

A Brief Introduction to Measurements and the Theory of Errors

Michael Goggin

Experiment is the ultimate test of a physical theory. No matter how clever a theory may be, it is wrong if it does not match experimental results. The numbers calculated in a theory are generally exact numbers, like the numbers you are familiar with from your math courses. Experimental results are not exact. There is a small amount of uncertainty related to the equipment and techniques used in the measurement. These uncertainties must be accounted for when comparing theory and experiment. The rest of this article will introduce you to the properties of measurements and the procedures for determining the uncertainties in the results.

Use the computer

Before we get started on the theoretical underpinnings of error analysis I want to be clear about something. *You should use a computer to do most of your calculations.* You will learn formulas that allow you to calculate things like averages and standard deviations on your calculator. But you should not use them except *maybe* to check the computer. Computers are wonderful tools that simplify many lab calculations. Take advantage of them. Learning how to do your calculations on a computer is a skill that you will use for the rest of your career.

Measurements and Significant Digits

A measurement will always consist of a number and a unit. The number tells us the precise value of the quantity being measured and the unit tells us what the quantity is. A measurement is meaningless without units. For example, we could measure the length of something and report the result as 2. Unless we know the units we don't know if the length is 2 mm or 2 m. The units are generally determined by the instrument with which the measurement is made. The size of the unit together with the number we measure tell us how much of the physical quantity we have measured. Every number that you report in your lab reports should have a unit with it.

The numerical value of any measurement will always contain some uncertainty. Suppose, for example, that you are measuring the width of a strip of wood, as shown in the figure below. In Figure 1 we are looking at the end of the wood strip.

The strip is almost 3 cm long. For most purposes this value would suffice, but it is obvious that the strip is a bit shorter than 3 cm. In fact it is closer to 2.95 cm. In science we try to be as precise as possible, however, because we may need to know a value more precisely later, if not now. Suppose that, in our example, we want to know the width of a table constructed with 100 identical strips. If we calculate the width using 3 cm per strip the result is 300 cm, but we get 295 cm when we use the value 2.95 cm per strip. The difference between the two is 5 cm (about 2 inches). We see that being precise can be important.

When scientists report measurements they use the concept of *significant digits* to show the precision with which the numbers were measured. The number of significant digits in a number is the number of digits we know for certain, plus one more digit, the one that has some uncertainty. In the above example we estimated the last digit in the value 2.95 cm. This last digit is the limit of the precision with which we can measure things using the ruler in the figure. We know that the first two digits are correct but the third digit has some uncertainty because it is based on our estimate of where the edge of the strip falls between the 2.90 cm mark and the 3.00 cm mark. Thus the number that we report, 2.95 cm, has three significant digits.

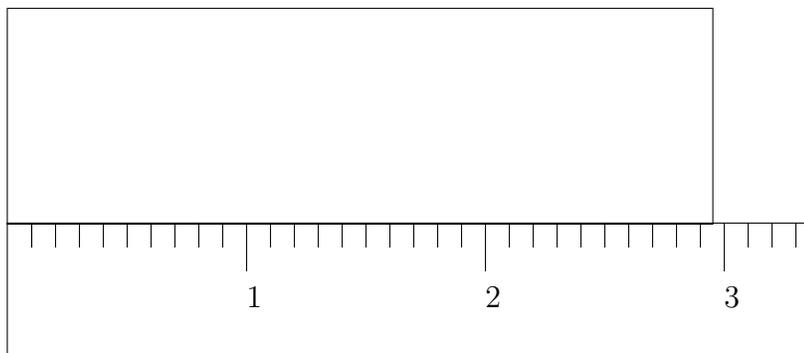


Figure 1: A sample length measurement.

If the edge of the strip *had* lined up with the 3.00 cm mark, the number of significant digits would still be three. We would write the result as 3.00 cm to indicate that we were able to measure out to the hundredths place.

The number of significant digits in a measurement is determined by the measuring instrument, it does not change. We can change the *total* number of digits in a number by changing the units with which we report the number, but the number of significant digits remains the same. For example, we could write 2.95 cm as 0.0295 m. The number of significant digits has not changed, it is still three. The leading zero in .0295 only serves to place the decimal point. Likewise, 29500 μm , has only three significant digits. But if we did not know the original measurement was 2.95 cm we would have no way of determining the number of significant digits in the number 29500 μm . After all the measurement could have been made using a micrometer. To deal with this kind of ambiguity we use *scientific notation*.

Scientific notation is a way to write numbers in which we only use the significant digits. We write numbers so that one significant digit is to the left of the decimal point and the remaining significant digits are to the right of the decimal place, we then multiply the number by ten raised to a power. The power of ten is determined by the number of places the decimal point needs to be shifted to give the number the correct magnitude. This will make more sense after we look at an example. Let's look at 29500 μm and assume it has three significant digits. In scientific notation this is written as $2.95 \times 10^4 \mu\text{m}$. The exponent of the ten is equal to four because we need to shift the decimal point four places to the right to get the original number back. Similarly, 0.0295 m is written as $2.95 \times 10^{-2} \text{m}$. Here we use a negative exponent because we need to shift the decimal point to the left. But what about our original measurement of 2.95 cm? This is a perfectly acceptable way of writing this measurement because there is no ambiguity as to the precision. If we *wanted* to write this number in scientific notation it would be written as $2.95 \times 10^0 \text{cm}$ (remember 10^0 is equal to one). It would be a bit silly, however, to write 2.95 cm as $2.95 \times 10^0 \text{cm}$. As long as there is no ambiguity in the precision of a number we won't write the number in scientific notation. However, if you are not sure whether or not to use scientific notation, then use it.

Examples of Scientific Notation and Significant Digits

In the following table the symbol 0 will be used to indicate a zero that we *know* is not significant but serves only to place the decimal point.

Number	# of Significant Digits	Scientific Notation
34.750	5	3.4750×10^1
652 <u>00</u>	3	6.52×10^4
<u>0</u> .00894	3	8.94×10^{-3}
123.734	6	1.23734×10^2

Often we wish to combine the results of different measurements to produce another physical quantity. For example, to find the area of a rug that is 3.37 m long and 2.43 m wide we would multiply the length

by the width to get 8.1891 m^2 . Our result has more precision than the original measurement, this does not make sense. The result of a calculation should never be more precise than the original numbers. In our example we would have to round our result to 8.19 m^2 . We will use the convention that a digit is rounded down if the following digit is between 0 and 4 and it will be rounded up if the following digit is between 5 and 9.

It may not be obvious to you that a calculated number shouldn't be more precise than the numbers used in the calculation. To make it clear let's look at the previous example in more detail. The length measurement of 3.37 m implies that the rug is between 3.36 m and 3.38 m because of the uncertainty in the last digit. Likewise, the width is between 2.42 m and 2.44 m . To find the range in values for the area we multiply the smallest length and width together and the largest length and width together. The area is between 8.1312 m^2 and 8.2472 m^2 . We see that we don't really know the area with much precision at all. The best we can do is say the area is $8.19 \pm 0.06 \text{ m}^2$. The final result has three significant digits, which is the same as the original numbers.

The rule for combining measurements is that the result has the same precision as the *least* precise measurement. For example, in multiplying $54.23 \text{ km/hr} \times 4.21 \text{ hr}$ to find the total distance traveled, the calculation gives an answer of 228.3083 km . This result has seven digits whereas the least precise number, 4.21 hr , has only three significant digits. Thus, the answer cannot have more than three significant digits and we round it to 228 km . Always perform the calculation using the direct measurement, and then round to the correct number of significant digits.

Systematic vs. Random Errors

There are two classes of error: systematic error and random error. Systematic error is due to a problem with the measurement equipment, the experimental setup, the reading of the meters, etc. In short, it is a problem with the system that can and should be compensated for. Systematic errors reduce the accuracy of a measurement. Random errors are due to random fluctuations in the environment of the experiment and are generally beyond the control of the experimenter. Random errors can be reduced but never eliminated. Random errors reduce the precision of the experiment.

You are probably asking yourself what's the difference between accuracy and precision? Accuracy is the property of a measurement that tells us how close the measured number is to the "true" value. The "true" value is the result we would obtain in a perfect experiment. Precision tells us the level of detail with which we know the value. Precision is also a measure of how repeatable our experiment is. As an example imagine that we measure the length of a desk top 6 times with a meter stick. The values we obtain are 1.821 m , 1.823 m , 1.822 m , 1.826 m , 1.820 m , and 1.824 m . We see that there is a small variation in the last digit. This means there is some uncertainty in that digit which is the result of random error. A systematic error would occur if the meter stick was worn down on one end. Suppose the meter stick is worn down 5 mm . That is, the end of the meter stick is actually at the 5 mm mark. That means we need to subtract 5 mm from each reading to get the length measurement. If we used an unworn meter stick the results would be 1.816 m , 1.818 m , 1.817 m , 1.821 m , 1.815 m , and 1.819 m . We see that the results from the worn meter stick are all off by 5 mm and any conclusions drawn from them would be in error.

Probability and Random Errors

As we have seen, when a large number of measurements are made, there will be random errors for each measurement. To reduce the measurements to a single value we record the *average value*. The average value is defined by

$$\text{average } x \equiv \langle x \rangle = (x_1 + x_2 + x_3 + \dots + x_n)/n \quad (1)$$

where n is the number of measurements taken, x is the label of the quantity being measured, in our example the length of the desk top, and the $x_1, x_2, \text{ etc.}$ are the individual measurements. The average value for the length of the desk top is given by

$$\langle x \rangle = (1.816 \text{ m} + 1.818 \text{ m} + 1.817 \text{ m} + 1.821 \text{ m} + 1.815 \text{ m} + 1.819 \text{ m})/6$$

or $\langle x \rangle = 1.818$ m. Note that we have used the data from the unworn meter stick.

There is obviously some uncertainty in our result since the actual data fluctuate about the average value. If the fluctuations are truly random, the data will fit a “bell curve,” much like the one that many instructors use when “curving” grades.

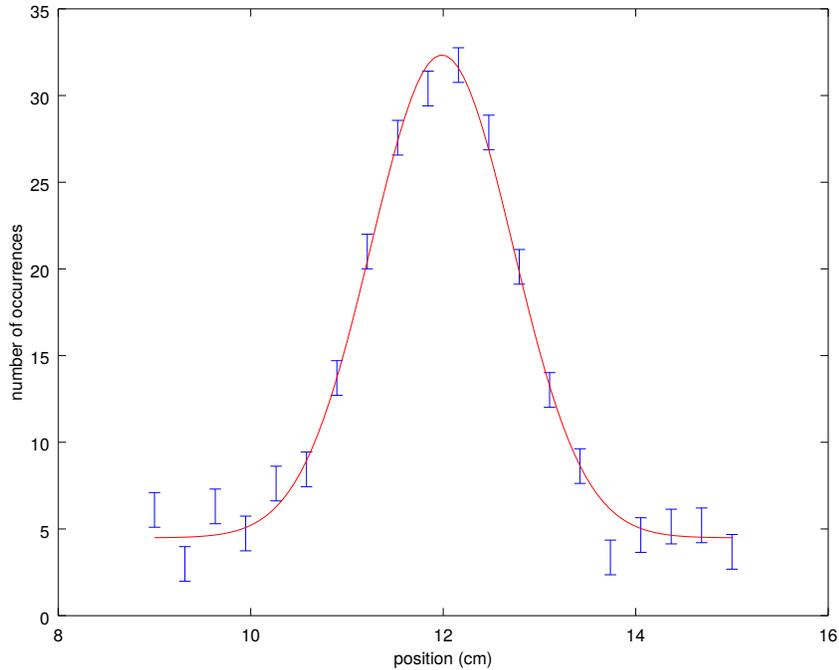


Figure 2: An example normal distribution. The number of times a particle is detected at a particular position along a linear detector. The counting uncertainty is ± 1 .

In Figure 2, above, the x -axis represents the values of the measurements taken, and the y -axis represents how many times each measurement occurs. Another way of representing the data is to divide every value in Fig 2. by the total number of data points. Then the y -axis represents the frequency of occurrence of a particular data value. In an ideal situation the data will center around the average value and decline in frequency the farther away one gets from the average. The width of the peak is a measure of the deviation of the measurements. The wider the peak, the further apart each measurement is from the average value. The width of the peak can be considered as a measurement of precision. A narrow peak is more precise than a wide peak. We can quantify this by calculating the *average deviation* of a set of measurements. The average deviation is simply the average of the *individual deviations*. The individual deviations are simply the absolute value of the difference between the *average value* and the *individual measurements*. The individual deviation of the n th measurement is given by

$$d_n = |w_n - \langle w \rangle| \quad (2)$$

In our example the individual deviations are

$$d_1 = |1.816 \text{ m} - 1.818 \text{ m}| = 0.002 \text{ m}$$

$$d_2 = |1.818 \text{ m} - 1.818 \text{ m}| = 0.000 \text{ m}$$

$$d_3 = |1.817 \text{ m} - 1.818 \text{ m}| = 0.001 \text{ m}$$

$$d_4 = |1.821 \text{ m} - 1.818 \text{ m}| = 0.003 \text{ m}$$

$$d_5 = |1.815 \text{ m} - 1.818 \text{ m}| = 0.003 \text{ m}$$

$$d_6 = |1.819 \text{ m} - 1.818 \text{ m}| = 0.001 \text{ m}$$

The average deviation is given by,

$$\langle d \rangle = (d_1 + d_2 + d_3 + \dots + d_n)/n. \quad (3)$$

More useful is the *standard deviation*, it is given by the symbol σ . The standard deviation is computed by squaring all of the individual deviations and adding them together. This number is then divided by the total number of measurements, n , and then the square root is taken. The equation is

$$\sigma_n = \sqrt{\frac{d_1^2 + d_2^2 + d_3^2 + \dots + d_n^2}{n}} \quad (4)$$

In our example, $\sigma_n = 1.97 \times 10^{-3}$. When there are a small number of measurements the standard deviation is better estimated by the following formula

$$\sigma_{n-1} = \sqrt{\frac{d_1^2 + d_2^2 + d_3^2 + \dots + d_n^2}{n-1}}. \quad (5)$$

In our example,

$$\sigma_{n-1} = \sqrt{\frac{(0.002 \text{ m})^2 + (0.000 \text{ m})^2 + (0.001 \text{ m})^2 + (0.003 \text{ m})^2 + (0.003 \text{ m})^2 + (0.001 \text{ m})^2}{6-1}}$$

which gives $\sigma_{n-1} = 2.19 \times 10^{-3}$ m. Note that the standard deviation has the same unit as the the measurement.

Equation (5) is the version of the standard deviation that is most appropriate for the data collected in this class. From now on when we talk about standard deviations we will be referring to equation (5). We will also drop the subscript for the rest of the semester and simply write σ for standard deviation. You should *not* use it explicitly when calculating standard deviations for your lab reports however. There are 12 laptop computers in the lab that you can use to run a spreadsheet to process your data. Excel has built-in functions to calculate averages and standard deviations. Most spreadsheet's standard deviation function calculates σ_{n-1} .

In addition, most scientific calculators have a statistical mode which will calculate the average value and the standard deviation. It is recommended that you have one of these and learn to use it, as these calculations can often be tedious. Your instructor will probably be able to help you even if you do not bring the calculator manual to class. It is possible to get significantly different answers for the standard deviation when using the calculator's functions versus doing the calculation by hand. The difference is due to round off. When comparing your results with others be sure to note how the calculations were made. In any case, there is no need to calculate averages and standard deviations by hand. Use the built-in functions in a spreadsheet or on your calculator.

Instrument Precision

The previous discussion assumes the variation in the data is larger than the precision with which a given measurement can be made. In such cases the standard deviation, σ , is larger than the instrument error, δ_i . Sometimes, however, $\sigma \leq \delta_i$. When this is the case, a more accurate measure of the uncertainty is given by,

$$\delta = \sqrt{\sigma^2 + \delta_i^2}. \quad (6)$$

You should be able to determine from your data and σ when you need to use this last expression.

To allow for this more general error in our measurements we will use δ to represent uncertainty in a measurement regardless of whether it is solely due to a standard deviation calculation or eq 6.

The uncertainty in a measurement by itself does not always give us a good "feel" for the precision of the experiment. For example, is an error of 1 km for a distance measurement small or large? If you answered "it depends" then you are on your way to becoming a physics professor. It depends on the size of the

measurement. It's terrible uncertainty for a measurement of the distance from OP to Barnett Hall. But it is pretty good for a measurement of Earth's radius.

A useful way of expressing uncertainty that accounts for different magnitudes of measurements is the *fractional uncertainty* (statisticians also call this the *coefficient of variation*) which is also known as the *percent uncertainty* when expressed as a percentage, both are related to the standard deviation. The fractional uncertainty is defined by

$$u = \frac{\text{uncertainty}}{\text{measured value}} = \frac{\sigma}{\langle x \rangle}. \quad (7)$$

In our example,

$$u = \frac{2.19 \times 10^{-3} \text{ m}}{1.818 \text{ m}} = 1.20 \times 10^{-3}.$$

Note that the fractional uncertainty is dimensionless, i.e. u has no units. The corresponding percent uncertainty is 0.120%. This is a small percentage which means that our measurement was reasonably precise. The fractional uncertainty will be important later on when we discuss error propagation.

When we perform an experiment we are generally interested in comparing the result with a theory, another experiment, or a standard value. In this class you will usually compare your results with standard values. Standard values are the result of many extremely precise and accurate experiments.

We already know how to determine the precision of an experiment: the smaller the uncertainty the better the precision. How do we determine the accuracy of an experiment? If a well known phenomenon, like the acceleration of gravity at sea level, is being measured then we compare the result to the standard value. If we are measuring something for the first time then we have to make sure the experimental setup gives accurate results for a standard phenomenon that is closely related to the phenomenon we are measuring. For example, if we want to measure the acceleration of gravity in Kirksville, we would want to test our experimental setup at sea level. Obviously, this method is not always practical. Another way to check accuracy is to perform the experiment several different ways. Whatever way we choose to check the accuracy of an experiment we will need to compare our result to some other value. Since we will not be doing cutting edge research in this class we will be concerned with comparisons to standard values.

Percent error is the deviation of an experimental value from the standard, or accepted value. It is simply the absolute value of the difference between the experimental value and the standard value, divided by the standard value. This result is then multiplied by 100%. The equation is

$$\%error = \frac{|x_m - x_s|}{x_s}(100\%), \quad (8)$$

where x_m is the measured value and x_s is the standard value. Suppose there is a standard value of 1.820 m for the desk measured in our previous examples. The percent error for our measurement would be given by

$$\%error = \frac{|1.818 \text{ m} - 1.820 \text{ m}|}{1.820 \text{ m}}(100\%) = 0.109\%.$$

As percent uncertainty is a measure of the precision of an experiment, percent error is a measure of its accuracy. An experiment can easily be accurate without being precise. For example, a very precise experiment with a small standard deviation (narrow peak) might not be close to the standard value at all, having a large percent error. Therefore, the experiment would be precise, but not accurate. Also, an experiment with an experimental result very close to the standard value might have a large uncertainty (wide peak in the distribution). This experiment would be accurate, but not precise. The comparison of these two quantities is one way to determine if a systematic error is present. A systematic error is one in which an inherent problem in the experiment is present: a faulty instrument, an incorrect setup, using an approximation when the conditions for its use do not apply, etc. Whatever the cause, every time the experiment is performed this error is present. A systematic error is said to be present if the absolute value of the difference between the standard value and the measured value is greater than 2 standard deviations.

$$|\langle x \rangle - x_s| > 2\sigma \quad (9)$$

This means that the standard value falls more than 2 standard deviations away from the experimental value, far outside the peak. For example, an experiment has a result of 512 ± 21 cm. If the standard value is

551 cm, the difference between the two values is 39 cm. Twice the standard deviation is 42 cm. (2×21 cm.) Since the difference is smaller, no systematic error is present. If the experimental value had been 502 cm, then the difference would be 49 cm, and a systematic error would be indicated because twice the standard deviation is still 42 cm, but the difference between the two values (49 cm) is now greater than twice the standard deviation.

Expressing Experimental Results

Now that we know about standard deviation and other measures of error we have a means of expressing an experimental result that reflects the uncertainty in the result. Whenever expressing an experimental result the final answer should always be expressed as the average value \pm the uncertainty, or $\langle x \rangle + \delta$. For our desk measurements we would express the result as $x = 1.818 \text{ m} \pm 0.002 \text{ m}$. Note that both the average value and uncertainty have units. The units are not necessarily identical. We could express the same result as $1.818 \text{ m} \pm 2 \text{ mm}$. Make sure the units are clear when expressing your results.

When we are given a number in the above form we treat the part to the left of the \pm sign as the measured value and the part to the right of the \pm sign as an uncertainty.

Error Propagation

When combining experimental results the uncertainty in the result is a function of the individual uncertainties. The method used for determining the resultant uncertainty depends on whether the numbers are being added(or subtracted) or multiplied(or divided).

When adding or subtracting numbers the uncertainty in the result is given by the sum of the uncertainties of the individual numbers in the sum or difference. If x_1, x_2, \dots, x_n denote the measurements and $\delta_1, \delta_2, \dots, \delta_n$ denote the uncertainties of the measurements then we calculate the resultant uncertainty by simply adding the δ 's as follows

$$\delta_r = \delta_1 + \delta_2 + \dots + \delta_n, \quad (10)$$

where δ_r is the resultant uncertainty. This is the same regardless of whether the x 's are being added or subtracted. For example if we have two lengths: $2.34 \text{ m} \pm 0.01 \text{ m}$ and $4.56 \pm 0.04 \text{ m}$. The result of adding these numbers is $(2.34 \text{ m} + 4.56 \text{ m}) \pm (0.01 \text{ m} + 0.04 \text{ m}) = 6.90 \text{ m} \pm 0.05 \text{ m}$. The uncertainty would be the same, 0.05 m, if the numbers were subtracted. For example, taking the difference we get $(4.56 \text{ m} - 2.34 \text{ m}) \pm (0.01 \text{ m} + 0.04 \text{ m}) = 2.22 \pm 0.05 \text{ m}$.

Things get a little more complicated when we multiply or divide numbers. For these situations the resultant fractional uncertainty, denoted u_r , is given by

$$u_r = \sqrt{u_1^2 + u_2^2 + \dots + u_n^2}. \quad (11)$$

where,

$$u_i = \frac{\delta_i}{x_i} \quad i = 1, 2, \dots, n \quad (12)$$

This keeps us from comparing numbers of drastically different magnitudes.

If we use equation (11) to calculate the fractional uncertainty for a result then we can use that result in equation (12) to find the resultant uncertainty, δ_r . Using the above notation we have

$$\delta_r = u_r x_r \quad (13)$$

where,

$$u_r = \sqrt{u_1^2 + u_2^2 + \dots + u_n^2}$$

and x_r is the result of multiplying or dividing the x 's. In the above example we could have $(2.34 \text{ m})(4.56 \text{ m})$ or $2.34 \text{ m}/4.56 \text{ m}$. For both cases the fractional uncertainty is the same,

$$u_r = \sqrt{\left(\frac{0.01 \text{ m}}{2.34 \text{ m}}\right)^2 + \left(\frac{0.04 \text{ m}}{4.56 \text{ m}}\right)^2} = 0.01$$

but the uncertainty depends on which operation we are performing. In the case of multiplication we get

$$\delta_r = (1 \times 10^{-2})(2.34 \text{ m})(4.56 \text{ m}) = 0.1 \text{ m}^2.$$

For the case of division we get

$$\delta_r = (1 \times 10^{-2}) \frac{2.34 \text{ m}}{4.56 \text{ m}} = 5 \times 10^{-3}$$

This has been a *brief* overview of measurements and errors. It is nowhere near complete. There are whole books on the subject. It has also been an introduction to the terminology and notation that will be used in this class. More information may be found in your textbook and in the library. These books may not all use the same notation, however. Be careful when comparing equations from different sources. When in doubt ask your instructor.