By repeating the measurement several times as we have, it is possible to both estimate our uncertainty in the final result, and minimize it.

The histogram on the following page shows the distribution of our measurements of the period. There was one measurement of 2.77, two measurements of 2.78, five of 2.79, and so on. If we took many more measurements, we might reasonably expect that our histogram might gradually begin to look like a smooth, bell-shaped curve. As long as we suspect that our distribution would eventually become a more-or-less symmetric bell-shaped curve, we can safely assume that the best value of our measurement is the *average* or *mean* of all of the individual measurements we have made, denoted by $\bar{x}$.



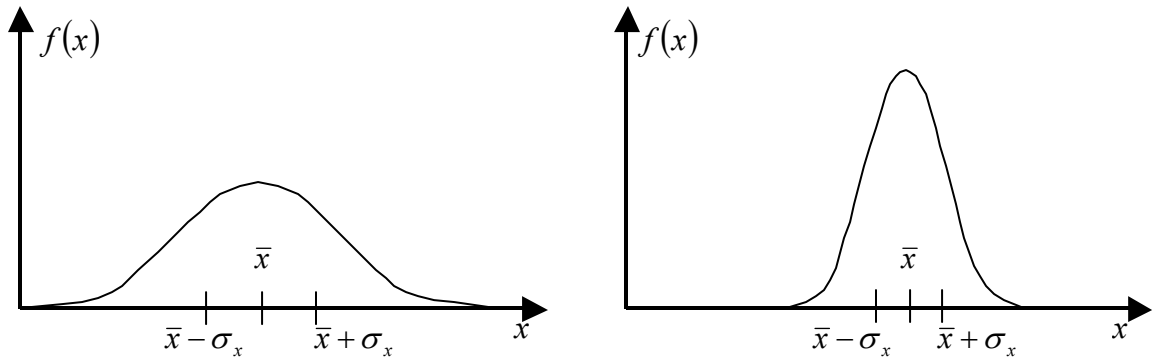
Rule for Averages: The best value of a repeated measurement is the *average* or *mean* of the individual measurements, $\bar{x} = \dfrac{x_1 + x_2 + \ldots + x_N}{N}$, provided the distribution function is a symmetric single peaked function. (But you probably already knew this.)

The standard deviation

Consider the two graphs below. Both are distribution functions $f(x)$ for repeated measurements of *x*, and both have the same average value $\bar{x}$. Clearly, the

difference between them is that one distribution is much wider than the other. The way we describe this is with the standard deviation σ_x of all of the measurements:

$$\sigma_x = \sqrt{\frac{\sum (\bar{x} - x_i)^2}{N}}$$



(You can think of the standard deviation as kind of a "typical difference" between the individual values and the average value. In fact, the equation looks a little like an average in that we're adding up a bunch of things and dividing by *N*, but we have to go through the business of squaring each item and then taking the square root of the sum in order to prevent our "typical difference" from averaging to zero.)

If a measurement is subject to many small random errors, so that the distribution function of the measurement is a nice, smooth, symmetric bell-shaped curve, then the standard deviation tells us where any *single* measurement is likely to fall. Though we won't prove it here, it turns out that for a smooth distribution like that, 68% of the measurements will fall between $\bar{x} - \sigma_x$ and $\bar{x} + \sigma_x$. Stated another way, any *single* measurement has a 68% of falling in the range of $\bar{x} \pm \sigma_x$.

Example: calculating a standard deviation.

Problem: Calculate the standard deviation of the following set of numbers: {3,3,4,4,6}.

Solution: The average of these numbers is

$$\bar{x} = \frac{\sum x_i}{N} = \frac{3+3+4+4+6}{5} = 4,$$

and from here we can calculate the standard deviation:

$$\sigma_x = \sqrt{\frac{\sum (\bar{x} - x_i)^2}{N}}$$
$$= \sqrt{\frac{(4-3)^2 + (4-3)^2 + (4-4)^2 + (4-4)^2 + (4-6)^2}{5}}$$
$$= 1.1$$

You will often see the definition of standard deviation with $N-1$ in the denominator instead of N:

$$\sigma_x = \sqrt{\frac{\sum(\bar{x}-x_i)^2}{N-1}} \quad \text{(alternate definition).}$$

There are subtle differences between the cases when you should use one version or the other, but as long as N is large enough, the difference won't matter much anyway. For purposes of this course, you may use them interchangeably. (If you ever really need to know the difference, plan on either spending an hour curled up with a statistics book or buying a cup of coffee for a friend who already has.)

Rule for Standard Deviations: The standard deviation $\sigma_x = \sqrt{\frac{\sum(\bar{x}-x_i)^2}{N}}$ or $\sigma_x = \sqrt{\frac{\sum(\bar{x}-x_i)^2}{N-1}}$ tells us about the width of a distribution. About 68% of the individual measurements should fall in the range $\bar{x} \pm \sigma_x$.

Standard Error

Suppose you have made a series of repeated measurements and calculated their average and standard deviation. After correctly figuring that your best estimate of the measurement is the average, *you might be tempted to write that your uncertainty in your measurement is the standard deviation, but you would be wrong.*

The uncertainty in your measurements is *not* given by the standard deviation, because the standard deviation doesn't necessarily go down with more measurements. Intuitively, we know that our results should be more and more precise as we average together more and more measurements. But with more and more measurements, the standard deviation stays the same. If we take a million additional measurements of something, the standard deviation might fluctuate either up or down a little bit, but it won't reflect the huge improvement in our precision that should come from averaging a large number of measurements.

The uncertainty of a series of measurements is obtained by dividing the standard deviation by $\sqrt{N}$. This quantity is called the *standard error* or *standard deviation of the mean*:

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

Intuitively, we should be comforted to see that our uncertainty decreases as we average more measurements.

Rule for Standard Errors: The uncertainty of a series of measurements is given by the standard error, $\sigma_{\bar{x}} = \frac{\sigma_x}{\sqrt{N}}$.

Example: what standard deviations and standard errors are good for.

Problem: Moustafa and Marika have just taken 5 measurements of the period of a pendulum using an electronic timer, and they have calculated $\bar{x}$, σ_x, and $\sigma_{\bar{x}}$. (a) What is their best estimate of their period, and its uncertainty? (b) If they took one additional measurement, what range would they expect it to lie in? (c) If they came back the next day to take five additional measurements, what range should they expect the *average* of those next five measurements to lie in?

Solution: (a) Their best estimate of the period is the average (mean) value of their measurements $\bar{x}$, and their uncertainty is the standard error $\sigma_{\bar{x}}$.

(b) If they took one additional measurement, it should fall within one standard deviation of their average: $\bar{x} \pm \sigma_x$. (Most measurements don't fall within the smaller range $\bar{x} \pm \sigma_{\bar{x}}$.)

(c) If they came back the next day to take five more measurements, their result should fall within the range $\bar{x} \pm \sigma_{\bar{x}}$. That's part of what their uncertainty estimate says: that on another day, someone doing the exact same thing should end up with the same answer, to within their uncertainty of the standard error. That is why the standard error is also called the *standard deviation of the mean*: because several different means calculated in the same way would have a distribution for which the standard deviation was $\sigma_{\bar{x}}$.

The Normal Distribution

If a measurement is subject to a large number of small, random uncertainties, it can be shown that the bell-shaped distribution function (as seen on a histogram like the previous graphs) will always have a particular mathematical form called the normal distribution, or Gaussian distribution, which has the form (which you don't have to memorize)

$$f(x) = e^{-(x-\mu_x)^2/2\sigma_x^2},$$

where μ_x is the mean of the distribution and σ_x is the standard deviation.

The normal distribution has many important mathematical properties that provide the theoretical justification for much of the error analysis we have done in this course. The mathematical form of the normal distribution allows us to add uncertainties in quadrature. It provides the justification for using the average as the best statistical estimate of a series of measurements. It's also the reason for the seemingly arbitrary 68% confidence level we adopted as our standard. Our use of the standard deviation and standard error as uncertainty estimates for a single measurement or a group of measurements also depends on a normal distribution. Even the least squares method of fitting data to a straight line in Error Analysis Interlude #4 was based entirely on the assumption that the distribution of the data points around the best fit line was a normal distribution.

Now that we know that all of the error analysis we have done so far is based on the assumption that our errors all follow a normal distribution, we can let you in on the

dirty little secret about error analysis: *most errors probably don't follow a normal distribution*. In fact, most distributions are slightly asymmetric about the center, and some may even be bimodal, or worse. Dealing with these cases could be the basis for many additional Error Analysis Interludes, or even whole courses.

Learning that most errors don't follow a normal distribution is like learning that most surfaces aren't frictionless, air resistance isn't always negligible, and g isn't exactly 9.8 m/sec^2 everywhere. These are still useful approximations to be able to use, but it's important to remember that they are just that: approximations.

In the end, error analysis is all about approximation in the following sense: even our best, most accurate, most precise measurements are only approximations of the physical reality that's out there. All measurements are somewhat uncertain, and error analysis is all about figuring out how much of an approximation our uncertain measurements really are.

The flip side is that measurements are also subject to some degree of certainty: "Error" analysis is also "correctness" analysis. When you make a measurement and calculate your uncertainty, you are putting up a big bold flag that says not only where your uncertainty begins, but also where your certainty begins. Anybody with a theory that relates to your measurement is going to have to deal with you first. Putting a line between what we do know and what we don't know yet is an important part of what science and engineering is all about, and error analysis is one of the ways we do it.