Measurement and Uncertainty Analysis Guide

"It is better to be roughly right than precisely wrong." – Alan Greenspan

Table of Contents

The Uncertainty of Measurements

Some numerical statements are exact: Mary has 3 brothers, and 2 + 2 = 4. However, all *measurements* have some degree of uncertainty that may come from a variety of sources. The process of evaluating the uncertainty associated with a measurement result is often called *uncertainty analysis* or sometimes *error analysis*.

The complete statement of a measured value should include an estimate of the level of confidence associated with the value. Properly reporting an experimental result along with its uncertainty allows other people to make judgments about the quality of the experiment, and it facilitates meaningful comparisons with other similar values or a theoretical prediction. Without an uncertainty estimate, it is impossible to answer the basic scientific question: "Does my result agree with a theoretical prediction or results from other experiments?" This question is fundamental for deciding if a scientific hypothesis is confirmed or refuted.

When making a measurement, we generally assume that some exact or true value exists based on how we define what is being measured. While we may never know this true value exactly, we attempt to find this ideal quantity to the best of our ability with the time and resources available. As we make measurements by different methods, or even when making multiple measurements using the same method, we may obtain slightly different results. So how do we report our findings for our best estimate of this elusive *true value*? The most common way to show the range of values that we believe includes the true value is:

$$\text{measurement} = (\text{best estimate} \pm \text{uncertainty}) \text{ units}$$

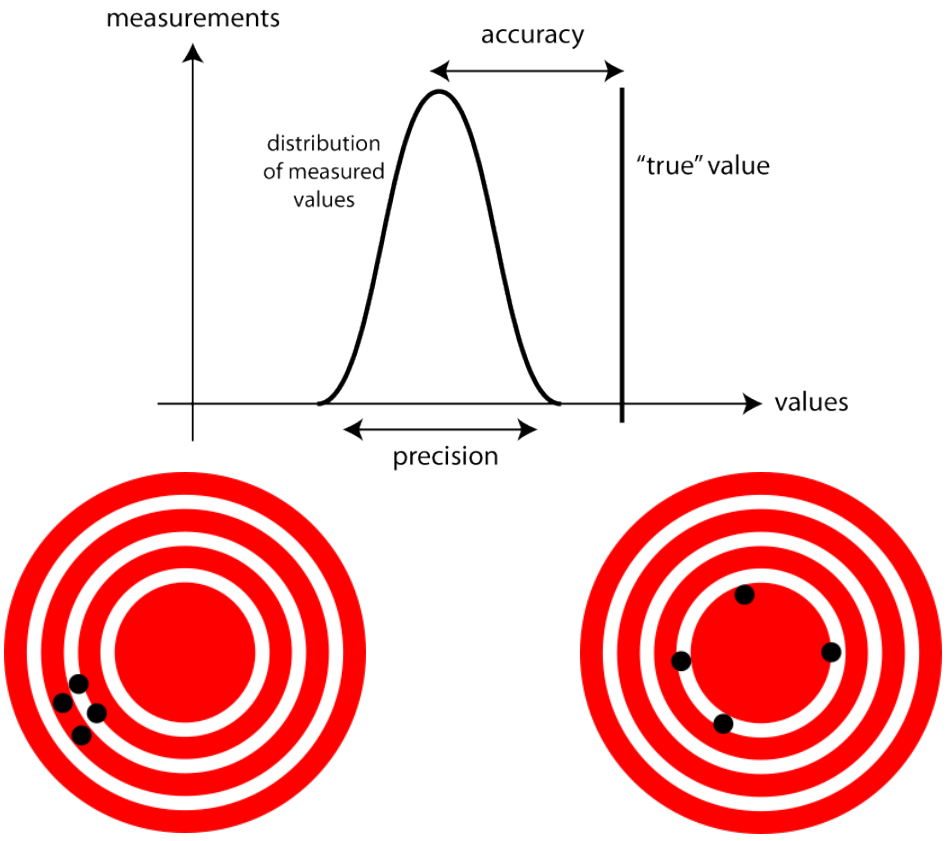
As an example, suppose you want to find the mass of a gold ring that you would like to sell to a friend. You do not want to jeopardize your friendship, so you want to get an accurate mass of the ring in order to charge a fair market price. You estimate the mass to be between 10 and 20 grams from how heavy it feels in your hand, but this is not a very precise estimate. After some searching, you find an electronic balance that gives a mass reading of 17.43 grams. While this measurement is much more *precise* than the original estimate, how do you know that it is *accurate*, and how confident are you that this measurement represents the true value of the ring's mass? Since the digital display of the balance is limited to 2 decimal places, you could report the mass as $m = 17.43 \pm 0.01$ g. Suppose you use the same electronic balance and obtain several more readings: 17.46 g, 17.42 g, 17.44 g, so that the average mass appears to be in the range of 17.44 ± 0.02 g. By now you may feel confident that you know the mass of this ring to the nearest hundredth of a gram, but how do you know that the true value definitely lies between 17.43 g and 17.45 g? Since you want to be honest, you decide to use another balance that gives a reading of 17.22 g. This value is clearly below the range of values found on the first balance, and under normal circumstances, you might not care, but you want to be

fair to your friend. So what do you do now? The answer lies in knowing something about the accuracy of each instrument.

To help answer these questions, we first define the terms *accuracy* and *precision*:

> **Accuracy** is the closeness of agreement between a measured value and a true or accepted value. Measurement *error* is the amount of inaccuracy.
>
> **Precision** is a measure of how well a result can be determined (without reference to a theoretical or true value). It is the degree of consistency and agreement among independent measurements of the same quantity; also the reliability or reproducibility of the result.

The accuracy and precision can be pictured as follows:



high precision, low accuracy

low precision, high accuracy

Figure 1. Accuracy vs Precision

The *uncertainty* estimate associated with a measurement should account for both the accuracy and precision of the measurement. **Precision** indicates the quality of the measurement, without any guarantee that the measurement is "correct." **Accuracy**, on the other hand, assumes that there is an ideal "true" value, and expresses how far your answer is from that "correct" answer. These concepts are directly related to *random* and *systematic* measurement uncertainties (next section).

Note: Unfortunately the terms *error* and *uncertainty* are often used interchangeably to describe both imprecision and inaccuracy. This usage is so common that it is impossible to avoid entirely. Whenever you encounter these terms, make sure you understand whether they refer to accuracy or precision, or both. In this document, we will emphasize the term "uncertainty" but will use the term "error," as necessary, to avoid confusion with commonly found examples and standard usage of the term.

In order to determine the *accuracy* of a particular measurement, we have to know the ideal, true value, sometimes referred to as the "gold standard." Sometimes we have a "textbook" measured value, which is well known, and we assume that this is our "ideal" value, and use it to estimate the *accuracy* of our result. Other times we know a theoretical value, which is calculated from basic principles, and this also may be taken as an "ideal" value. But physics is an empirical science, which means that the theory must be validated by experiment, and not the other way around. We can escape these difficulties and retain a useful definition of *accuracy* by assuming that, even when we do not know the true value, we can rely on the best available *accepted* value with which to compare our experimental value.

For the gold ring example, there is no accepted value with which to compare, and both measured values have the same precision, so there is no reason to believe one more than the other. We could look up the accuracy specifications for each balance as provided by the manufacturer, but the best way to assess the accuracy of a measurement is to compare it with a known **standard**. For this situation, it may be possible to calibrate the balances with a standard mass that is accurate within a narrow tolerance and is traceable to a **primary mass standard** at the National Institute of Standards and Technology (NIST). Calibrating the balances should eliminate the discrepancy between the readings and provide a more *accurate* mass measurement.

Precision is often reported quantitatively by using **relative** or **fractional uncertainty**:

$$\text{Relative Uncertainty} = \left| \frac{\text{uncertainty}}{\text{measured quantity}} \right| \qquad (1)$$

For example, $m = 75.5 \pm 0.5$ g has a fractional uncertainty of: $\frac{0.5g}{75.5g} = 0.00\overline{6} = 0.7\%$

Accuracy is often reported quantitatively by using **relative error**:

$$\text{Relative Error} = \frac{\text{measured value - expected value}}{\text{expected value}} \qquad (2)$$

If the expected value for *m* is 80.0 g, then the relative error is: $\frac{75.5 - 80.0}{80.0} = -0.056 = -5.6\%$

Critical Notes:

- The minus sign indicates that the measured value is *less* than the expected value – *unless explicitly stated, the term "relative error" does not in and of itself refer to a magnitude*.
- The denominator is neither the measured value nor the average of the measured and expected value – *the relative error can only be cited when there is a known expected value or gold standard.*

Types of Uncertainty

Measurement uncertainties may be classified as either **random** or **systematic**, depending on how the measurement was obtained (an instrument could cause a random uncertainty in one situation and a systematic uncertainty in another).

Random uncertainties are statistical fluctuations (in either direction) in the measured data. These uncertainties may have their origin in the measuring device, or in the fundamental physics underlying the experiment. The random uncertainties may be masked by the precision or accuracy of the measurement device. Random uncertainties can be evaluated through statistical analysis and can be reduced by averaging over a large number of observations (see "standard error" later in this document).

Systematic uncertainties are reproducible inaccuracies that are consistently in the "same direction," and could be caused by an artifact in the measuring instrument, or a flaw in the experimental design (because of these possibilities, it is not uncommon to see the term "systematic error"). These uncertainties may be difficult to detect and cannot be analyzed statistically. If a systematic uncertainty or error is identified when calibrating against a standard, applying a correction or correction factor to compensate for the effect can reduce the bias. Unlike random uncertainties, systematic uncertainties cannot be detected or reduced by increasing the number of observations.

When making careful measurements, the goal is to reduce as many sources of uncertainty as possible and to keep track of those that cannot be eliminated. It is useful to know the types of uncertainties that may occur, so that we may recognize them when they arise. Common sources of uncertainty in physics laboratory experiments include:

Incomplete definition (may be systematic or random) - One reason that it is impossible to make exact measurements is that the measurement is not always clearly defined. For example, if two different people measure the length of the same string, they would probably get different results because each person may stretch the string with a different tension. The best way to minimize definition uncertainty is to carefully consider and specify the conditions that could affect the measurement.

Failure to account for a factor (usually systematic) – The most challenging part of designing an experiment is trying to control or account for all possible factors except the one independent variable that is being analyzed. For instance, you may inadvertently ignore air resistance when measuring free-fall acceleration, or you may fail to account for the effect of the Earth's magnetic field when measuring the field near a small magnet. The best way to account for these sources of uncertainty is to brainstorm with your peers about all the factors that could possibly affect your

result. This brainstorm should be done *before* beginning the experiment in order to plan and account for the confounding factors before taking data. Sometimes a *correction* can be applied to a result *after* taking data to account for an uncertainty that was not detected earlier.

Environmental factors (systematic or random) - Be aware of uncertainty introduced by the immediate working environment. You may need to take account of or protect your experiment from vibrations, drafts, changes in temperature, and electronic noise or other effects from nearby apparatus.

Instrument resolution (random) - All instruments have finite precision that limits the ability to resolve small measurement differences. For instance, a meter stick cannot be used to distinguish distances to a precision much better than about half of its smallest scale division (typically 0.5 mm). One of the best ways to obtain more precise measurements is to use a *null difference* method instead of measuring a quantity directly. *Null* or *balance* methods involve using instrumentation to measure the difference between two similar quantities, one of which is known very accurately and is adjustable. The adjustable reference quantity is varied until the difference is reduced to zero. The two quantities are then balanced, and the magnitude of the unknown quantity can be found by comparison with a measurement standard. With this method, problems of source instability are eliminated, and the measuring instrument can be very sensitive and does not even need a scale. This type of measurement is more sophisticated and will typically not be used in the introductory physics courses.

Calibration (systematic) – Whenever possible, the calibration of an instrument should be checked before taking data. If a calibration standard is not available, the accuracy of the instrument should be checked by comparing with another instrument that is at least as precise, or by consulting the technical data provided by the manufacturer. Calibration errors are usually linear (measured as a fraction of the full scale reading), so that larger values result in greater absolute errors.

Zero offset (systematic) - When making a measurement with a micrometer caliper, electronic balance, or electrical meter, always check the zero reading first. Re-zero the instrument if possible, or at least measure and record the zero offset so that readings can be corrected later. It is also a good idea to check the zero reading throughout the experiment. Failure to zero a device will result in a constant offset that is more significant for smaller measured values than for larger ones.

Physical variations (random) - It is always wise to obtain multiple measurements over the widest range possible. Doing so often reveals variations that might otherwise go undetected. These variations may call for closer examination, or they may be combined to find an average value.

Parallax (systematic or random) - This error can occur whenever there is some distance between the measuring scale and the indicator used to obtain a measurement. If the observer's eye is not squarely aligned with the pointer and scale, the reading may be too high or low (some analog meters have mirrors to help with this alignment).

Instrument drift (systematic) - Most electronic instruments have readings that drift over time. The amount of drift is generally not a concern, but occasionally this source of uncertainty can be significant.

Lag time and **hysteresis** (systematic) - Some measuring devices require time to reach equilibrium, and taking a measurement before the instrument is stable will result in a measurement that is too high or low. A common example is taking temperature readings with a thermometer that has not reached thermal equilibrium with its environment. A similar effect is *hysteresis*, wherein the instrument readings lag behind and appear to have a "memory" effect, as data are taken sequentially moving up or down through a range of values. Hysteresis is most commonly associated with materials that become magnetized when a changing magnetic field is applied.

Last but not least, some uncertainties are the result of carelessness, poor technique, or bias on the part of the experimenter. The experimenter may use a measuring device incorrectly, or may use poor technique in taking a measurement, or may introduce a bias into measurements by expecting (and inadvertently forcing) the results to agree with the expected outcome. Gross uncertainties of this nature can be referred to as **mistakes** or **blunders**, and should be avoided and corrected if discovered. As a rule, these uncertainties are *excluded* from any uncertainty analysis discussion because it is generally assumed that the experimental result was obtained by following correct and well-intentioned procedures – there is no point to performing an experiment and then reporting that it was known to be done incorrectly. *The term **human error** should be avoided in uncertainty analysis discussions because it is too general to be useful.*

Estimating Experimental Uncertainty for a Single Measurement

Any measurement will have some uncertainty associated with it, no matter the precision of the measuring tool. How is this uncertainty determined and reported? The uncertainty of a single measurement is limited by the precision and accuracy of the measuring instrument, along with any other factors that might affect the ability of the experimenter to make the measurement.

For example, if you are trying to use a meter stick to measure the diameter of a tennis ball, the uncertainty might be ±5 mm, but if you use a Vernier caliper, the uncertainty could be reduced to maybe ±2 mm. The limiting factor with the meter stick is parallax, while the second case is limited by ambiguity in the definition of the tennis ball's diameter (it's fuzzy!). In both of these cases, the uncertainty is greater than the smallest divisions marked on the measuring tool (likely 1 mm and 0.05 mm respectively). Unfortunately, there is no general rule for determining the uncertainty in all measurements. The experimenter is the one who can best evaluate and quantify the uncertainty of a measurement based on all the possible factors that affect the result. Therefore, the person making the measurement has the obligation to make the best judgment possible and report the uncertainty in a way that clearly explains what the uncertainty represents:

Measurement = (measured value ± standard uncertainty) (unit of measurement)

where "± **standard uncertainty"** indicates approximately a 68% confidence interval (see sections on Standard Deviation and Reporting Uncertainties).

Example: Diameter of tennis ball = 6.7 ± 0.2 cm

Estimating Uncertainty in Repeated Measurements

Suppose you time the period of oscillation of a pendulum using a digital instrument (that you assume is measuring accurately) and find that T = 0.44 seconds. This single measurement of the period *suggests* a precision of ±0.005 s, but this instrument precision may not give a complete sense of the uncertainty, and you should avoid reporting the uncertainty in this fashion if possible. If you repeat the measurement several times and examine the variation among the measured values, you can get a better idea of the uncertainty in the period. For example, here are the results of 5 measurements, in seconds: 0.46, 0.44, 0.45, 0.44, 0.41. For this situation, the best estimate of the period is the **average**, or **mean**:

$$\text{Average (mean)} = \frac{x_1 + x_2 + \ldots + x_N}{N}$$

Whenever possible, repeat a measurement several times and average the results. This average is generally the best estimate of the "true" value (unless the data set is skewed by one or more *outliers* which should be examined to determine if they are bad data points that should be omitted from the average or valid measurements that require further investigation). Generally, the more repetitions you make of a measurement, the better this estimate will be, but be careful to avoid wasting time taking more measurements than is necessary for the precision required.

Consider, as another example, the measurement of the width of a piece of paper using a meter stick. Being careful to keep the meter stick parallel to the edge of the paper (to avoid a systematic error which would cause the measured value to be consistently higher than the correct value), the width of the paper is measured at a number of points on the sheet, and the values obtained are entered in a data table. Note that the last digit is only a rough estimate, since it is difficult to read a meter stick to the nearest tenth of a millimeter (0.01 cm) – we retain the last digit for now to make a point later.

Observation	Width (cm)
#1	31.33
#2	31.15
#3	31.26
#4	31.02
#5	31.20

Table 1. Five Measurements of the Width of a Piece of Paper

$$\text{Average} = \frac{\text{sum of observed widths}}{\text{number of observations}} = \frac{155.96\text{ cm}}{5} = 31.19\text{ cm}$$

This average is the best available estimate of the width of the piece of paper, but it is not exact. We would have to average an infinite number of measurements to approach the true mean value, and even then, we are not guaranteed that the mean value is *accurate* because there is still likely some systematic uncertainty from the measuring tool, which is difficult to calibrate perfectly unless it is the gold standard. So how do we express the uncertainty in our average value?

One way to express the variation among the measurements is to use the **average deviation.** This statistic tells us on average (with 50% confidence) how much the individual measurements vary from the mean.

$$\text{Average Deviation, } \bar{d} = \frac{|x_1 - \bar{x}| + |x_2 - \bar{x}| + \ldots + |x_N - \bar{x}|}{N}$$

The average deviation would seem to be a sufficient measure of uncertainty; however, it is important to understand the *distribution* of measurements. The

Central Limit Theorem proves that as the number of independent measurements increases, and assuming that the variations in these measurements are random (i.e., there are no systematic uncertainties), the distribution of measurements will approach the *normal* distribution, more commonly known as a *bell curve*. In this course, we will assume that our measurements, performed in sufficient number, will produce a bell curve (normal) distribution. In this case, the *standard deviation* is the correct way to characterize the spread of the data. The standard deviation is always slightly greater than the average deviation, and is used because of its mathematical association with the normal distribution.

Standard Deviation

To calculate the standard deviation for a sample of *N* measurements:

1. Sum all the measurements and divide by *N* to get the **average**, or **mean**.
2. Subtract this **average** from each of the *N* measurements to obtain *N* "**deviations.**"
3. **Square** each of the *N* **deviations** and add them together.
4. Divide this result by (*N*–1) and take the square root.

To convert this into a formula, let the *N* measurements be called $x_1, x_2, \ldots, x_N$. Let the average of the *N* values be called $\bar{x}$. Then each deviation is given by

$$\delta x_i = x_i - \bar{x}, \text{ for } i = 1, 2, ..., N$$

The standard deviation is then:

$$s = \sqrt{\frac{\left(\delta x_1^2 + \delta x_2^2 + \ldots + \delta x_N^2\right)}{(N-1)}} = \sqrt{\frac{\sum \delta x_i^2}{(N-1)}}$$

In the meter stick and paper example, the average paper width $\bar{x}$ is 31.19 cm. The deviations are:

Observation	Width (cm)	Deviation (cm)	
#1	31.33	+0.14	= 31.33 - 31.19
#2	31.15	-0.04	= 31.15 - 31.19
#3	31.26	+0.07	= 31.26 - 31.19
#4	31.02	-0.17	= 31.02 - 31.19
#5	31.20	+0.01	= 31.20 - 31.19

Table 1 (completed). Five Measurements of the Width of a Sheet of Paper

The *average* deviation is: $\bar{d}$ = 0.09 cm

The *standard* deviation is: $s = \sqrt{\frac{(0.14)^2 + (0.04)^2 + (0.07)^2 + (0.17)^2 + (0.01)^2}{5-1}} = 0.12 \text{ cm}$

The significance of the standard deviation is this: if you now make one more measurement using the same meter stick, you can reasonably expect (with about 68% confidence) that the new measurement will be within 0.12 cm of the estimated average of 31.19 cm. In fact, it is reasonable to use the standard deviation as the uncertainty associated with this *single* new measurement. However, the uncertainty

of the *average* value is the *standard deviation of the mean*, which is always *less* than the standard deviation (see next section).

Consider an example of 100 measurements of a quantity, for which the average or mean value is 10.50 and the standard deviation is s = 1.83. Figure 2 below is a *histogram* of the 100 measurements, which shows how often a certain range of values was measured. For example, in 20 of the measurements, the value was in the range 9.50 to 10.50, and most of the readings were *close* to the mean value of 10.50. The standard deviation s for this set of measurements is roughly how far from the average value *most* of the readings fell. For a large enough sample, approximately 68% of the readings will be within one standard deviation ("1-sigma") of the mean value, 95% of the readings will be in the interval $\bar{x} \pm 2s$ ("2-sigma") , and nearly all (99.7%) of the readings will lie within 3 standard deviations ("3-sigma") of the mean. The smooth curve superimposed on the histogram is the *normal* distribution predicted by theory for measurements involving random errors. As more and more measurements are made, the histogram will better approximate a bell-shaped curve, but the standard deviation of the distribution will remain approximately the same.

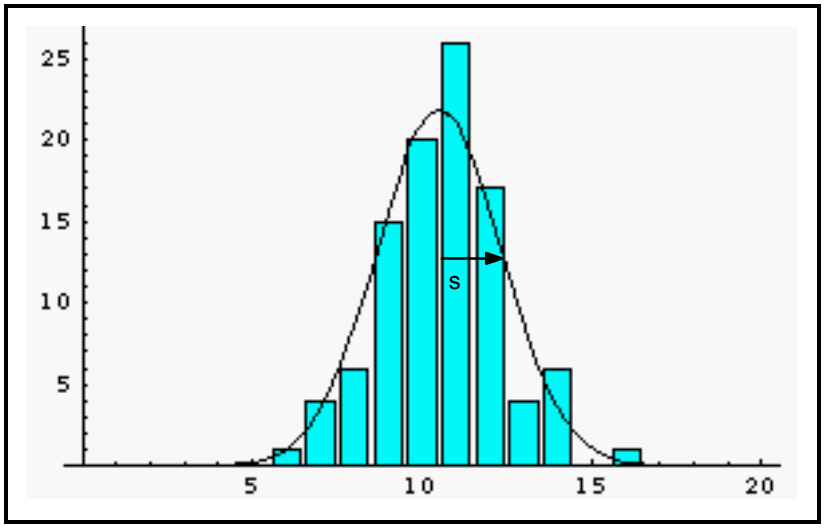


Figure 2. A Normal Distribution (Bell Curve) Based on 100 Measurements

Standard Deviation of the Mean (Standard Error)

When reporting the average value of *N* measurements, the uncertainty associated with this average value is the *standard deviation of the mean*, often called the *standard error* (SE).

$$\text{Standard Deviation of the Mean, or Standard Error (SE)}, \quad \sigma_{\bar{x}} = \frac{s}{\sqrt{N}} \qquad (3)$$

The standard error is smaller than the standard deviation by a factor of $1/\sqrt{N}$. This reflects the fact that we expect the uncertainty of the average value to get smaller when we use a larger number of measurements. In the previous example, we have divided the standard deviation of 0.12 by √5 to get the standard error of 0.05 cm. The final result should then be reported as "average paper width = 31.19 ± 0.05 cm."

When to Use Standard Deviation vs Standard Error

For repeated measurements, the significance of the standard deviation *s* is that you can reasonably expect (with about 68% confidence) that the next measurement will be within *s* of the estimated average. It is reasonable to use the standard deviation as the uncertainty associated with this measurement; however, as more measurements are made, the value of the standard deviation may be refined but it will not significantly decrease as the number of measurements is increased.

In contrast, if you are confident that the systematic uncertainty in your measurement is very small, then it is reasonable to assume that your finite sample of all possible measurements is not biased away from the "true" value. In this case, the uncertainty of the average value can be expressed as the standard deviation of the mean, which is always less than the standard deviation by a factor of √N.

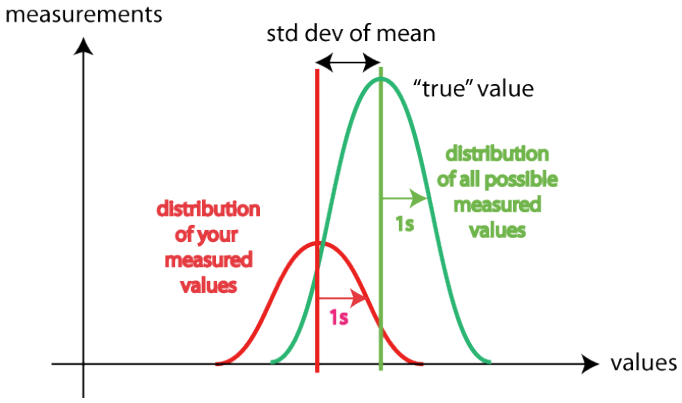


Figure 3. Standard Deviation vs Standard Error

If you are *not* confident that the systematic uncertainty in your measurement is very small, the uncertainty that should be reported is the standard combined uncertainty (*Uc*) that includes all known uncertainty estimates (see section on *Combining and Reporting Uncertainties* later in this document).

Anomalous Data

The first step you should take in analyzing data (and even while taking data) is to examine the data set as a whole to look for patterns and **outliers**. Anomalous data points that lie *outside* the general trend of the data may suggest an interesting phenomenon that could lead to a new discovery, or they may simply be the result of a mistake or random fluctuations. In any case, an outlier requires closer examination to determine the cause of the unexpected result. Extreme data should never be "thrown out" without clear justification and explanation, because you may be discarding the most significant part of the investigation! However, if you can clearly justify omitting an inconsistent datum, then you may exclude the outlier from your analysis so that the average value is not *skewed* from the "true" mean. There are a number of statistical measures that help quantify the decision to discard outliers, but they are beyond the scope of this document. Be aware of the possibility of anomalous data, and address the topic as needed in the discussion included with a lab report or lab notebook.

Fractional Uncertainty

When a reported value is determined by taking the average of a set of independent readings, the fractional uncertainty is given by the ratio of the uncertainty divided by the average value. For this example,

$$\text{Fractional Uncertainty} = \frac{\text{uncertainty}}{\text{average}} = \frac{0.05\ \text{cm}}{31.19\ \text{cm}} = 0.0016 \approx 0.2\%$$

The fractional uncertainty is dimensionless but is often reported as a percentage or in parts per million (ppm) to emphasize the fractional nature of the value. A scientist might also make the statement that this measurement "is good to about 1 part in 500" or "precise to about 0.2%". The fractional uncertainty is important because it is used in *propagating* uncertainty in calculations using the result of a measurement, as discussed in the next section.

Biases and the Factor of N–1

You may find it surprising that the best value (average) is calculated by normalizing (dividing) by N, whereas the standard deviation is calculated by normalizing to N–1. The reason is because normalizing to N is known to *underestimate* the correct value of the width of a normal distribution, unless N is large. This underestimate is referred to as a *bias* and is the result of incomplete sampling (that is, the population of measurements falls short of the entire population of measurements that could be taken). If the number of samples is less than 10 or so, even the N–1 term (known as Bessel's correction) can still induce a bias. Determining the exact correction to minimize or eliminate bias depends on the distribution of the data, and there is no simple exact equation that can be applied; however, for small sample sizes that are quite common in introductory physics classes, a correction of N–1.5 may be more appropriate. If you use a correction factor of 1.5 in your lab reports, you must make this clear in your analysis and cite this Guide as a reference.

Significant Figures

The number of significant figures in a value can be defined as all the digits between and including the first non-zero digit from the left, through the last digit. For instance, 0.44 has two significant figures, and the number 66.770 has 5 significant figures. Zeroes are significant except when used to locate the decimal point, as in the number 0.00030, which has 2 significant figures. Zeroes may or may not be significant for numbers like 1200, where it is not clear whether two, three, or four significant figures are indicated. To avoid this ambiguity, such numbers should be expressed in scientific notation (e.g., 1.20×10^3 clearly indicates three significant figures).

When using a calculator, the display will often show many digits, only some of which are *meaningful* (significant in a different sense). For example, if you want to estimate the area of a circular playing field, you might pace off the radius to be 9 meters and use the formula $A = \pi r^2$. When you compute this area, the calculator will report a value of 254.4690049 m^2. It would be extremely misleading to report this number as the area of the field, because it would suggest that you know the area to an absurd degree of precision – to within a fraction of a square millimeter! Since the radius is only known to one significant figure, it is considered best practice to also express the final answer to only one significant figure: Area = 3×10^2 m^2.

From this example, we can see that the number of significant figures reported for a value *implies* a certain degree of precision and can *suggest* a rough estimate of the relative uncertainty:

1 significant figure suggests a relative uncertainty of about 10% to 100%
2 significant figures suggest a relative uncertainty of about 1% to 10%
3 significant figures suggest a relative uncertainty of about 0.1% to 1%

To understand this connection more clearly, consider a value with 2 significant figures, like 99, which suggests an uncertainty of ±1, or a relative uncertainty of ±1/99 = ±1% (some might argue that the implied uncertainty in 99 is ± 0.5 since the range of values that would round to 99 are 98.5 to 99.4; however, since the uncertainty here is only a rough estimate, there is not much point arguing about the factor of two.) The smallest 2-significant-figure number, 10, also suggests an uncertainty of ±1, which in this case is a relative uncertainty of ±1/10 = ±10%. The ranges for other numbers of significant figures can be reasoned in a similar manner.

Warning: this procedure is open to a wide range of interpretation; therefore, one should use caution when using significant figures to imply uncertainty, and the method should only be used if there is no other better way to determine uncertainty. An explicit warning to this effect should accompany the use of this method.

Subject to the above warning, significant figures can be used to find a possibly appropriate precision for a calculated result for the four most basic math functions.

- For multiplication and division, the number of significant figures that are reliably known in a product or quotient is the same as the smallest number of significant figures in any of the original numbers.

Example:	6.6	(2 significant figures)
	× 7328.7	(5 significant figures)
	$48369.42 = 48 \times 10^3$	(2 significant figures)

- For addition and subtraction, the result should be rounded off to the last decimal place reported for the least precise number.

Examples:	223.64	5560.5
	+54	+0.008
	278	5560.5

Critical Note: if a calculated number is to be used in further calculations, it is *mandatory* to keep *guard digits* to reduce rounding errors that may accumulate. The final answer can then be rounded according to the above guidelines. The number of guard digits required to maintain the integrity of a calculation depends on the type of calculation. For example, the number of guard digits must be larger when performing power law calculations than when adding.

Uncertainty, Significant Figures, and Rounding

For the same reason that it is dishonest to report a result with more significant figures than are reliably known, the uncertainty value should also not be reported with excessive precision. For example, it would be unreasonable to report a result in the following way:

measured density = 8.93 ± 0.475328 g/cm^3 **WRONG!**

The uncertainty in the measurement cannot possibly be known so precisely! In most experimental work, the confidence in the uncertainty estimate is not much better than about ±50% because of all the various sources of error, none of which can be known exactly. Therefore, uncertainty values should be stated to only one significant figure (or perhaps 2 significant figures if the first digit is a 1). Because experimental uncertainties are inherently imprecise, they should be rounded to one, or at most two, significant figures.

To help give a sense of the amount of confidence that can be placed in the standard deviation as a measure of uncertainty, the following table indicates the relative uncertainty associated with the standard deviation for various sample sizes. Note that in order for an uncertainty value to be reported to 3 significant figures, more than 10,000 readings would be required to justify this degree of precision!

N	Relative Uncertainty*	Significant Figures Valid	Implied Uncertainty
2	71%	1	±10% to 100%
3	50%	1	± 10% to 100%
4	41%	1	± 10% to 100%
5	35%	1	± 10% to 100%
10	24%	1	± 10% to 100%
20	16%	1	± 10% to 100%
30	13%	1	± 10% to 100%
50	10%	2	± 1% to 10%
100	7%	2	± 1% to 10%
10000	0.7%	3	±0.1% to 1%

Table 2. Valid Significant Figures in Uncertainties

*The relative uncertainty is given by the approximate formula: $\frac{\sigma_\sigma}{\sigma} = \frac{1}{\sqrt{2(N-1)}}$

When an explicit uncertainty estimate is made, the uncertainty term indicates how many significant figures should be reported in the measured value (not the other way around!). For example, the uncertainty in the density measurement above is about 0.5 g/cm^3, which suggests that the digit in the tenths place is uncertain, and should be the last one reported. The other digits in the hundredths place and beyond are insignificant, and should not be reported:

measured density = 8.9 ± 0.5 g/cm^3 **RIGHT!**

An experimental value should be rounded to be consistent with the magnitude of its uncertainty. This generally means that the last significant figure in any reported value should be in the same decimal place as the uncertainty.

In most instances, this practice of rounding an experimental result to be consistent with the uncertainty estimate gives the same number of significant figures as the rules discussed earlier for simple propagation of uncertainties for adding, subtracting, multiplying, and dividing.

Caution: When conducting an experiment, it is important to keep in mind that *precision is expensive* (both in terms of time and material resources). Do not waste

your time trying to obtain a precise result when only a rough estimate is required. The cost increases exponentially with the amount of precision required, so the potential benefit of this precision must be weighed against the extra cost.

Practical Tips for Measuring Uncertainty

I. "Stacking"

Assume you are asked to measure the mass of a *typical* penny (according to the US Mint, currently made pennies have a nominal mass of 2.5 grams) with a scale whose accuracy is known to be ±0.2 gram. Measuring one penny might yield a measurement of 2.4 ± 0.2 grams, and this would be the only measurement possible for that one penny. Likewise, another penny might yield a measurement of 2.5 ± 0.2 grams.

Is there a way to get a more precise measurement? In this case, yes, because you are asked to find the mass of a *typical* penny. By *stacking* pennies and measuring more than one of them at the same time, dividing by the number of pennies measured can provide a more precise answer. For example, assume that you measure 5 pennies separately with these results (all with an accuracy of ±0.2 g): 2.4, 2.4, 2.5, 2.4, 2.6. The relative uncertainty of each measurement is about 8%. Further assume that when you measure all five at the same time, the value is 12.3 ± 0.2 g, yielding a relative uncertainty of about 2% for the stack. The mean value for a typical penny is therefore (retaining guard digits) 2.460 g. But what do we assign as the uncertainty? One might argue that the uncertainty is still 0.2; however, the uncertainty can also be divided by 5, based on the upper-lower bound method, for which the sum of the individual measurements can be plausibly written as:

$$\text{sum} = (2.4 \pm 0.04) + (2.4 \pm 0.04) + (2.5 \pm 0.04) + (2.4 \pm 0.04) + (2.6 \pm 0.04) = 12.3 \pm 0.2 \text{ g}$$

This is arguably the same as the stacked value of 12.3 ± 0.2 g. Therefore, we can reasonably divide both the stacked value *and its uncertainty* by *N*, and can thus reasonably assert the value for a typical penny as 2.46 ± 0.04 g.

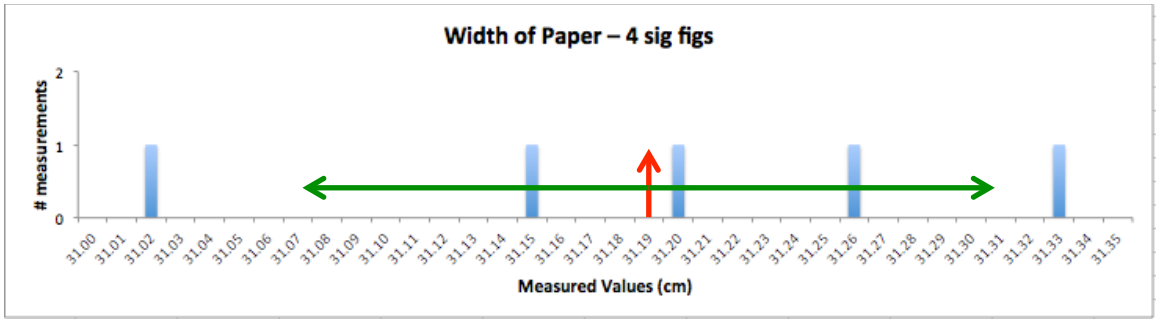
II. Always Minimize Your Sig Figs

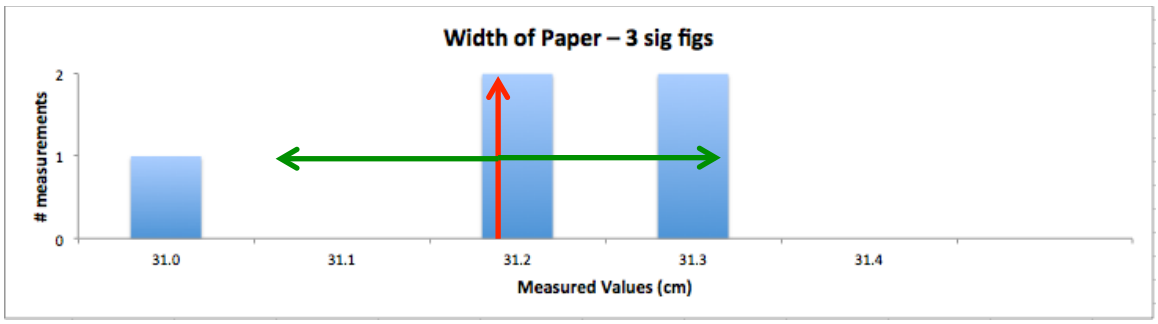
This "stacking" method can be used for any type of measurement that requires typical values to be found from repeated measurements of similar objects or time intervals; however, the method is not foolproof. Suppose instead that the scale has a quoted accuracy that is much better than its resolution (e.g., accuracy = 0.2%, and resolution = 0.1 g). Such devices are not designed to measure small values (the 0.2% accuracy for values on the order of 1 g is smaller than the resolution). In this case, the sum, in full precision, would be:

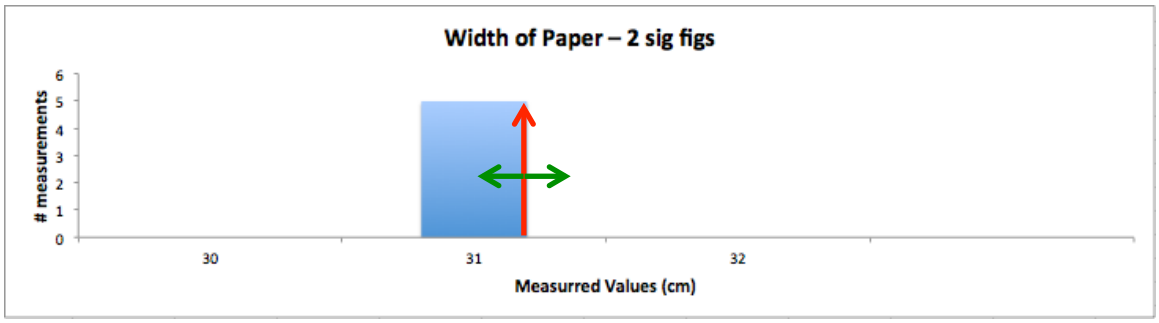
sum = (2.4 ± 0.005) + (2.4 ± 0.005) + (2.5 ± 0.005) + (2.4 ± 0.005) + (2.6 ± 0.005) = 12.300 ± 0.025 g

Note that the 0.025 uncertainty is 4 times less than the resolution (in contrast to the previous example, where the 0.2 uncertainty is twice the resolution). It would therefore be incorrect to assert the answer as 2.460 ± 0.005 g, because the accuracy cited far exceeds the resolution of the instrument.

This result is easier to visualize by looking at actual distributions. Suppose that you have a device that reports measurements to 4 sig figs. Consider the previous example of the meter stick used to measure the width of a piece of paper, where 5 measurements were used to determine that the width is 31.19 ± 0.12 cm (Table 1). Table 2 asserts that only 1 sig fig in the standard deviation is justified for 5 data points. Thus, this result should properly be reported as 31.2 ± 0.1 cm, consistent with our belief that it is difficult to read a meter stick to the nearest tenth of a millimeter. Histograms of the measurements to 4, 3, and 2 sig figs are shown below. The red and green arrows represent the average and standard deviation, respectively.







Question: which choice of sig figs provides the best representation of a bell curve whose height is represented by the red arrow and its width is represented by the green arrow? The most likely answer of course is *none*: none of the distributions looks much like a bell curve! The reason for this is that not enough data points exist to create a bell curve. Without knowing that the actual distribution of measurements would look like a bell curve, we cannot be sure that these data do create a normal distribution, for which the concept of a standard deviation makes sense. Given this limitation, the correct solution is to take many more data points; however, for 4 sig figs, many many data points (likely thousands or tens of thousands) would be required to fill *every* bin in such a way that a bell curve could be approximated. Likewise, with only 2 sig figs, it's probable that every datum will reduce to "31" with *no uncertainty*. The compromise in this case is 3 sig figs: we represent the uncertainty with a minimum number of sig figs. Without more time to take data and use more powerful statistical techniques, we will instead choose the smallest number of sig figs to cite the uncertainty. In return, we will assume that the average and standard deviation are the most reasonable approach to representing a distribution of data.

Any assertion of an uncertainty beyond 1 sig fig is only justified for N=50 or above (Table 1). Therefore, for the second stacked penny result, the sum would be appropriately rounded to 12.30 ± 0.03 g, yielding a final value for a typical penny of 2.460 ± 0.006 g. The only workaround to a better result in this case is to stack *many* more pennies, until the stacked uncertainty is large enough to compensate for the poor resolution.

Propagation of Uncertainty

Suppose we want to determine a quantity *f*, which depends on *x* and maybe several other variables *y, z,* etc. We want to know the uncertainty in *f* if we measure *x, y*, etc, with uncertainties σ_X, σ_Y, etc. That is, we want to find out how the uncertainty in one set of variables (usually the independent variables) propagates to the uncertainty in another set of variables (usually the dependent variables). There are two primary methods of performing this propagation procedure:

- ***upper-lower bound***
- ***quadrature***

The upper-lower bound method is simpler in concept, but tends to overestimate the uncertainty, while the quadrature method is more sophisticated (and complicated) but provides a better statistical estimate of the uncertainty.

The Upper-Lower Bound Method of Uncertainty Propagation

This method uses the uncertainty ranges of each variable to calculate the maximum and minimum values of the function. You can also think of this procedure as examining the best and worst case scenarios. For example, suppose you measure an angle to be $\theta = 25° \pm 1°$ and you need to find $f = \cos\theta$, then:

$$f_{\max} = \cos(26°) = 0.8988 \qquad f_{\min} = \cos(24°) = 0.9135$$

Then, $f = 0.906 \pm \mathbf{0.007}$ (where 0.007 is half the difference between $f_{\max}$ and $f_{\min}$)

Note that even though θ was only measured to 2 significant figures, *f* is known to 3 figures.

As another important example, consider the division of two variables. A common example is the calculation of average speed:

$$v_{avg} = \frac{\Delta x}{\Delta t}$$

Let's say an experiment done repeatedly measures distance travelled of 30 ± 0.5 m during a time of 2 ± 0.1 sec. To find the upper and lower bound of v_{avg}, the uncertainties must be set to create the "worst case scenario" for the uncertainty in v_{avg}:

$$v_{avg-\max} = \frac{30 + 0.5 \text{ m}}{2 - 0.1 \text{ sec}} = 16.05 \text{ m/s} \qquad v_{avg-\min} = \frac{30 - 0.5 \text{ m}}{2 + 0.1 \text{ sec}} = 14.05 \text{ m/s}$$

The best (expected) value for the average speed is 30/2 = 15.00 m/s. The upper bound is 1.05 m/s higher but the lower bound is 0.95 m/s lower (different from 1.05). This is a typical outcome when using the upper-lower bound method. The uncertainty should be expressed as the *most conservative* value. Thus:

v_{avg} = 15.00 ± 1.05 m/s **PREFERRED!**

Note that it is not correct to take the difference between the upper and lower bound and divide by two:

v_{avg} = 15.05 ± 1.00 m/s **NOT PREFERRED!**

Although the last result satisfies symmetry between the bounds, it explicitly calculates an *incorrect* value of the best-known expected value of the average speed.

Many times, the difference between the so-called "preferred" and "not preferred" approaches is not significant enough to be an issue. For example, if it is appropriate to round the uncertainty in the above values to one sig fig, the answer is 15 ± 1 m/s, regardless of the approach. Nevertheless, you should be aware of this pitfall.

The upper-lower bound method is especially useful when the functional relationship is not clear or is incomplete. One practical application is forecasting the expected range in an expense budget. In this case, some expenses may be fixed, while others may be uncertain, and the range of these uncertain terms could be used to predict the upper and lower bounds on the total expense.

Quadrature

The quadrature method yields a *standard uncertainty* estimate (with a 68% confidence interval) and is especially useful and effective in the case of several variables that weight the uncertainty non-uniformly. The method is derived with several examples shown below.

For a single-variable function $f(x)$, the deviation in f can be related to the deviation in x using calculus:

$$\delta f = \left(\frac{df}{dx}\right)\delta x$$

Taking the square and the average yields:

$$\overline{\delta f^2} = \left(\frac{df}{dx}\right)^2 \overline{\delta x^2}$$

Using the definition of σ yields:

$$\sigma_f = \left|\frac{df}{dx}\right|\sigma_x$$

Examples for power laws:

$f = \sqrt{x}$

$$\frac{df}{dx} = \frac{1}{2\sqrt{x}}$$

$$\sigma_f = \frac{\sigma_x}{2\sqrt{x}} \quad \text{or} \quad \frac{\sigma_f}{f} = \frac{1}{2}\frac{\sigma_x}{x}$$

$f = x^2$

$$\frac{df}{dx} = 2x$$

$$\frac{\sigma_f}{f} = 2\frac{\sigma_x}{x}$$

Note that by judiciously normalizing, it is easy to express the *relative* (fractional) uncertainty in one variable with respect to the *relative* (fractional) uncertainty in another. Note also that the weighting is directly related to the power exponent of the function. Now reconsider the trig example from the upper-lower bound section:

$$f = \cos\theta$$

$$\frac{df}{d\theta} = -\sin\theta$$

$$\sigma_f = |\sin\theta|\sigma_\theta$$

Note that in this situation, σ_θ must be in radians. For $\theta = 25° \pm 1°$ (0.727 ± 0.017)

$$\sigma_f = |\sin\theta|\sigma_\theta = (0.423)(\pi/180) = 0.0074$$

This is the same result as upper-lower bound method. The *fractional* uncertainty follows immediately as:

$$\frac{\sigma_f}{f} = |\tan\theta|\sigma_\theta$$

The deeper power of the quadrature method is evident in the case where *f* depends on two or more variables; the derivation above can be repeated with minor modification. For two variables, *f*(*x*, *y*):

$$\delta f = \left(\frac{\partial f}{\partial x}\right)\delta x + \left(\frac{\partial f}{\partial y}\right)\delta y$$

The partial derivative $\frac{\partial f}{\partial x}$ means differentiating *f* with respect to *x* holding the other variables fixed. Taking the square and the average yields the generalized law of propagation of uncertainty by quadrature:

$$(\delta f)^2 = \left(\frac{\partial f}{\partial x}\right)^2 (\delta x)^2 + \left(\frac{\partial f}{\partial y}\right)^2 (\delta y)^2 + 2\left(\frac{\partial f}{\partial x}\right)\left(\frac{\partial f}{\partial y}\right)\overline{\delta x \delta y} \qquad (4)$$

If the measurements of *x* and *y* are *uncorrelated*, then $\overline{\delta x \delta y} = 0$, and this reduces to its most common form:

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

Addition and Subtraction Example: $f = x \pm y$

$$\frac{\partial f}{\partial x} = 1, \quad \frac{\partial f}{\partial y} = \pm 1 \quad \rightarrow \quad \sigma_f = \sqrt{\sigma_x^2 + \sigma_y^2}$$

When adding (or subtracting) *independent* measurements, the absolute uncertainty of the sum (or difference) is the root sum of squares (RSS) of the individual absolute uncertainties. When adding *correlated* measurements, the uncertainty in the result is simply the sum of the absolute uncertainties, which is always a larger uncertainty estimate than adding in quadrature (RSS). Adding or subtracting a constant does not change the absolute uncertainty of the calculated value as long as the constant is an exact value.

Multiplication example: $f = xy$

$$\frac{\partial f}{\partial x} = y, \quad \frac{\partial f}{\partial y} = x \quad \rightarrow \quad \sigma_f = \sqrt{y^2\sigma_x^2 + x^2\sigma_y^2}$$

Dividing the above equation by $f = xy$ yields:

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

Division example: $f = x/y$

$$\frac{\partial f}{\partial x} = \frac{1}{y}, \quad \frac{\partial f}{\partial y} = -\frac{x}{y^2} \quad \rightarrow \quad \sigma_f = \sqrt{\left(\frac{1}{y}\right)^2 \sigma_x^2 + \left(\frac{x}{y^2}\right)^2 \sigma_y^2}$$

Dividing the previous equation by $f = x/y$ yields:

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

When multiplying (or dividing) independent measurements, the relative uncertainty of the product (quotient) is the RSS of the individual relative uncertainties. When multiplying *correlated* measurements, the uncertainty in the result is just the sum of the relative uncertainties, which is always a larger uncertainty estimate than adding in quadrature (RSS). Multiplying or dividing by a constant does not change the relative uncertainty of the calculated value.

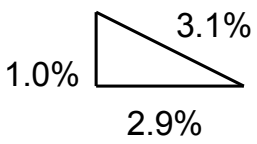
Note that the relative (fractional) uncertainty in f has the same form for multiplication and division: the relative uncertainty in a product or quotient depends on the *relative* uncertainty of each individual term.

As another example, consider propagating the uncertainty in the speed $v = at$, where the acceleration is $a = 9.8 \pm 0.1$ m/s^2 and the time is $t = 1.2 \pm 0.1$ s.

$$\frac{\sigma_v}{v} = \sqrt{\left(\frac{\sigma_a}{a}\right)^2 + \left(\frac{\sigma_t}{t}\right)^2} = \sqrt{\left(\frac{0.1}{9.8}\right)^2 + \left(\frac{0.1}{1.2}\right)^2} = \sqrt{(0.010)^2 + (0.029)^2} = 0.031 \text{ or } 3.1\%$$

Notice that the relative uncertainty in t (2.9%) is significantly greater than the relative uncertainty for a (1.0%), and therefore the relative uncertainty in v is essentially the same as for t (about 3%).

Graphically, the RSS is like the Pythagorean theorem:
The total uncertainty is the length of the hypotenuse of a right triangle with legs the length of each uncertainty component.



Timesaving approximation: "A chain is only as strong as its weakest link."
If one of the uncertainty terms is more than 3 times greater than the other terms, the root-squares formula can be skipped, and the combined uncertainty is simply the largest uncertainty. This shortcut can save a lot of time without losing any accuracy in the estimate of the overall uncertainty.

The quadrature method can be generalized to ***all*** power laws in the following way:

$$f = x^n y^m$$

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

The proof of this is shown in the Appendix.

The uncertainty estimate from the upper-lower bound method is generally larger than the standard uncertainty estimate found from the quadrature method, but both methods will give a reasonable estimate of the uncertainty in a calculated value.

Note: Once you have an understanding of the quadrature method, it is ***not required*** to perform the partial derivative every time you are presented with a propagation of uncertainty problem in any of the above forms! Instead, simply apply the correct formula for the relative uncertainties.

Combining and Reporting Uncertainties

In 1993, the International Standards Organization (ISO) published the first official worldwide *Guide to the Expression of Uncertainty in Measurement*. Before this time, uncertainty estimates were evaluated and reported according to different conventions depending on the context of the measurement or the scientific discipline. Here are a few key points from this 100-page guide, which can be found in modified form on the NIST website (see References).

When reporting a measurement, the measured value should be reported along with an estimate of the total **combined standard uncertainty** U_c of the value. The total uncertainty is found by combining the uncertainty components based on the two types of uncertainty analysis:

Type A evaluation of standard uncertainty – method of evaluation of uncertainty by the statistical analysis of a series of observations. This method primarily includes *random* uncertainties.

Type B evaluation of standard uncertainty – method of evaluation of uncertainty by means other than the statistical analysis of series of observations. This method includes *systematic* uncertainties and errors and any other factors that the experimenter believes are important.

The individual uncertainty components u_i should be combined using the *law of propagation of uncertainties,* commonly called the "root-sum-of-squares" or "RSS" method. When this is done, the combined standard uncertainty should be equivalent to the standard deviation of the result, making this uncertainty value correspond with a 68% confidence interval. If a wider confidence interval is desired, the uncertainty can be multiplied by a **coverage factor** (usually k = 2 or 3) to provide an uncertainty range that is believed to include the true value with a confidence of 95% (for k = 2) or 99.7% (for k = 3). If a coverage factor is used, there should be a clear explanation of its meaning so there is no confusion for readers interpreting the significance of the uncertainty value.

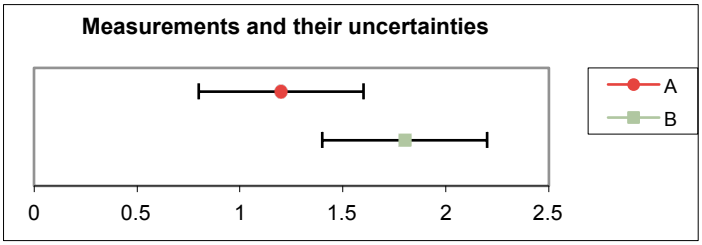
You should be aware that the ± uncertainty notation might be used to indicate different confidence intervals, depending on the scientific discipline or context. For example, a public opinion poll may report that the results have a **margin of error** of ±3%, which means that readers can be 95% confident (not 68% confident) that the reported results are accurate within 3 percentage points. Similarly, a manufacturer's **tolerance** rating generally assumes a 95% or 99% level of confidence.

Conclusion: "When do measurements agree with each other?"

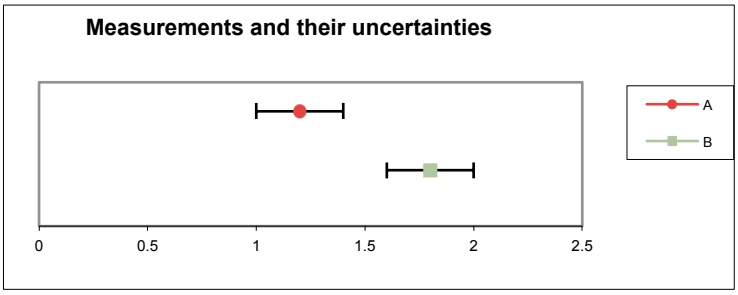
We now have the resources to answer the fundamental scientific question that was asked at the beginning of this error analysis discussion: "Does my result agree with a theoretical prediction or results from other experiments?"

Generally speaking, a measured result agrees with a theoretical prediction if the prediction lies within the range of experimental uncertainty. Similarly, if two measured values have *standard uncertainty* ranges that overlap, then the measurements are said to be *consistent* (they agree). If the uncertainty ranges do not overlap, then the measurements are said to be *discrepant* (they do not agree). However, you should recognize that these overlap criteria can give two opposite answers depending on the evaluation and confidence level of the uncertainty. It would be unethical to arbitrarily inflate the uncertainty range just to make a measurement agree with an expected value. A better procedure would be to discuss the size of the difference between the measured and expected values within the context of the uncertainty, and try to discover the source of the discrepancy if the difference is truly significant. Example:

A = 1.2 ± 0.4
B = 1.8 ± 0.4



These measurements **agree** within their uncertainties, despite the fact that the *percent difference* between their central values is 40%. In contrast, if the uncertainty is halved (±0.2), these same measurements **do not agree** since their uncertainties do not overlap:



Further investigation would be needed to determine the cause for the discrepancy. Perhaps the uncertainties were underestimated, there may have been a systematic error that was not considered, or there may be a true difference between these values.

An alternative method for determining agreement between values is to calculate the difference between the values divided by their combined standard uncertainty. This ratio gives the number of standard deviations separating the two values. If this ratio is less than 1.0, then it is reasonable to conclude that the values agree. If the ratio is more than 2.0, then it is highly unlikely (less than about 5% probability) that the values are the same.

Example from above with $u = 0.4$: $\frac{|1.2 - 1.8|}{0.57} = 1.1$ A and B likely ***agree***

Example from above with $u = 0.2$: $\frac{|1.2 - 1.8|}{0.28} = 2.1$ A and B likely ***do not agree***

Making Graphs

When graphs are required in laboratory exercises, you will be instructed to "plot A vs. B" (where A and B are variables). By convention, A *(the dependant variable)* should be plotted along the vertical axis (ordinate), and B *(the independent variable*) should be plotted along the horizontal axis (abscissa). Graphs that are intended to provide numerical information should be drawn on ruled graph paper. Use a sharp pencil (not a pen) to draw graphs, so that mistakes can be corrected easily. It is acceptable to use a computer (see the Excel tutorial below) to produce graphs. The following graph is a typical example in which distance vs. time is plotted for a freely falling object. Examine this graph and note the following important rules for graphing:

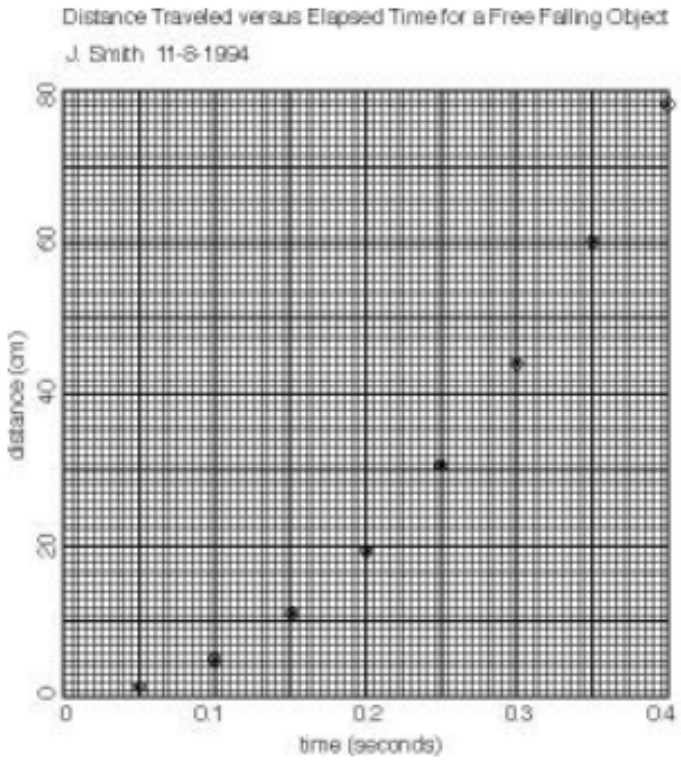


Figure 3. Plot of Distance vs Time

Title. Every graph should have a title that clearly states which variables appear on the plot. If the graph is not attached to another identifying report, write your name and the date on the plot for convenient reference.

Axis labels. Each coordinate axis of a graph should be labeled with the word or symbol for the variable plotted along that axis and the units (in parentheses) in which the variable is plotted.

Choice of Scale. Scales should be chosen in such a way that data are easy to plot and easy to read. On coordinate paper, every 5th and/or 10th line should be selected as

major division lines that represent a decimal multiple of 1, 2, or 5 (e.g., 0, l, 2, 0.05, 20, 500, etc.) – other choices (e.g., 0.3) make it difficult to plot and also read data. Scales should be made no finer than the smallest increment on the measuring instrument from which data were obtained. For example, data from a meter stick (which has l mm graduations) should be plotted on a scale no finer than l division = l mm, because a scale finer than 1 div/mm would provide no additional plotting accuracy, since the data from the meter stick are only accurate to about 0.5 mm. Frequently the scale must be considerably coarser than this limit, in order to fit the entire plot onto a single sheet of graph paper. In the illustration above, scales have been chosen to give the graph a roughly square boundary; avoid choices of scale that make the axes very different in length. Note that it is not always necessary to include the origin ('zero') on a graph axis; in many cases, only the portion of the scale that covers the data need be plotted.

Data Points. Enter data points on a graph by placing a suitable symbol (e.g., a small dot with a small circle around the dot) at the coordinates of the point. If more than one set of data is to be shown on a single graph, use other symbols (e.g., Δ) to distinguish the data sets. If drawing by hand, a drafting template is useful for this purpose.

Curves. Draw a simple smooth curve through the data points. The curve will not necessarily pass through all the points, but should pass as close as possible to each point, with about half the points on each side of the curve; this curve is intended to guide the eye along the data points and to indicate the trend of the data. A French curve is useful for drawing curved line segments. *Do not connect the data points by straight-line segments in a dot-to-dot fashion.* This curve now indicates the average trend of the data, and any *predicted* (interpolated or extrapolated) values should be read from this curve rather than reverting back to the original data points.

Straight-line Graphs. In many of the exercises in this course, you will be asked to *linearize* your experimental results (plot the data in such a way that there is a linear, or straight-line relationship between graphed quantities). In these situations, you will be asked to fit a straight line to the data points and to determine the slope and y-intercept from the graph. In the example given above, it is expected that the falling object's distance varies with time according to $d = \frac{1}{2}gt^2$. It is difficult to tell whether the data plotted in the first graph above agrees with this prediction; however, if d vs. t^2 is plotted, a straight line should be obtained with slope = $\frac{1}{2}g$ and y-intercept = 0.

Using Excel for Data Analysis in Physics Labs

Students have a number of software options for analyzing lab data and generating graphs with the help of a computer. It is the student's responsibility to ensure that the computational results are correct and consistent with the requirements stated in this lab manual. Any suitable software can be used to perform these analyses and generate tables and plots for lab reports and assignments; however, since Microsoft Excel is widely available on all CCI laptops and in university computer labs, students are encouraged to use this spreadsheet program. In addition, there may be assignments during the semester that specifically require an Excel (or platform-equivalent) spreadsheet to be submitted.

Getting Started

This tutorial will lead you through the steps to create a graph and perform linear regression analysis using an Excel spreadsheet. The techniques presented here can be used to analyze virtually any set of data you will encounter in your physics studio.

To begin, open Excel. A blank worksheet should appear. Enter the sample data and column headings shown below into cells A1 through D6. Save the file to a disk or to your personal file space on the campus network.

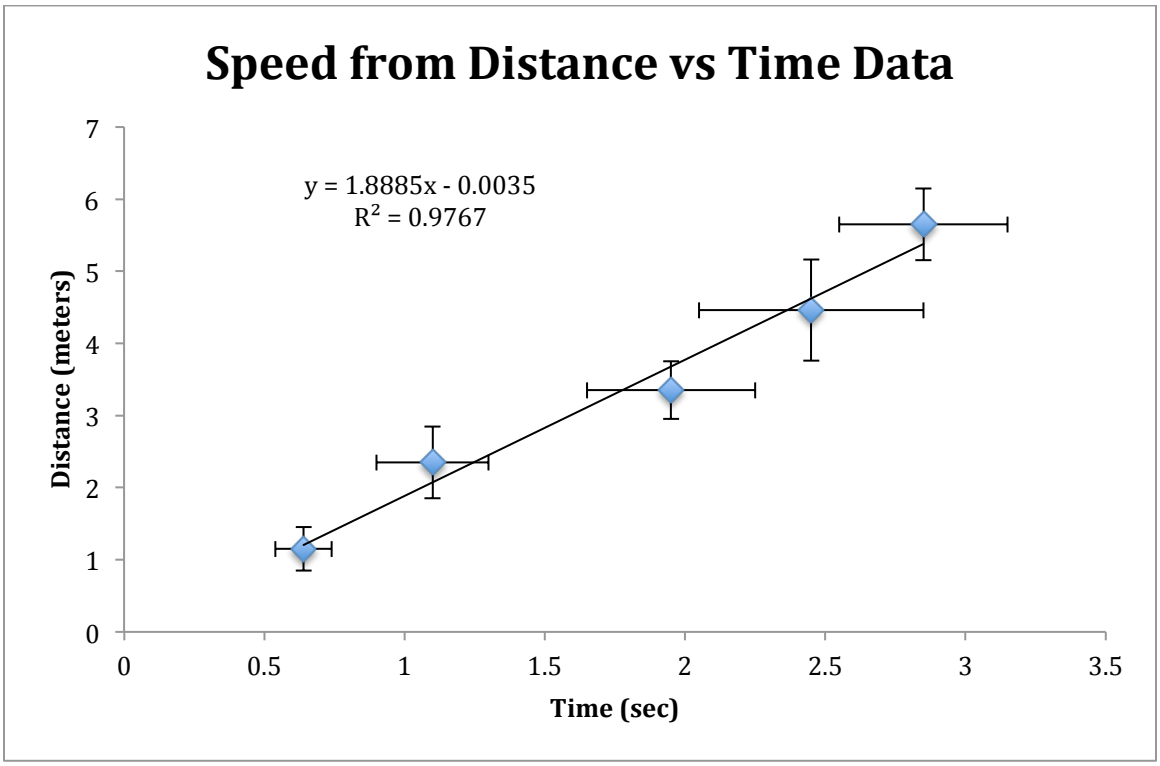
Time (sec)	Distance (m)	Time ±	Distance ±
0.64	1.15	0.1	0.3
1.1	2.35	0.2	0.5
1.95	3.35	0.3	0.4
2.45	4.46	0.4	0.7
2.85	5.65	0.3	0.5

Note that the uncertainties for the time and distance (denoted ±) have been included. These are not necessary for a basic plot, but the studio lab reports and assignments *require* an uncertainty analysis, so you should get into the habit of including them.

Creating and Editing a Graph

Note: these instructions may not be precisely correct for different versions of Excel or on different platforms; however, these instructions should be roughly correct for all modern versions of Excel on all platforms. Other spreadsheet programs should have similar features. We suggest you use on-line resources or consult with your classmates or your TA for specific questions or issues.

You will be creating a graph of these data whose finished form looks like this:



Follow these steps to accomplish this:

1: Use your mouse to select all the cells that contain the data that you want to graph (in this example, columns A and B). To graph these data, select **Chart** on the toolbar.

2: Click on **Scatter** plots and choose **XY (Scatter)** or **Marked Scatter** with no lines. A default plot should appear in the spreadsheet, and it should be both moveable and resizable.

3: Using the **Chart Layout** tool, experiment with setting the title, axes, axis titles, gridlines and legends. At a minimum, we *require* that the plot be titled and that the x- and y-axes are descriptively labeled with units. We strongly *suggest* that all gridlines and the legend be removed for clarity.

Most graph features can be modified by double-clicking on the feature you want to change. You can also right-click on a feature to get a menu. Try changing the color of the plot area, the numbers on the axes, and the appearance of the data points. It is recommended that you always format the background area to white using the "Automatic" option.

Adding Error Bars

Right-click on a datum point and choose **Format Data Series...** and select **Error Bars**. Click on the appropriate **Error Bar** *tab* (**X** or **Y**) and choose **Both** under **Display** and **Cap** for the **End style**. Fixed values or percentages can be set, for example, but if you have separate columns of uncertainty values for each datum, as shown above, then select **Custom** to specify the values. In the subsequent custom error bar window, select the positive error value field and then click and drag in the corresponding Excel column of uncertainties. Repeat for the negative value and click OK. Your custom error bars will then be applied. Repeat for the other axis. Note that if you create separate columns for the positive and negative error bars, they can be set independently. Also note that error bars may not be visible if they are smaller than the size of the datum point on the plot.

Adding a Trendline

The primary reason for graphing data is to examine the mathematical relationship between the two variables plotted on the x- and y-axes. To add a trendline and display its corresponding equation, right-click on any datum point and **Add Trendline**. Choose the graph shape that best fits your data and is consistent with your theoretical prediction (usually Linear). Click on the "Options" tab and check the boxes for "Display equation on chart" and "Display R-squared value on chart." A good fit is indicated by an R^2 value close to 1.

Caution: When searching for a mathematical model that explains your data, it is very easy to use the trendline tool to produce *nonsense*. This tool should be used to find the simplest mathematical model that explains the relationship between the two variables you are graphing. Look at the equation and shape of the trendline critically:

- Does it make sense in terms of the physical principle you are investigating?
- Is this the best possible explanation for the relationship between the two variables?

Use the simplest equation that passes through most of the error bars on your graph. You may need to try a couple of trendlines before you get the most appropriate one. To clear a trendline, right-click on its regression line and select **Clear**.

Determining the Uncertainty in Slope and Y-intercept

The R^2 value indicates the quality of the least-squares fit, but this value does not give the error in the slope directly. Given the best fit line $y = mx + b$, with n data points, the standard error (uncertainty) in the slope m can be determined from the R^2 value by using the following formula:

$$\sigma_m = m\sqrt{\frac{(1/R^2)-1}{n-2}}$$

Likewise, the uncertainty in the y-intercept b is:

$$\sigma_b = \sigma_m\sqrt{\frac{\sum x^2}{n}}$$

These values can be computed directly in Excel or by using a calculator. For this sample set of data, σ_m =0.1684 m/s, and σ_b = 0.333 m. Note that a value of R^2 of *exactly* 1 leads to slope and intercept uncertainties of *zero*. Carefully exam the Excel R^2 value – although it may *display* as exactly 1, it likely is not *exactly* 1. ***If your value is indeed exactly 1, it indicates an error in how you have plotted your data***.

The uncertainty in the slope and y-intercept can also be found by using the LINEST function in Excel. Using this function is somewhat tedious and is best understood from the Help feature in Excel.

Interpreting the Results

Once a regression line has been found, the equation must be interpreted in terms of the context of the situation being analyzed. This sample data set came from a cart moving along a track. We can see that the cart was moving at nearly a constant speed since the data points tend to lie in a straight line and do not curve up or down. The speed of the cart is simply the slope of the regression line, and its uncertainty is found from the equation above: $v = 1.9 \pm 0.2$ m/s. (Note: If we had plotted a graph of time versus distance, then the speed would be the inverse of the slope: $v = 1/m$) The y-intercept gives us the initial position of the cart: $x_0 = -0.0035 \pm 0.33$ m, which is essentially zero.

Final Step – Copying Data and Graphs into a Word Document

Copy your plots and data table from Excel to Word. Just select the graph (or cells) and use the Edit menu or keyboard shortcuts to copy and paste.

Using LINEST in Excel

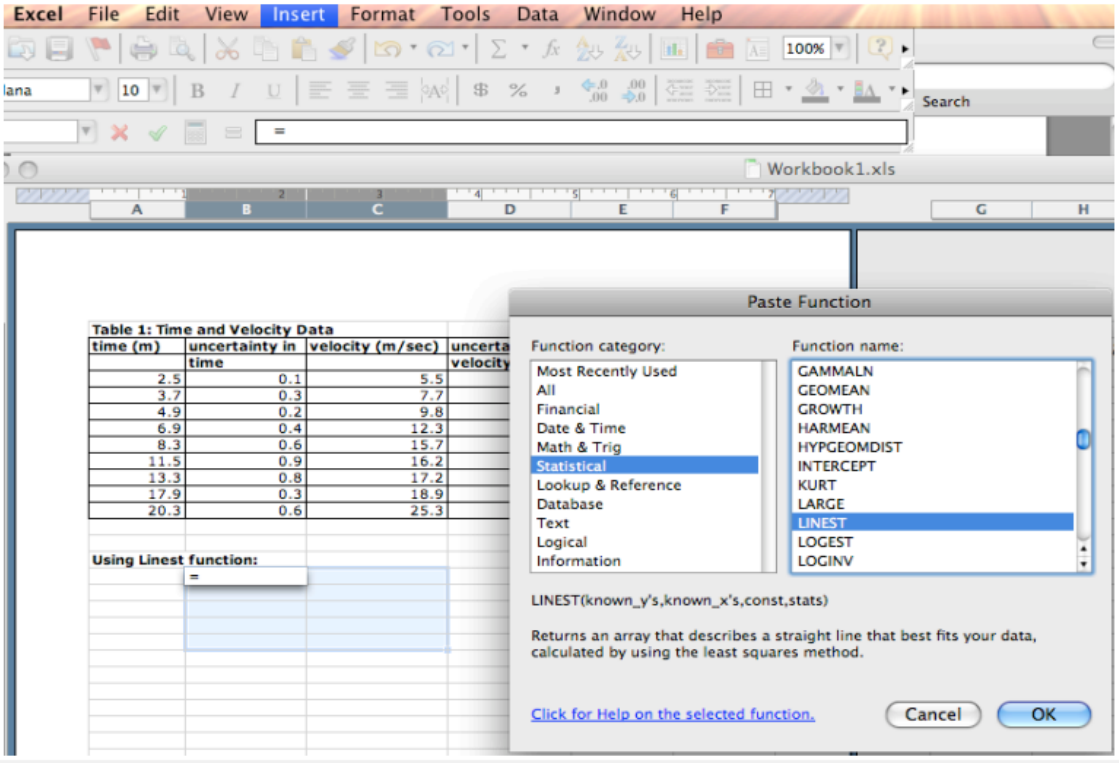
LINEST is an alternative linear least square fitting function in Excel. The results of the LINEST analysis are virtually identical to the linear trendline analysis described above; however, LINEST provides a single-step calculation of both the slope and intercept uncertainties, instead of the multi-step procedure described above.

- Start with a table for time and velocity (right).
- The LINEST function returns several outputs; to prepare, select a 2 by 5 array below the data, as shown.
- ~~Note, velocity was mistakenly labeled as having units of m/sec in the table to the right.~~
- Under the **Insert** menu, selection **Function**, then **Statistical**, and finally LINEST as shown below (and hit OK).

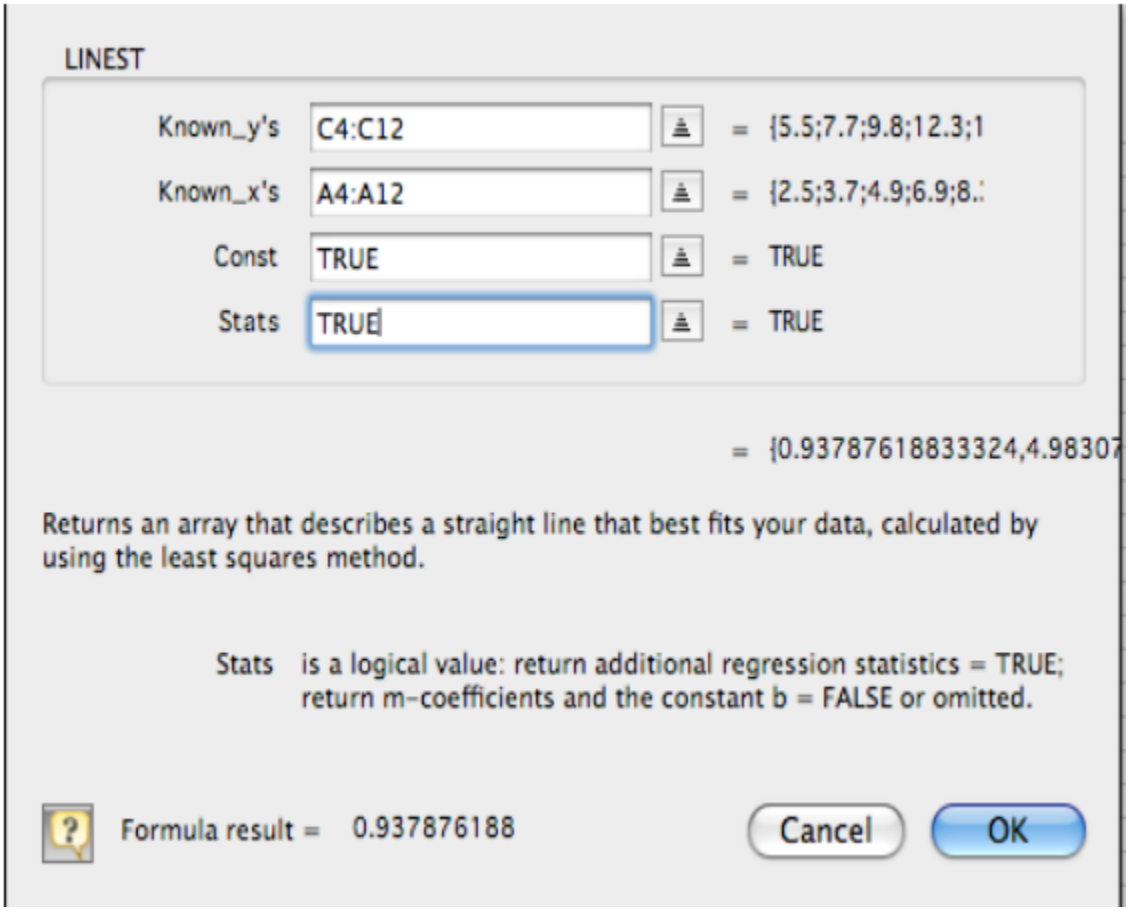
Table 1: Time and Velocity Data

time (m)	uncertainty in time	velocity (m/sec)	uncertainty in velocity
2.5	0.1	5.5	0.3
3.7	0.3	7.7	0.2
4.9	0.2	9.8	0.7
6.9	0.4	12.3	0.9
8.3	0.6	15.7	0.2
11.5	0.9	16.2	0.3
13.3	0.8	17.2	0.6
17.9	0.3	18.9	0.6
20.3	0.6	25.3	0.9

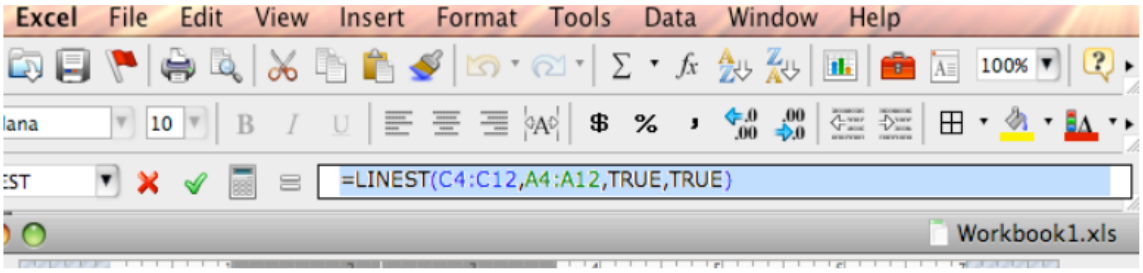
Using Linest function:



Select the y-values and x-values from the table and write 'TRUE' for the last two options (below). Again, press OK.



The LINEST function is an array function; therefore, you must tell Excel you are done with the array. Highlight the formula in the formula bar as shown below. Press Ctrl+Shift simultaneously with Enter (Mac users, press Command and Enter)



The result is that the array of blank cells that was selected previously is now filled with data (although the new data isn't labeled). See the image below for the labels that can be added after the fact (e.g., "Slope," "R^2 Value," etc):

Table 1: Time and Velocity Data			
time (m)	uncertainty in time	velocity (m/sec)	uncertainty in velocity
2.5	0.1	5.5	0.3
3.7	0.3	7.7	0.2
4.9	0.2	9.8	0.7
6.9	0.4	12.3	0.9
8.3	0.6	15.7	0.2
11.5	0.9	16.2	0.3
13.3	0.8	17.2	0.6
17.9	0.3	18.9	0.6
20.3	0.6	25.3	0.9
Using Linest function:			
Slope	0.937876188	4.983072931	Y-intercept
Slope Uncerainty	0.102847501	1.188614381	y-intercept uncertainty
R^2 Value	0.922358443	1.828360018	
	83.15790383	7	
	277.9885864	23.4003025	

References:

Baird, D.C. Experimentation: An Introduction to Measurement Theory and Experiment Design, 3rd. ed. Prentice Hall: Englewood Cliffs, 1995.

Bevington, Phillip and Robinson, D. Data Reduction and Error Analysis for the Physical Sciences, 2nd. ed. McGraw-Hill: New York, 1991.

ISO. *Guide to the Expression of Uncertainty in Measurement.* International Organization for Standardization (ISO) and the International Committee on Weights and Measures (CIPM): Switzerland, 1993.

Lichten, William. *Data and Error Analysis., 2nd. ed.* Prentice Hall: Upper Saddle River, NJ, 1999.

NIST. *Essentials of Expressing Measurement Uncertainty*. http://physics.nist.gov/cuu/Uncertainty/

Taylor, John. *An Introduction to Error Analysis, 2nd. ed.* University Science Books: Sausalito, 1997.

Appendix: Propagation of Uncertainty by Quadrature

Consider a quantity *f* to be calculated by multiplying two measured quantities *x* and *y* whose uncertainties are σ_x and σ_y, respectively. The question is how to propagate those uncertainties to the calculated quantity *f*. From the chain rule of calculus, the change in *f* due to changes in *x* and *y* is:

$$\delta f = \frac{\partial f}{\partial x}\delta x + \frac{\partial f}{\partial y}\delta y$$

Squaring and averaging yields:

$$\overline{(\delta f)^2} = \left(\frac{\partial f}{\partial x}\right)^2 \overline{(\delta x)^2} + \left(\frac{\partial f}{\partial y}\right)^2 \overline{(\delta y)^2} + 2\frac{\partial f}{\partial x}\frac{\partial f}{\partial y}\overline{\delta x \delta y}$$

For uncorrelated measurements, $\overline{\delta x \delta y}$ is zero. Consider the average square change in quantities to be the uncertainty in each of *x*, *y*, and *f*; that is, $\overline{(\delta f)^2} = \sigma_f^2$, etc. Then:

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

To generalize it to arbitrary powers of *x* and *y*, consider the function $f = x^n y^m$; substituting this into the last equation and dividing by *f* yields the relative uncertainty:

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

The partial derivatives are $\frac{\partial f}{\partial x} = n x^{n-1} y^m$ and $\frac{\partial f}{\partial y} = m x^n y^{m-1}$. Substituting these yields:

$$\frac{\sigma_f}{f} = \sqrt{\left(n x^{n-1} y^m\right)^2 \frac{\sigma_x^2}{\left(x^n y^m\right)^2} + \left(m x^n y^{m-1}\right)^2 \frac{\sigma_y^2}{\left(x^n y^m\right)^2}}$$

This expression looks complicated but it simplifies to something rather simple:

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

The result is that the relative (fractional) uncertainty in f is the root of the squared sum (RSS) of individual uncertainties in x and y. Examples in the body of this document include:

$f = xy$	$\frac{\sigma_f}{f} = \sqrt{(1)^2\left(\frac{\sigma_x}{x}\right)^2 + (1)^2\left(\frac{\sigma_y}{y}\right)^2}$
$f = x/y$	$\frac{\sigma_f}{f} = \sqrt{(1)^2\left(\frac{\sigma_x}{x}\right)^2 + (-1)^2\left(\frac{\sigma_y}{y}\right)^2}$
$f = xy^2$	$\frac{\sigma_f}{f} = \sqrt{\left(\frac{\sigma_x}{x}\right)^2 + 4\left(\frac{\sigma_y}{y}\right)^2}$

Note that the result for multiplication and division is the same (division is just a power law with a negative exponent). Also note that variables that appear with a higher power are weighted more heavily in the propagation.

For some functions, especially non-linear trig functions, you may have to perform the derivatives to find how the uncertainty propagates; however, for many functions, performing the derivatives each time is not required – merely apply the equation highlighted in yellow.