

Appendix A: Error Analysis for the Physics Labs

One of the main goals of the Physics Lab is that you learn about error analysis and the role it plays in experimental science. This appendix briefly reviews the topics you will need to know about error analysis.

(1) The Different types of errors

When you measure something, there are different types of errors you can make:

Illegitimate errors involve making gross mistakes in the experimental setup, in taking or recording data, or in calculating results. Examples of illegitimate errors include: measuring time t when you were supposed to be measuring temperature T , misreading a measurement on a scale so that you think it is 2.0 when it should be 12.0, typing 2.2 into your spreadsheet when you meant to type 20.2, or using the formula "momentum = mv^2 " rather than "momentum = mv ".

Random errors involve errors in measurement due to random changes or fluctuations in the process being measured or in the measuring instrument. Random measuring errors are very common. For example, suppose you measure the length of an object using a ruler and cannot decide whether the length is closer to 10 or 11 mm. If you simply cannot tell which it is closer to, then you will tend to make a random error of about ± 0.5 mm in your choice. Another example of a random error is when you try to read a meter on which the reading is fluctuating. We say that "noise" causes the reading to change with time, and this leads to a random error in determining the true reading of the meter.

A **sampling error** is a special kind of random error that occurs when you make a finite number of measurements of something which can take on a range of values. For example, the students in the university have a range of ages. We could find the exact average age of a student in the University by averaging together the ages of all of the students. Suppose instead that we only took a sub-group, or random sample of students, and averaged together their ages. This sample average would not in general be equal to the exact average age of all of the students in the University. It would tend to be more or less close to the exact average depending on how large or small the group was. The difference between the sample average and the exact average is an example of a sampling error. In general, the larger the sample, the closer the sample average will approach the true average.

A **systematic error** is a repeated and consistent error that occurs in all of your measurements due to the design of the apparatus. Examples of systematic errors include: measuring length with a ruler which is too short, measuring time with a stopwatch which runs too fast, or measuring voltage with a voltmeter which is not properly calibrated. Systematic errors can be very difficult to detect because your results will tend to be consistent, repeatable and precise. The best way to find a systematic error is to compare your results with results from a completely different apparatus.

(2) What is error analysis good for?

If you're like most students who have worked on these labs, you may find yourself wondering why you have to go through all of the trouble of using error analysis. In science and engineering, if you don't understand why you are calculating something, then you really are wasting your time. The most important thing to understand about error analysis is what it can do for you in the lab. There are four main things that you should use error analysis for:

- (i) Finding silly mistakes, such as typing a wrong number into your spreadsheet, using the wrong units, entering a wrong formula, or reading a scale wrong. These illegitimate errors must be corrected before your data can be meaningfully interpreted.
- (ii) Finding a more accurate value for a quantity by making several measurements.
- (iii) Determining the precision of your experimental results.
- (iv) Finding whether your results agree with theory or other experimental results.

It is important to realize that all of the formulas which follow only work for random errors! Despite this, performing error analysis on the random errors in an experiment can often reveal the presence of illegitimate errors. In fact, in these labs you will probably find this to be the most useful thing you can do with error analysis. How to use error analysis to find illegitimate errors is discussed in section 8.

(3) How to Estimate Experimental Errors

Random experimental errors describe our limited ability to measure the true values of physical quantities. To estimate the error in a measurement, you need to consider the measuring equipment that you used and how you actually made the measurement. There are a few cases you will encounter in the labs:

Ruler: One of the most common cases you will encounter involves using a ruler to measure a length. A ruler has a series of marks that are separated by a fixed increment. Suppose that you can very carefully align one end of the object with one of the marks, then the uncertainty in measuring the length of the object comes from the difficulty of figuring out exactly where the other end of the object falls on the ruler. If you can measure to the nearest mark (*i.e.* decide which mark is closer to the length), then the worst mistake you will make is one-half of the distance between the marks. For the rulers used in the lab, the smallest division is usually 1 mm, so you should be able to measure to a precision of about ± 0.5 mm. Since you will not always make the worst mistake, this is actually an overestimate of the typical mistake you can expect to make. A better estimate is about 0.3 mm or about $1/3$ of the distance between the marks. This is called the “ $1/3$ rule”. If you have good eyesight, are careful reading the value, and carefully align the ruler with the object, then you might be able to measure to ± 0.25 mm. To determine whether you should use one half, one-third or one quarter of a unit as the error, you need to consider how carefully you read the scale when you took the data.

Pointer and scale: The next most common case you will encounter is taking a reading off of a scale that has a pointer that indicates the value. Such scales are found on analog voltmeters, ammeters, pressure gauges, and thermometers. The rule for estimating the error in such a “pointer and scale” instrument is exactly the same as for using a ruler. Most people can figure out which mark the pointer is closest to and this implies a worst case error of one-half the

smallest division marked on the scale. A better estimate for the uncertainty would be 1/3 of the increment between divisions on the scale.

Vernier calipers and instruments with vernier scales: A vernier caliper is an instrument for measuring lengths. The operation of this instrument is described in some detail in Appendix B. It makes use of a "vernier scale" which allows one to make more accurate measurements than a simple pointer and scale apparatus. A vernier scale can usually be read to only about ± 1 division.

Digital readouts: The next most common case involves measurements made with a digital readout. These readouts always have a finite number of digits. If the reading is stable, and there are no other sources of error, then the estimated experimental error can be taken as being equal to $\pm 1/2$ unit on the rightmost digit on the scale. If the system is designed such that it displays the digit that is nearest to the true value, then the worst mistake the instrument will make is $\pm 1/2$ unit on the rightmost digit on the scale and a better estimate for the uncertainty is $\pm 1/2$ unit. One needs to be careful however. Often when making a measurement, the rightmost digits on a digital scale will fluctuate randomly due to noise. In this case, the estimated error should be taken as about ± 1 unit on the rightmost digit which does not fluctuate. Also, if you time an event with a stopwatch, you will need to take into account the fact that starting and stopping the timer is not very precise. To determine the error in this case, you will need to take repeated measurements, as discussed below. Also, some instruments change by multiples of ± 1 unit on the smallest scale. To be sure of what your instrument does, you need to check the instrument's operating manual or use the procedure in the next paragraph.

Determining the error from the data: The above techniques work fairly well when you only have one measurement of a quantity. In some of the labs you make many measurements of the same quantity. In this case, it is possible to directly determine the random experimental error in each observation, rather than simply estimating it. Suppose you make N measurements of x , lets call them $x_1, x_2, x_3, x_4, \dots, x_N$. We will generally denote a given measurement, the i -th one, as x_i . If N is large enough, then the experimental error in one measurement can be taken as the standard deviation (see the discussion below on the standard deviation)

$$\Delta x \approx \sigma_x = \left[\frac{1}{(N-1)} \sum_{i=1}^N (x_i - \langle x \rangle)^2 \right]^{1/2}, \quad [\text{A.1}]$$

where $\langle x \rangle$ is the average value of x , and the symbol Σ is the Greek letter Sigma and the notation means that the expression which follows the Σ should be summed up while letting i range from 1 to N . For example, suppose that you made 5 measurements ($N=5$) and $x_1=0, x_2=2, x_3=1.5, x_4=2.5$, and $x_5=0.5$. A simple calculation shows that the average is $\langle x \rangle = 1.3$. The estimated error in each point is thus:

$$\begin{aligned} \Delta x \approx \sigma_x &= \left[\frac{1}{(5-1)} \sum_{i=1}^5 (x_i - \langle x \rangle)^2 \right]^{1/2} \\ &= \left[\frac{1}{4} \left((0-1.3)^2 + (2-1.3)^2 + (1.5-1.3)^2 + (2.5-1.3)^2 + (0.5-1.3)^2 \right) \right]^{1/2} \\ &= \frac{1}{2} \left[(1.3)^2 + (0.7)^2 + (0.2)^2 + (1.2)^2 + (0.8)^2 \right]^{1/2} = 1.04. \end{aligned}$$

Notice that $\Delta x = \pm 1.04$ is quite reasonable since this is about how far each measurement is from the average; $x_1=0$ is 1.3 below the average, $x_2=2$ is 0.7 above the average, $x_3=1.5$ is 0.2 above the average, $x_4=2.5$ is 1.2 above from the average, and $x_5=0.5$ is 0.8 below the average.

Radioactive decay and counting random events: Some experiments involve counting how many random events happen in a certain period of time, for example, counting how many atoms decay in a second in a radioactive material. How can we assign an error to such a measurement? Assigning an error can seem puzzling because in each second there is a definite integer-number of counts. The idea is that if we repeated the measurements many times, we would tend to find a different number of counts in the same time interval, even if the sample and detector were prepared in exactly the same way. By repeating the measurements many times, we could find the average number of counts in each time interval. In general, the average number of decays in a given time interval will be different from the number of decays found in an individual measurement. The difference between the average number of decays in a given interval and the number of decays found in one measurement can be thought of as the error in the measurement. For random decays, the rule is very simple; if you measure N events, then the error in the measurement is $\pm N^{1/2}$. Thus if you measure 100 counts, the error is $\pm 100^{1/2} = \pm 10$ counts.

(4) How to Propagate Errors

Suppose you measure x and t and use these values to calculate a velocity $v = x/t$. If you have errors in x and t , what is the error in v ? This is a problem involving the "propagation of errors". Such problems arise whenever you need to find the error in a quantity which is itself found by combining together measurements which have errors in them.

To proceed let us first consider the simple case where $v = x/t$ and the only error is Δx in x ; i.e. $\Delta t = 0$. Now in general, we do not know the true value of x , but if the error in x were really Δx , then the true value of x would be

$$x_{true} = x + \Delta x.$$

In this case, the true value of the velocity is not $v=x/t$ but rather

$$v_{true} = \frac{(x + \Delta x)}{t}.$$

The difference between the true value of the velocity and our calculated value $v=x/t$ would then be:

$$\Delta v = v_{true} - v = \frac{(x + \Delta x)}{t} - \frac{x}{t} = \frac{\Delta x}{t}. \quad [A.2]$$

This is the correct expression for the error in v . Now notice that

$$\frac{\partial v}{\partial x} = \frac{\partial}{\partial x} \left(\frac{x}{t} \right) = \frac{1}{t}. \quad [A.3]$$

If you are not familiar with the " $\partial/\partial x$ " symbol you need to talk with your TA or professor. It is a partial derivative and means that you take the ordinary derivative of v with respect to x while keeping all other variables fixed. You need to use a partial derivative here because v is a function of more than one variable, and we want to find the change in v produced just when x is changed. Using Equations A.3 and A.2, we can write the error in v as

$$\Delta v = \frac{\partial v}{\partial x} \Delta x. \quad [\text{A.4}]$$

You should recognize this as just the ordinary result from calculus for finding the change in a function v when its argument x changes by a small amount Δx .

Now suppose that there are random errors in both x and t , of magnitude Δx and Δt respectively, and we are trying to find the error in $v = x/t$. A derivation of this result is beyond the scope of this class, and we will simply quote the answer; viz.,

$$\Delta v = \left[\left(\frac{\partial v}{\partial x} \Delta x \right)^2 + \left(\frac{\partial v}{\partial t} \Delta t \right)^2 \right]^{1/2}. \quad [\text{A.5}]$$

There are three things to notice about this expression. First, if $\Delta t = 0$ then it reduces to Equation A.4. Second, this expression does **not** correspond to the usual rule from calculus for finding the change in a function v when x and t change by small amounts. This is because we are assuming that the errors in x and t are random and uncorrelated, so that they can work together or oppose each other in producing changes in v . Finally notice that this expression can be simplified by evaluating the derivatives. We can use A.3 to replace $\partial v/\partial x$ and also use:

$$\frac{\partial v}{\partial t} = -\frac{x}{t^2}. \quad [\text{A.6}]$$

We can then rewrite Equation A.5 as

$$\begin{aligned} \Delta v &= \left[\left(\frac{1}{t} \Delta x \right)^2 + \left(\frac{-x}{t^2} \Delta t \right)^2 \right]^{1/2} = \left[\left(\frac{x \Delta x}{t^2} \right)^2 + \left(\frac{x \Delta t}{t^2} \right)^2 \right]^{1/2} \\ &= \left[\left(v \frac{\Delta x}{x} \right)^2 + \left(v \frac{\Delta t}{t} \right)^2 \right]^{1/2} = v \left[\left(\frac{\Delta x}{x} \right)^2 + \left(\frac{\Delta t}{t} \right)^2 \right]^{1/2}. \end{aligned} \quad [\text{A.7}]$$

The above ideas can be generalized to include functions with errors in an arbitrary number of arguments. For example, if f is a function of x , y , z , t , r , and B , and these have random errors Δx , Δy , Δz , Δt , Δr , and ΔB respectively, then the error in f is just

$$\Delta f = \left[\left(\frac{\partial f}{\partial x} \Delta x \right)^2 + \left(\frac{\partial f}{\partial y} \Delta y \right)^2 + \left(\frac{\partial f}{\partial z} \Delta z \right)^2 + \left(\frac{\partial f}{\partial t} \Delta t \right)^2 + \left(\frac{\partial f}{\partial r} \Delta r \right)^2 + \left(\frac{\partial f}{\partial B} \Delta B \right)^2 \right]^{1/2}. \quad [\text{A.8}]$$

The following examples illustrate some special cases that you will encounter in the labs.

- (i) Suppose $f = ax + b$, where a and b are constants and x has an uncertainty Δx . The uncertainty in f can be found from Equation A.8 by noting that x is the only variable and $\partial f/\partial x = a$. Thus,

$$\Delta f = \left[\left(\frac{\partial f}{\partial x} \Delta x \right)^2 \right]^{1/2} = \left[(a \Delta x)^2 \right]^{1/2} = a \Delta x.$$

- (ii) Suppose $f = x + y$, where x has error Δx and y has error Δy . Then $\partial f/\partial x = 1$ and $\partial f/\partial y = 1$ and

$$\Delta f = \left[\left(\frac{\partial f}{\partial x} \Delta x \right)^2 + \left(\frac{\partial f}{\partial y} \Delta y \right)^2 \right]^{1/2} = \sqrt{(\Delta x)^2 + (\Delta y)^2}.$$

(iii) Suppose $f = x^n y^m$, where n and m are constants. The derivatives are just $\partial f / \partial x = n x^{n-1} y^m = n f / x$ and $\partial f / \partial y = m x^n y^{m-1} = m f / y$. Thus,

$$\Delta f = \left[\left(\frac{\partial f}{\partial x} \Delta x \right)^2 + \left(\frac{\partial f}{\partial y} \Delta y \right)^2 \right]^{1/2} = \left[\left(\frac{n f}{x} \Delta x \right)^2 + \left(\frac{m f}{y} \Delta y \right)^2 \right]^{1/2} = f \left[n^2 \left(\frac{\Delta x}{x} \right)^2 + m^2 \left(\frac{\Delta y}{y} \right)^2 \right]^{1/2}.$$

(5) The mean value $\langle x \rangle$ and the error in the mean $\Delta \langle x \rangle$

Suppose you make N measurements of the quantity x and denote the result of the first measurement by x_1 , the second measurement by x_2, \dots and the N -th measurement by x_N . The average or mean value of x is denoted by $\langle x \rangle$ and is just:

$$\langle x \rangle = \frac{1}{N} \sum_{i=1}^N x_i. \quad [\text{A.9}]$$

Excel Tip: In Excel you can use the command "=average" (.) to automatically calculate the mean or average of a set of data. For example, suppose you wanted to find the mean of some data that was in cells D13 to D27. You would enter the command =average(D13:D27) in the cell where you want the average value to appear. This is nice because Excel figures out how many points N there are and you don't have to keep track.

Notice that the above definition of $\langle x \rangle$ is just our normal definition of the average of a set of numbers. Why is the mean value important? It turns out that if all of the measurements have the same experimental uncertainty, then the mean value is the best estimate of the true value of x .

The error in the mean value can be found by propagating errors, as in section 4. Assuming that each of the measurements x_i are independent variables, and that the error in each measurement is Δx , one finds

$$\Delta \langle x \rangle = \frac{\Delta x}{\sqrt{N}}. \quad [\text{A.10}]$$

This result says that the error in the mean value, $\Delta \langle x \rangle$, is smaller than the error in any one measurement Δx , by a factor of $N^{1/2}$. For example, suppose you make 100 measurements. The error in the mean, $\Delta \langle x \rangle$, will be 10 times smaller than the error Δx in an individual measurement. What this means is that you can obtain very precise measurements, even with imprecise instruments, provided you take many data points.

(6) The standard deviation.

Suppose you make N measurements of the quantity x and denote the result of the first measurement by x_1 , the second measurement by x_2, \dots and the N -th measurement by x_N . The standard deviation of x is denoted by σ_x and is defined to be

$$\sigma_x = \left[\frac{1}{N-1} \sum_{i=1}^N (x_i - \langle x \rangle)^2 \right]^{1/2}. \quad [\text{A.11}]$$

Excel Tip: In Excel you can use the command "=STDEV(..)" to automatically calculate the standard deviation of a set of data. For example, suppose you wanted to find the standard deviation of some data which was in cells D13 to D27. You would enter the command =STDEV(D13..D27) in the cell where you want the standard deviation to appear.

The standard deviation tells you how far a typical data point is from the average. If your data has a lot of scatter in it, then you will find a large standard deviation. If all of your measurements are practically identical, then the standard deviation will be quite small. Typically you would expect that a given measured value of x might be different from the true value by about Δx , the uncertainty in the measurement. Thus if the average value is a good estimate of the true value, you expect

$$x_i - \langle x \rangle \approx \Delta x$$

Substituting this into Equation A.8, one finds that for large N

$$\sigma_x \approx \left[\frac{1}{N-1} \sum_{i=1}^N (\Delta x)^2 \right]^{1/2} = \left[\frac{N}{N-1} (\Delta x)^2 \right]^{1/2} = (\Delta x) \left[\frac{N}{N-1} \right]^{1/2} \approx \Delta x.$$

Which is where Equation A.1 came from.

There are some things about σ_x which can be confusing:

- (i) Why is there a factor of (N-1) in the denominator instead of N? To understand why there is an N-1, suppose we make just one measurement; call it x_1 . In this case the average is simply $\langle x \rangle = x_1$. Notice however that Equation C.8 says that the standard deviation is undefined because $N=1$. The reason it is undefined is because with only one measurement, it is not possible to say how much spread there is in the data. That requires at least two measurements.
- (ii) Why do we take the square each of the terms? If we did not take the square, but just added together all of the terms $x_i - \langle x \rangle$, we would get zero. To see this, just look at the definition of $\langle x \rangle$! The point is that data which falls below the average is balanced by data which falls above that average. Roughly speaking, by taking the square, we make all the terms positive and end up finding the (root mean square) distance of a typical data point from the average, independent of whether it is above or below the average.
- (iii) Does σ_x get bigger if we measure more data points? No. The standard deviation does not tend to get bigger (or smaller) as you take more data points. Physically speaking, σ_x is just the typical distance a data point is from the average. As you take more data, you tend to get a more accurate value for the true value of σ_x , not a bigger value.

(7) The weighted mean $\langle x \rangle$ and the error in the weighted mean $\Delta \langle x \rangle$

In some of the labs, you will need to find the best estimate for a measured parameter by combining together measurements which have different sizes of errors. For example, in 261 Lab 2, some of your measurements will be made with a ruler and some with vernier calipers. The measurements made with the calipers will be much more accurate than those taken with the ruler. If we want to combine together data from measurements with different errors, we need to use the weighted average, which is defined by

$$\langle x \rangle = \frac{\sum_{i=1}^N \left[\frac{x_i}{(\Delta x)_i^2} \right]}{\sum_{i=1}^N \left[\frac{1}{(\Delta x)_i^2} \right]}. \quad [\text{A.12}]$$

Notice that in this expression, each measurement gets multiplied by $1/\Delta x_i^2$ before it is added to the other measurements. The factor $1/\Delta x_i^2$ can be thought of as the importance or "weight" of the measurement. Thus if $\Delta x_1 = 1$ mm and $\Delta x_2 = 0.1$ mm, then the second measurement is 100 times more important than the first. From this you can see that it really pays to make more accurate measurements, they carry a lot of weight! If we want to simplify Equation A.12, we can define the weight of the i -th measurement as w_i , where

$$w_i = 1/\Delta x_i^2, \quad [\text{A.13}]$$

and rewrite Equation A.12 as

$$\langle x \rangle = \frac{\sum_{i=1}^N w_i x_i}{\sum_{i=1}^N w_i}. \quad [\text{A.14}]$$

Notice that the denominator is just the sum of all of the weights, so that it acts to normalize out the total weight of all the measurements.

The error in the weighted mean is given by

$$\Delta \langle x \rangle = \frac{1}{\left[\sum_{i=1}^N \left[\frac{1}{(\Delta x)_i^2} \right] \right]^{1/2}} = \frac{1}{\left[\sum_{i=1}^N w_i \right]^{1/2}}. \quad [\text{A.15}]$$

Notice that the error in the weighted mean is just the square root of the same term which appears in the denominator of the weighted average. If you want to remember this result, notice that it is just says that the error in the average is one over the square root of the total weight. Squaring and rearranging would give you an equation which says that the total weight is one over the square of the error in the mean. This is the same relationship as Equation A.13, except now it is for the mean rather than an individual measurement.

Excel Tip: Excel does not have a command for calculating the weighted mean or the error in the weighted mean. The easiest way to calculate the weighted mean is to set up three columns, the first with the data in it, the second with 1 over the square of the error in each measurement, and the third

with the product of the first and second columns. To get the weighted mean you then sum the third column and divide by the sum of the second column.

(8) Everything you need to know about χ^2 .

This section briefly discusses everything you need to know about χ^2 .

How to speak like a native

The symbol χ is the Greek letter chi. It is pronounced “kie” and rhymes with “pie”. Thus χ^2 is pronounced “kie-squared”. Similarly, the symbol ν is the Greek letter nu and is pronounced “new”. Thus the expression χ^2/ν is pronounced “kie-squared over new”.

How to calculate χ^2

In order to calculate, you need three things: N measured data points (x_1, x_2, \dots, x_N), a theory which tells you how big each of the data points was supposed to be (we'll call this $x_{i,theory}$ for the i-th data point), and estimated errors for each measurement ($\Delta x_1, \Delta x_2, \dots, \Delta x_N$). χ^2 can then be found from the formula

$$\chi^2 = \sum_{i=1}^N \left[\frac{x_i - x_{i,theory}}{\Delta x_i} \right]^2.$$

For example, suppose that you made 5 measurements ($N=5$, $x_1=1.1$, $x_2=1.2$, $x_3=1.5$, $x_4=1.5$, and $x_5=1.3$), that the theory says that x should have been 1.25, and that the estimated error in each measurement was $\Delta x_i=0.15$. Then

$$\begin{aligned} \chi^2 &= \sum_{i=1}^5 \left[\frac{x_i - 1.25}{0.15} \right]^2 \\ &= \left[\frac{1.1 - 1.25}{0.15} \right]^2 + \left[\frac{1.2 - 1.25}{0.15} \right]^2 + \left[\frac{1.5 - 1.25}{0.15} \right]^2 + \left[\frac{1.5 - 1.25}{0.15} \right]^2 + \left[\frac{1.3 - 1.25}{0.15} \right]^2 = 6.74 \end{aligned}$$

Excel Tip: Excel does not have a command for calculating χ^2 , but the easiest way to calculate it is to set up five columns, the first with the data in it, the second with the theory in it, the third with the difference between the theory and data, the fourth with the error in it, and the fifth with the square of the third column divided by the square of the error in the data. To get χ^2 , you then sum the last column. For example:

i	x_i	$x_{i,theory}$	$x_i - x_{i,theory}$	Δx	$(x_i - x_{i,theory})^2 / \Delta x^2$
1	1.1	1.25	0.15	0.15	1
2	1.2	1.25	0.05	0.15	0.09
3	1.5	1.25	0.25	0.15	2.78
4	1.5	1.25	0.25	0.15	2.78
5	1.3	1.25	0.05	0.15	0.09
					sum = $\chi^2 = 6.74$

The degrees of freedom ν

In order to use χ^2 you also need to know ν , the "degrees of freedom" in your experiment. In these labs there are only two cases of practical importance:

- (i) You did not use any fitting parameters or averages to compute the theoretical values used in χ^2 . If the theory is given and you did not use any of your data to fit the theory to the data, then the degrees of freedom is equal to the number N of data points, i.e. $\nu=N$.
- (ii) Fitting parameters or averages were used to find the theory. If you computed your theoretical values (used in χ^2) by averaging your N data points or by using a fitting routine, then $\nu = N-\alpha$, where α is the number of fitting parameters you used. For example, if $x_{\text{theory}} = \langle x \rangle$ then you computed one parameter, the average, from your data, so that $\alpha = 1$ and $\nu = N-1$. If you used Excel to fit your data to a straight line, and used the slope and intercept of the line to compute theoretical values, then you used two fitting parameters (slope and intercept), so that $\nu = N-2$.

χ^2/ν and what it tells you about your experiment

Once you have found χ^2 and ν , it is a simple matter to divide the two and find χ^2/ν the "reduced value of χ^2 ". What makes χ^2/ν so important is what it can tell you about your experiment. There are three cases:

- (i) $\chi^2/\nu \approx 1$. This means that your results and theory are consistent to within your experimental errors.
- (ii) $\chi^2/\nu \gg 1$. This means that something is wrong! There are three possibilities:
 - (a) You have made an illegitimate error in measuring the data, calculating the theory, recording your data or errors, or analyzing your data. To determine if this is what happened, you need to go back and look at your data and theory and see if the numbers are reasonable. Be especially careful of units.
 - (b) You have underestimated the size of the errors you are making in your measurements, or, the quantity you are trying to measure has a distribution of possible values. To determine if this is what happened, look at the scatter in your data and see if it is much larger than your estimated error.
 - (c) If you can rule out (a) and (b) above, then having χ^2/ν much bigger than 1 means that the theory is wrong.
- (iii) $\chi^2/\nu \ll 1$. This also means that something is wrong! There are only two possibilities:
 - (a) You have overestimated the size of your errors, that is, you have actually measured things much better than your claimed errors. You need to go back and get a more accurate, and less conservative, estimate of your errors.
 - (b) You have made an illegitimate error in calculating χ^2/ν . Most likely you have neglected units or made a simple computational error.

What is $P(\chi^2, v)$ and how do you find it

You may be wondering how much bigger or smaller than 1 that χ^2/v has to be for something to be clearly wrong, or how close to 1 does χ^2/v have to be for you to say that the theory and experiment agree. This can only be answered by considering the quantity $P(\chi^2/v)$. Suppose you have found a value for χ^2 and that you have v degrees of freedom. Then $P(\chi^2, v)$ is the probability that random errors will cause χ^2 to be larger than you found.

Excel Tip: The easiest way to get $P(\chi^2, v)$ is directly from Excel by typing in the command =CHIDIST(χ^2, v). For example, if your value of χ^2 is 3 and you have five degrees of freedom, then you would type in =CHIDIST (3,5). See the help manual in Excel for more information.

Values does $P(\chi^2, v)$ can also be obtained from the following table.

		$P(\chi^2, v)$									
		0.99	0.95	0.9	0.70	0.50	0.30	0.10	0.05	0.01	0.001
v	1	0.00016	0.0040	0.0518	0.148	0.455	1.074	2.706	3.841	6.635	10.83
	2	0.0100	0.0515	0.105	0.357	0.693	1.204	2.303	2.996	4.605	6.908
	3	0.0383	0.117	0.195	0.475	0.789	1.222	2.084	2.605	3.780	5.423
	4	0.0742	0.178	0.265	0.549	0.839	1.220	1.945	2.372	3.319	4.617
	5	0.111	0.229	0.322	0.600	0.870	1.213	1.847	2.214	3.017	4.102
	6	0.145	0.273	0.367	0.638	0.891	1.205	1.774	2.099	2.802	3.743
	7	0.177	0.310	0.405	0.667	0.907	1.198	1.717	2.010	2.639	3.475
	8	0.206	0.342	0.436	0.691	0.918	1.191	1.670	1.938	2.511	3.266
	9	0.232	0.369	0.463	0.710	0.927	1.184	1.632	1.880	2.407	3.097
	10	0.256	0.394	0.487	0.727	0.934	1.178	1.599	1.831	2.321	2.959
	12	0.298	0.436	0.525	0.751	0.945	1.168	1.546	1.752	2.185	2.740
	15	0.349	0.484	0.570	0.781	0.956	1.155	1.487	1.666	2.039	2.513
	20	0.413	0.543	0.622	0.813	0.967	1.139	1.421	1.571	1.878	2.266
	30	0.498	0.616	0.687	0.850	0.978	1.118	1.342	1.459	1.696	1.990
	50	0.594	0.695	0.754	0.886	0.987	1.038	1.26	1.350	1.523	1.733

To find $P(\chi^2, v)$ from the table by using the following procedure:

- The left hand column lists v , the degrees of freedom. Find the row which has the degrees of freedom for your experiment.
- Next, proceed from left to right along this row until you reach the cell which has your value of χ^2/v . Notice that the values increase as you go along a row from left to right. If your value of χ^2/v isn't listed in the table, just choose the cell in your row which has the value which is closest to your value of χ^2/v .
- Once you have found the cell with your value of χ^2/v , look at the cell at the very top of this column. This is the value of $P(\chi^2, v)$ for your χ^2 and v .

For example, suppose that you have $N=9$ measurements, you used one fitting parameter, and you computed $\chi^2=12$. The degrees of freedom is thus $v=N-1=9-1=8$ and the reduced χ^2 is $\chi^2/v=12/8=1.5$. Since $v=8$, we first find the $v=8$ row. Looking along this row, we see the values 0.206, 0.342, 0.436, 0.691, 0.918, 1.191, 1.670, etc. Since our value of χ^2/v is 1.5, the closest value listed in the row is 1.670. Looking at the top of this column we find the number 0.10. Thus $P(\chi^2=12, v=8)$ is equal to about 0.1. This means that there is about a 10% chance that random errors would produce a value of χ^2 which is larger than 12.

What does $P(\chi^2, v)$ tell you about your experiment

As discussed above, the value of χ^2/v can tell you if something is wrong with your experiment or if it agrees with theory. $P(\chi^2, v)$ does the same thing, except that it provides a more precise statement of the agreement between theory and experiment. This section uses a few examples to help you understand $P(\chi^2, v)$.

- (i) $P(\chi^2, v) \approx 0.5$ means everything is OK. Suppose you have 8 degrees of freedom ($v = 8$) and have found that $\chi^2 = 8$ so that $\chi^2/v = 1$. From our discussion of χ^2/v above, we know that when χ^2/v is close to 1, our results are consistent with the theory. We expect then that $P(\chi^2, v)$ should tell us the same thing. If you use table above, you will find that $P(\chi^2 = 8, v = 8) \approx 0.40$ (because $\chi^2/v = 1$ is about halfway between the $P = 0.3$ and $P = 0.5$ columns). What this means is that if you repeated the experiment many times, you would find that 40% of the time your random errors (with the size you used in calculating χ^2) will produce a χ^2 which is bigger than 8. Logically this also means that 60% of the time random errors will produce a χ^2 which is smaller than 8. This means that random errors are just about as likely to produce a bigger or a smaller value of χ^2 . If you think about this, you will realize that this is just what you would expect if the theory is in agreement with the data.
- (ii) If $P(\chi^2, v)$ is smaller than 0.05 then something may be wrong. Suppose you have 8 degrees of freedom ($v = 8$) and have found that $\chi^2 = 16$ so that $\chi^2/v = 2$. From our discussion of χ^2/v above, we know that when χ^2/v is much bigger than 1, our results are not consistent with the theory. If you use the table above, you would find that $P(\chi^2 = 16, v = 8) \approx 0.05$. What this means is that if you repeated the experiment many times, you would find that about 5% of the time your random errors (with the size you used in calculating χ^2) will produce a χ^2 which is bigger than 16. Logically, this also means that 95% of the time random errors should produce a χ^2 which is smaller than 16. This means that random errors are much more likely to produce a value of χ^2 which is smaller than you found. Another way to say this is that 5% isn't very big, so that it is rather unlikely that random errors could have produced a χ^2 as large as you found. This is equivalent to saying that the theory is not in good agreement with the data.
- (iii) If $P(\chi^2, v)$ exceeds about 0.95 then something may be wrong. Suppose you have 8 degrees of freedom ($v = 8$) and have found that $\chi^2 = 2.7$ so that $\chi^2/v = 0.33$. From our discussion of χ^2/v above, we know that when χ^2/v is much smaller than 1, we have probably made an illegitimate error or used too large of an error estimate. From the table, you will find that $P(\chi^2 = 2.4, v = 8) \approx 0.95$. What this means is that if you repeated the experiment many times, about 95% of the time your random errors (with the size you used in calculating χ^2) will produce a χ^2 which is bigger than 2.7. Logically, this also means that only 5% of the time random errors should produce a χ^2 which is smaller than 2.7. This means that random errors are much more likely to produce a value of χ^2 which is larger than you found. Another way to say this is that you should have expected random errors to produce a value of χ^2 which was larger than you found. This is equivalent to saying that the data has too little scatter in it to be consistent with your estimated errors, and this suggests that something may be wrong with your estimated errors or your analysis.

There are a few other things that may be useful to understand about $P(\chi^2, \nu)$. First, it provides the same information as χ^2/ν , but it does a better job. In particular, it tells you precisely how much bigger or smaller χ^2/ν has to be compared to 1 to decide something is wrong. Second, $P(\chi^2, \nu)$ must be understood statistically. If you find $P(\chi^2=8, \nu=8)=0.5$, this does not mean that the theory is definitely correct. Rather, it means that the theory is statistically consistent with the data. Similarly, $P(\chi^2, \nu)=0.01$ does not mean that your theory is definitely wrong or that you have necessarily done anything wrong; in fact, 1% of the time you should expect to find $P(\chi^2, \nu)=0.01$. It just says that you wouldn't expect this to happen more than 1% of the times you tried the experiment, so you better check things to see if something else isn't wrong.

The 5% to 95% confidence limits

Finally, the standard or conventional choice is to say that if $P(\chi^2, \nu)$ is smaller than 0.05 or larger than 0.95, then this signifies a significant disagreement between theory, data and errors. That is why we chose the specific values of 0.05 and 0.95 in the above discussion of the meaning of $P(\chi^2, \nu)$. In other words, if you find $0.05 < P(\chi^2, \nu) < 0.95$ then there is not a statistically significant disagreement between your data and the theory and you can say that your data and theory are in agreement, given the size of your error bars. On the other hand, if you find $P(\chi^2, \nu) > 0.95$ or $P(\chi^2, \nu) < 0.05$ then there is a significant disagreement between your data and the theory or accepted value.

χ^2 and the two-sigma rule

Suppose you have measured a value for a parameter x to an uncertainty σ_x and need to figure out if it agrees with an accepted value x_{theory} . From these three values you can calculate χ^2 using:

$$\chi^2 = \left[\frac{x - x_{theory}}{\sigma_x} \right]^2$$

In this case there is just one degree of freedom so that $\nu = 1$. For the data and the theory to be in agreement means you would need to have $P(\chi^2, \nu) > 0.05$. Examining the table above for $\nu = 1$, we see that $P(\chi^2, \nu) > 0.05$ means that $\chi^2 < 3.841$. Plugging this into the above expression for χ^2 gives:

$$\left[\frac{(x - x_{theory})}{\sigma_x} \right]^2 < 3.841$$

which gives:

$$|x - x_{theory}| < 1.96\sigma_x \approx 2\sigma_x$$

This is the “**two-sigma rule**”. It says that a measured value and the accepted value are in agreement provided they differ by no more than about twice the uncertainty in the measured value. Another way to say this is that if your data point is no more than two error bars from the accepted value then the disagreement is not significant.

(9) Summary of Some Error Analysis Results

In the following formulas, x_i denotes the i -th measurement of the quantity x , Δx_i denotes the estimated random experimental error in x_i , and a total of N measurements were made.

Propagation of Errors; suppose f is a function of the variables x_1, x_2, \dots, x_N .

The error in f due to errors in x_1, x_2, \dots, x_N is

$$\Delta f = \sqrt{\sum_{i=1}^N \left[\frac{\partial f}{\partial x_i} \Delta x_i \right]^2}$$

Standard Deviation, in Excel use STDEV(..),

$$\sigma_x = \left[\frac{1}{N-1} \sum_{i=1}^N (x_i - \langle x \rangle)^2 \right]^{1/2}.$$

Note: For most experiments, one expects that $\sigma_x \approx \Delta x_i$, i.e. the standard deviation is about equal to the estimated experimental error in one measurement of x .

Mean, in Excel use AVERAGE(..),
(use if each measurement has the same experimental error)

$$\langle x \rangle = \frac{1}{N} \sum_{i=1}^N x_i$$

Error in the Mean

(each measurement has experimental error Δx):

$$\Delta \langle x \rangle = \frac{\Delta x}{\sqrt{N}}$$

Weighted Mean:

$$\langle x \rangle = \frac{\sum_{i=1}^N \left[\frac{x_i}{[\Delta x_i]^2} \right]}{\sum_{i=1}^N \left[\frac{1}{[\Delta x_i]^2} \right]}$$

Error in Weighted Mean:

$$\Delta \langle x \rangle = \frac{1}{\left[\sum_{i=1}^N \left[\frac{1}{[\Delta x_i]^2} \right] \right]^{1/2}}$$

χ^2 (chi-squared):

$$\chi^2 = \sum_{i=1}^N \left[\frac{x_i - x_{theory}}{\Delta x_i} \right]^2$$

reduced χ^2 :

$$\frac{\chi^2}{\nu} \quad \dots \text{ should equal about 1 for a good fit}$$

Degrees of freedom in χ^2 $\nu = N - \alpha$ (where N is the number of measurements and α is the number of fitting parameters used to find the theoretical value from the data)

Data and theory are in agreement if $0.05 < P(\chi^2, \nu) < 0.95$

“**two-sigma rule**”: a measurement and accepted value are in agreement if they differ by less than two error bars.