

---

# Introduction to Labs and Uncertainty

---

## 1. Introduction

There is no such thing as a perfect measurement. All measurements have errors and uncertainties, no matter how hard we might try to minimize them. Understanding possible errors is an important issue in any experimental science. The conclusions we draw from the data, and especially the strength of those conclusions, will depend on how well we control the uncertainties.

Let's consider an *example*: You're trying to measure something and from theory, you know the expected value should be 2.3. You make two measurements and get two very different values: 2.5 and 1.5. We can see immediately that 2.5 is rather close to the expected value while 1.5 is quite far off. However, we have not taken error into account. Each measurement has a certain amount of uncertainty, or wiggle room. Basically, there's an interval surrounding your measurement where the true value is expected to lie. If your measurements give experimental uncertainties of 0.1 and 1.0 respectively, the new measured values may be expressed  $2.5 \pm 0.1$  and  $1.5 \pm 1.0$ . The expected value falls within the range of the second measurement, but not the first!

Analyzing data and error in experiments is essential in making conclusions about the physical laws we are testing. The advent of computers and software made to manipulate large data sets has revolutionized scientist's ability to make conclusions from experimental data. In this lab course, we will be using Microsoft Excel to record data sets from the experiments and determine experimental uncertainties in calculated quantities. We will learn to use excel to propagate uncertainties and plot error bars with our data. You can download a personal copy of Microsoft Excel with your student email address from Office 365 Education <sup>1</sup>. Please note the sections introducing new Excel tools pertain to the newest version of Excel. If you are using a personal laptop with a different version of Excel, you are responsible for adapting the instructions to your version of Excel.

The purpose of this introduction is to give some basic information about statistics

---

<sup>1</sup><https://products.office.com/en-us/student?ms.officeurl=getoffice365>

and uncertainty. The techniques studied here will be essential for the rest of this two-semester lab course. These tools are important in order to arrive at good judgments in any field (like medicine) in which it is necessary to understand not just numerical results, but the uncertainties associated with those results.

## 2. Theory

### 2.1 Types of Uncertainties

Uncertainty in a measurement can arise from three possible origins: the measuring device, the procedure of how you measure, and the observed quantity itself. Usually the largest of these will determine the uncertainty in your data.

Uncertainties can be divided into two different types: systematic uncertainties and random (statistical) uncertainties<sup>2</sup>.

#### Systematic Uncertainties

Systematic uncertainties or systematic errors always bias results in one specific direction. They will cause your measurement to consistently be higher or lower than the accepted value.

An *example* of a systematic error follows. Assume you want to measure the length of a table in cm using a meter stick. However, the stick is made of metal that has contracted due to the temperature in the room, so that it is less than one meter long. Therefore, all the intervals on the stick are smaller than they should be. Your numerical value for the length of the table will then always be larger than its actual length no matter how often or how carefully you measure. Another example might be measuring temperature using a mercury thermometer in which a bubble is present in the mercury column.

Systematic errors are usually due to imperfections in the equipment, improper or biased observation, or the presence of additional physical effects not taken into account. (An example might be an experiment on forces and acceleration in which there is friction in the setup and it is not taken into account!)

In performing experiments, try to estimate the effects of as many systematic errors as you can, and then remove or correct for the most important. By being aware of the sources of systematic error beforehand, it is often possible to perform experiments with sufficient care to compensate for weaknesses in the equipment.

---

<sup>2</sup>If you were to engage in further research, random uncertainty is typically referred to as statistical uncertainty.

## Random Uncertainties

In contrast to systematic uncertainties, random uncertainties are an unavoidable result of measurement, no matter how well designed and calibrated the tools you are using. Whenever more than one measurement is taken, the values obtained will not be equal but will exhibit a spread around a mean value, which is considered the most reliable measurement. That spread is known as the random uncertainty. Random uncertainties are unbiased – meaning it is equally likely that an individual measurement is too high or too low.

From your everyday experience you might be thinking, “Stop! Whenever I measure the length of a table with a meter stick I get exactly the same value no matter how often I measure it!” This may happen if your meter stick is insensitive to random measurements, because you use a coarse scale (like mm) and you always read the length to the nearest mm. But if you would use a meter stick with a finer scale, or if you interpolate to fractions of a millimeter, you would definitely see the spread. As a general rule, if you do not get a spread in values, you can improve your measurements by using a finer scale or by interpolating between the finest scale marks on the ruler.

How can one reduce the effect of random uncertainties? Consider the following *example*. Ten people measure the time of a sprinter using stopwatches. It is very unlikely that each of the ten stopwatches will show exactly the same result. Even if all of the people started their watches at exactly the same time (unlikely) some of the people will have stopped the watch early, and others may have done so late. You will observe a spread in the results. If you *average* the times obtained by all ten stopwatches, the *mean* value will be a better estimate of the true value than any individual measurement, since the uncertainty we are describing is random, the effects of the people who stop early will compensate for those who stop late. In general, making multiple measurements and averaging can reduce the effect of random uncertainty.

*Remark:* We usually specify any measurement by including an estimate of the random uncertainty. (Since the random uncertainty is unbiased we note it with a  $\pm$  sign). So if we measure a time of 7.6 seconds, but we expect a spread of about 0.2 seconds, we write as a result:

$$t = (7.6 \pm 0.2) \text{ s} \tag{1}$$

indicating that the uncertainty of this measurement is 0.2 s or about 3%.

## 2.2 Accuracy and Precision

An important distinction in physics is the difference between the *accuracy* and the *precision* of a measurement. Accuracy refers to the closeness of a measured value to a standard or known value. For example, if in lab you obtain a weight measurement

of 3.2 kg for a given substance, but the actual or known weight is 10 kg, then your measurement is not accurate. In this case, your measurement is not close to the known value.

Precision refers to the closeness of two or more measurements to each other. Using the example above, if you weigh a given substance five times, and get 3.2 kg each time, then your measurement is very precise. Precision is independent of accuracy. You can be very precise but inaccurate, as described above. You can also be accurate but imprecise.

For example, if on average, your measurements for a given substance are close to the known value, but the measurements are far from each other, then you have accuracy without precision.

A good analogy for understanding accuracy and precision is to imagine a basketball player shooting baskets. If the player shoots with accuracy, his aim will always take the ball close to or into the basket. If the player shoots with precision, his aim will always take the ball to the same location which may or may not be close to the basket. A good player will be both accurate and precise by shooting the ball the same way each time and each time making it in the basket.

## 2.3 Numerical Estimates of Uncertainties

For this laboratory, we will estimate uncertainties with three approximation techniques, which we describe below. You should note which technique you are using in a particular experiment.

### Upper Bound (Analog)

Most of our measuring devices in this lab have scales that are coarser than the ability of our eyes to measure.

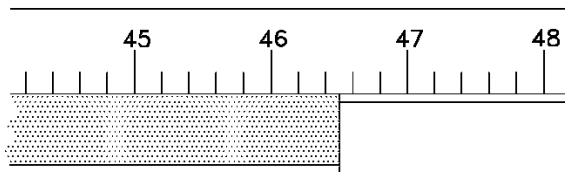


Figure 1: Measuring Length

For example in the figure above, where we are measuring the length of an object against a meter stick marked in cm, we can definitely say that our result is somewhere between 46.4 cm and 46.6 cm. We assume as an *upper* bound of our uncertainty, an

amount equal to *half* this width (in this case 0.1 cm). The final result can be written as:

$$\ell = (46.5 \pm 0.1) \text{ cm} \quad (2)$$

There will be many circumstances when the error is more complicated than simply the coarseness of the measuring tool. For example, if you find yourself measuring something that is very long or hard to line up properly with a meter stick. In this case, you may need to use some judgement of the best possible measurement to make and the uncertainty will be greater than the millimeter precision of your meter stick. **It is always best to slightly overestimate error and allow yourself some wiggle room if you feel that better represents your measurement!**

### Upper Bound (Digital)

Many measurements will be made with the use of digital equipment, such as digital calipers, scales or timers. This kind of equipment also has some inherent uncertainty which we quantify by using an *upper* bound equal to the *smallest digit* that our equipment can measure (unless specified otherwise). For instance, in the figure below, we use a scale that can measure down to two decimals of a gram, hence the associated uncertainty is  $\pm 0.01\text{g}$ .

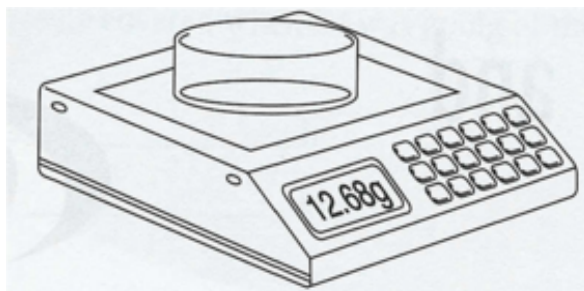


Figure 2: Digital Equipment

The final result is written as:

$$W = (12.68 \pm 0.01)\text{g} \quad (3)$$

### Estimation from the Spread (2/3 method)

For data in which there is random uncertainty, we usually observe individual measurements to cluster around the mean and drop in frequency as the values get further from the mean (in both directions).<sup>3</sup> Find the interval around the mean that contains about

---

<sup>3</sup>There is a precise mathematical procedure to obtain uncertainties (standard deviations) from a number of measured values. Here we will apply a simple “rule of thumb” that avoids the more

2/3 of the measured points: *half* the size of this interval is a good estimate of the uncertainty in each measurement.

The reasons for choosing a range that includes 2/3 of the values come from the underlying statistics of the normal (or Gaussian) distribution (see Figure 5). This choice allows us to accurately add and multiply values with errors and has the advantage that the range is not affected much by outliers and occasional mistakes. A range that always includes all of the values is generally less meaningful.

*Example:* You measure the following values of a specific quantity:

9.7, 9.8, 10, 10.1, 10.1, 10.3

The mean of these six values is 10.0. The interval from 9.8 to 10.1 includes 4 of the 6 values; we therefore estimate the uncertainty to be 0.15. The result is that the best estimate of the quantity is 10.0 and the uncertainty of a single measurement is 0.2.<sup>45</sup>

### Square-Root Estimation in Counting

For inherently random phenomena that involve counting individual events or occurrences, we measure only a single number  $N$ . This kind of measurement is relevant to counting the number of radioactive decays in a specific time interval from a sample of material, for example. It is also relevant to counting the number of left-handed people in a random sample of the population. The (absolute) uncertainty of such a single measurement,  $N$ , is estimated as the square root of  $N$  (a counting measurement is expressed as  $N \pm \sqrt{N}$ ). As an example, if we measure 50 radioactive decays in 1 second we should present the result as  $50 \pm 7$  decays per second. (The quoted uncertainty indicates that a subsequent measurement performed identically could easily result in numbers differing by 7 from 50.)

## 2.4 Relative and Absolute Uncertainty

There are two ways to record uncertainties: the absolute value of the uncertainty or the uncertainty relative to the mean value. So in the example above, you can write  $c = (5.1 \pm 0.3)$  cm or equally well  $c = 5.1$  cm  $(1.00 \pm 0.06)$ . You can see that if you multiply out the second form you will obtain the first, since  $5.1 \times 0.06 = 0.3$ . The second form may look a bit odd, but it tells you immediately that the uncertainty is 6% of the measured value. The number 0.3 cm is the absolute uncertainty and has the

---

complicated mathematics of that technique. The uncertainty using the standard deviation for the group of values in our example below is 0.2.

<sup>4</sup>Note that about 5% of the measured values will lie *outside*  $\pm$  twice the uncertainty

<sup>5</sup>While the above method for calculating uncertainty is good enough for our purposes, it oversimplifies a bit the task of calculating the uncertainty of the *mean* of a quantity.

same units as the mean value (cm). The 0.06 (or 6%) is the relative uncertainty and has no units since it is the ratio of two lengths. It's important to use proper notation when describing uncertainty to remove any unwanted ambiguity, so make sure it's clear when you are using relative or absolute errors.

## 2.5 Propagation of Uncertainties

Often, we are not directly interested in a measured value, but we want to use it in a formula to calculate another quantity. In many cases, we measure many of the quantities in the formula and each has an associated uncertainty. We deal here with how to propagate uncertainties to obtain a well-defined uncertainty on a computed quantity.

### Adding/Subtracting Quantities

When we **add or subtract** quantities, the combined uncertainty is the **sum of the absolute uncertainties** of the constituent parts<sup>6</sup>.

Take as an example measuring the length of a dog. We measure the distance between the left wall and the tail of the dog and subtract the distance from the wall to the dog's nose. So the total length of the dog is:

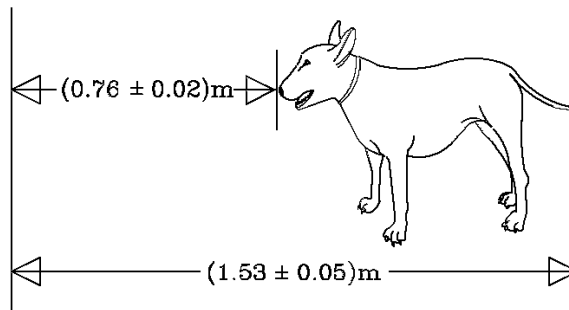


Figure 3: Measuring a Dog

$$\begin{aligned}
 \text{Length} &= (1.53 \pm 0.05) \text{ m} - (0.76 \pm 0.02) \text{ m} \\
 &= (1.53 - 0.76) \pm (0.05 + 0.02) \text{ m} \\
 &= (0.77 \pm 0.07) \text{ m}
 \end{aligned} \tag{4}$$

<sup>6</sup>The propagation of random uncertainties is actually slightly more complicated, but the procedure outlined here usually represents a good approximation, and it never underestimates the uncertainty. See the appendix for more information.

## Multiplying/Dividing Quantities

When we **multiply or divide** quantities, the combined **relative** uncertainty is the **sum of the relative uncertainties** of the constituent parts.<sup>7</sup>

Take as an example the area of a rectangle, whose individual sides are measured to be:

$$\begin{aligned} a &= 25.0 \pm 0.5 \text{ cm} = 25.0 \text{ cm} (1.00 \pm 0.02) \\ b &= 10.0 \pm 0.3 \text{ cm} = 10.0 \text{ cm} (1.00 \pm 0.03) \end{aligned} \quad (5)$$

The area is obtained as follows:

$$\begin{aligned} \text{Area} &= (25.0 \pm 0.5 \text{ cm}) \cdot (10.0 \pm 0.3 \text{ cm}) \\ &= 25.0 \text{ cm} (1.00 \pm 0.02) \cdot 10.0 \text{ cm} (1.00 \pm 0.03) \\ &= (25.0 \text{ cm} \cdot 10.0 \text{ cm}) (1.00 \pm (0.02 + 0.03)) \\ &= 250.0 \text{ cm}^2 (1.00 \pm 0.05) \\ &= 250.0 \pm 12.5 \text{ cm}^2 \\ &= 250 \pm 10 \text{ cm}^2 \end{aligned} \quad (6)$$

Note that the final step has rounded both the result and the uncertainty to an appropriate number of significant digits, given the uncertainty on the lengths of the sides.

Remarks: Note that uncertainties on quantities used in a mathematical relationship always increase the uncertainty on the result. The quantity with the biggest uncertainty usually dominates the final result. Often one quantity will have a much bigger uncertainty than all the others. In such cases, we can simply use this main contribution.

## Multiplication by a Constant

Multiplying a value by a constant leaves the relative error unchanged. This is equivalent to multiplying the absolute error by the same constant. For example, suppose we are trying to find the circumference of a circle knowing its radius as  $r = 1.0 \pm 0.1 \text{ cm}$  with error. We would calculate the circumference with error as follows.

$$\begin{aligned} C &= 2\pi r \\ C &= 2\pi(1.0 \pm 0.1) \\ C &= 6.3 \pm 0.6 \text{ cm} \end{aligned} \quad (7)$$

---

<sup>7</sup>Our calculation of the uncertainty actually overestimates it. The correct method does not add the absolute/relative uncertainty, but rather involves evaluating the square root of the sum of the squares. For more information please refer to the appendix.



## Powers and Roots

When raising a value to a certain power, its **relative uncertainty is multiplied by the exponent**. This applies to roots as well, since taking the root of a number is equivalent to raising that number to a fractional power.

Squaring a quantity involves multiplying its relative uncertainty by 2, while cubing a quantity causes its relative uncertainty to be multiplied by 3.

Taking the square root of a quantity (which is equivalent to raising the quantity to the  $1/2$  power) causes its relative uncertainty to be multiplied by  $1/2$ . For example, if you know the area of a square to be:

$$\text{Area} = 100 \pm 8 \text{ m}^2 = 100 \text{ m}^2 (1.00 \pm 0.08) \quad (8)$$

then it follows that the side of the square is:

$$\text{Side} = 10 \text{ m} (1.00 \pm 0.04) = 10.0 \pm 0.4 \text{ m} \quad (9)$$

The most general rule for finding the error in powers and roots is mathematically represented as follows.

$$f(x) = x^n \quad (10)$$

$$\frac{\sigma_{f(x)}}{f(x)} = |n| \frac{\sigma_x}{x} \quad (11)$$

Where  $\sigma$  is the *absolute* uncertainty and  $f(x)$  is some power or root of  $x$ .

## Other Functions

If you need to calculate the error of a calculation that does not fit into one of these rules (such as trigonometric functions or logarithmic ones), here is a manual method that you can use.

Based upon the error of the quantity that you determined, you can find the maximum and minimum values of the quantity that you are calculating. The value that you found should be roughly midway between these two quantities. Then if you split the difference between the maximum and minimum you should obtain a reasonable estimate of the error. Mathematically, you would do so as follows.

$$\sigma_{f(x)} = \frac{f(x + \sigma_x) - f(x - \sigma_x)}{2} \quad (12)$$

Here is an example: Suppose you measure an angle to be  $(47.3 \pm 0.5)^\circ$  and you want to determine the error of  $\sin(47.3 \pm 0.5)^\circ$ . You find that  $\sin(47.3) = 0.735$ . Based upon your reported uncertainty, you know that your angle could be as large as

$47.8^\circ$  and as small as  $46.8^\circ$ , and therefore you should calculate  $\sin(47.8) = 0.741$  and  $\sin(46.8) = 0.729$ . So your calculated value is  $0.735$  but it can be as low as  $0.729$  and as high as  $0.741$  and therefore, if you halve the difference between  $0.729$  and  $0.741$  you get a reasonable error estimate of  $0.006$ . So you should report your value as  $0.735 \pm 0.006$ .

## 2.6 Best-Fit Line

In most research laboratories, plotting measurements is found to be the preferred method of reviewing the data and quantitatively measuring the relationship between the experimental variables. This is effective because we often have some idea of the expected relationship between the variables *a priori*. In these labs, this expected relationship is almost always arranged to be a straight line. But even if we know that the ideal points fit on a precise straight line, experimentally measured data points will not always lie on a single line – because the measurements always have intrinsic uncertainty. Therefore when the points are plotted, we should include error bars on both axes to indicate the uncertainties in the data. Because real measurements do not all lie on a single straight line, there are a variety of possible lines you might choose to fit the data.

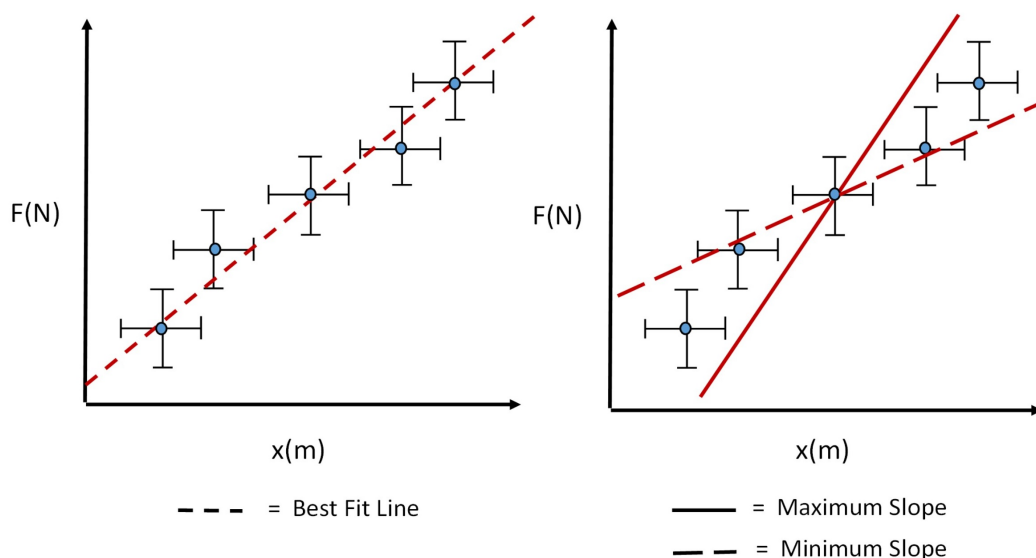


Figure 4: Left: an example best-fit line. Right: the maximum and minimum possible slope from our data used to calculate uncertainty in the best-fit line. Notice how we have drawn the lines on the outer bounds of the error bars to achieve the maximum and minimum possible slope within the error bars.

How do we know which line represents the best fit? There is an exact mathematical

procedure to obtain the best-fit line, but this is usually a very tedious calculation which is outside the scope of this lab. For experiments in this course, you will be using Excel's built-in fitting function for data. The process for doing so will be explained below.

Remark: Often in our experiments the data points will not look as nice as in the examples here. One or several points may not be close to any best-fit line you try. Such anomalous points may occur, for example, because of a mistake in measuring. In such cases, it is acceptable to ignore these anomalies when estimating the best-fit line (and of course you must note this fact down in your lab report). Dropping anomalous points must be done with extreme care and only rarely (if you know the point is not physically meaningful).<sup>8</sup> It is better to choose a line with as many points above the line as below. If you are not sure of your measurements, it is better to re-measure or to take more data points.

## 2.7 Numerical Statistics

The previous discussions of uncertainty and error tell us how we can quantitatively describe our inability to make perfect single measurements. However, in real physics experiments, very rarely do we draw conclusions from a single data point. As such, it is essential that we know how to quantify error in sets of data. The 2/3 methods as discussed in Section 2.3 provides a good estimation of data statistics, but we can more rigorously calculate data set statistics. In statistics, a data set can be well described by the following four fundamental quantities: mean, median, mode, and standard deviation. The mean of a data set is the sum of all numbers in the data set divided by the number of points in the set. It is defined in the following manner.

$$\text{Average} \equiv \bar{x} = \sum_i \frac{x_i}{N} \quad (13)$$

The median of a data set is the middle value in a set of numbers listed in increasing order. The mode is the number that occurs the most number of times in the data set. The standard deviation describes how the numbers in the data set are distributed around the mean. It is defined as follows.

$$\text{standard deviation} \equiv \sigma = \sqrt{\sum_i \frac{(x_i - \bar{x})^2}{N}} \quad (14)$$

These four statistical quantities give us enough information to characterize the distribution of our data set. For example, let's consider the two following Data Sets.

---

<sup>8</sup>More than once, data points that did not behave as theory predicted turned out to be new effects and led to Nobel prizes!

										Mean	Median	Mode	$\sigma$
Set 1	9	8	11	13	10	10	12	6	9	9.8	10	10	2.2
Set 2	11	0	10	40	2	3	10	10	4	9.8	10	10	11.4

Notice how both Data Sets have the same mean, median, and mode, which tells us the data points in each set are centered around the mean value of 9.8. However, the standard deviations are quite different. The large standard deviation in Data Set 2 tells us there must be outliers in the data set which increase the distribution. Whereas, the relatively small standard deviation in the first data set tells us the numbers in set 1 are clustered closely together. Standard deviation is especially important because it tells us exactly how distributed the numbers are around the mean value, which gives an indication of how error affects the spread of data points (for example, see figure 5 for the canonical "bell curve" distribution, also known as the gaussian distribution). The 2/3 method as discussed in Section 2.3 is an approximation to the standard deviation since  $2/3 \sim 66\%$ , which roughly corresponds to the first standard deviation (see figure 5).

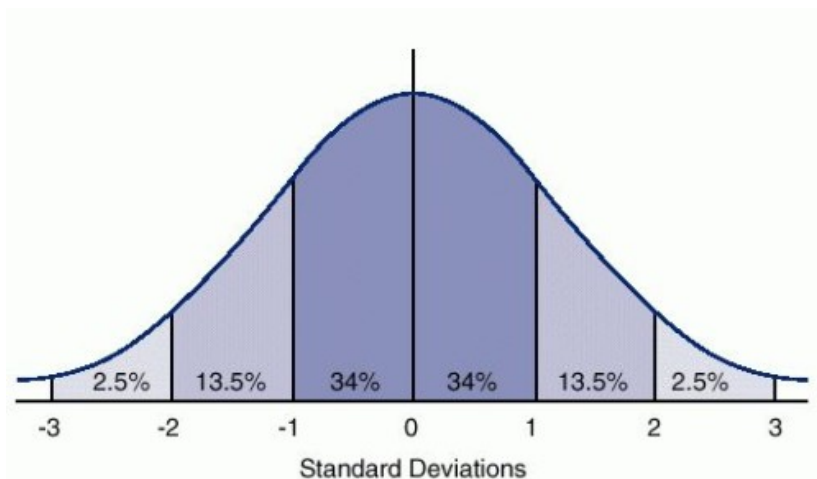


Figure 5: An example of a Gaussian distribution, also known as a bell curve.  $\sim 68\%$  of the data points are within 1 standard deviation,  $\sim 95\%$  of that data points are contained within 2 standard deviations,  $\sim 99.5\%$  of the data points are contained within 3 standard deviations, etc...

## 2.8 Number of Significant Digits

The number of significant digits in a result refers to the number of digits that are relevant. The digits may occur after a string of zeroes. For example, the measurement

of 2.3 mm has two significant digits. This does not change if you express the result in meters as 0.0023 m. The number 100.10, by contrast, has 5 significant digits<sup>9</sup>.

When you record a result, you should use the calculated error to determine how many significant digits to keep. Let's illustrate the procedure with the following example. Assume you measure the diameter of a circle to be  $d = 1.6232$  cm, with an uncertainty of 0.102 cm. You now round your uncertainty to one or two significant digits (up to you). So (using one significant digit) we initially quote  $d = (1.6232 \pm 0.1)$  cm. Now we compare the mean value with the uncertainty, and keep only those digits that the uncertainty indicates are relevant. Finally, we quote the result as  $d = (1.6 \pm 0.1)$  cm for our measurement.

Suppose further that we wish to use this measurement to calculate the circumference  $c$  of the circle with the relation  $c = \pi \cdot d$ . If we use a standard calculator, we might get a 10 digit display indicating:

$$c = 5.099433195 \pm 0.3204424507 \text{ cm} \quad (15)$$

This is not a reasonable way to write the result! The uncertainty in the diameter had only one significant digit, so the uncertainty of the circumference calculated from the diameter cannot be substantially better. Therefore we should record the final result as:

$$c = 5.1 \pm 0.3 \text{ cm} \quad (16)$$

(If you do intermediate calculations, it is a good idea to keep as many figures as your calculator can store. The above argument applies when you report your results!)

## 3. Working With Excel

### 3.1 Plotting with Excel

An important set of data analysis tools in Excel are plotting and linear fit functions. You will need to plot and fit data many times throughout this lab course, so make sure you are familiar with this section. Below is a walkthrough of plotting and fitting a set of data with error in excel.

1. Before plotting, you need to have 4 columns with data: x data, y data, x error data, and y error data. Make sure you have entered the information into excel.
2. First select your x data and y data (you can select multiple boxes in excel by holding down the ctrl button while selecting). Make sure to select your x data first or your x and y axes will be switched.

---

<sup>9</sup>Another way to find the number of significant digits is to convert to scientific notation, and count the number of digits in the mantissa (also significand or coefficient). For example: for  $1.2 \times 10^2$ , there are two significant digits in 1.2.

- Choose the subheading “insert”, then “Scatter”, then “Scatter with straight lines and markers”. Now your x and y data should be plotted without error bars (see figure 6).

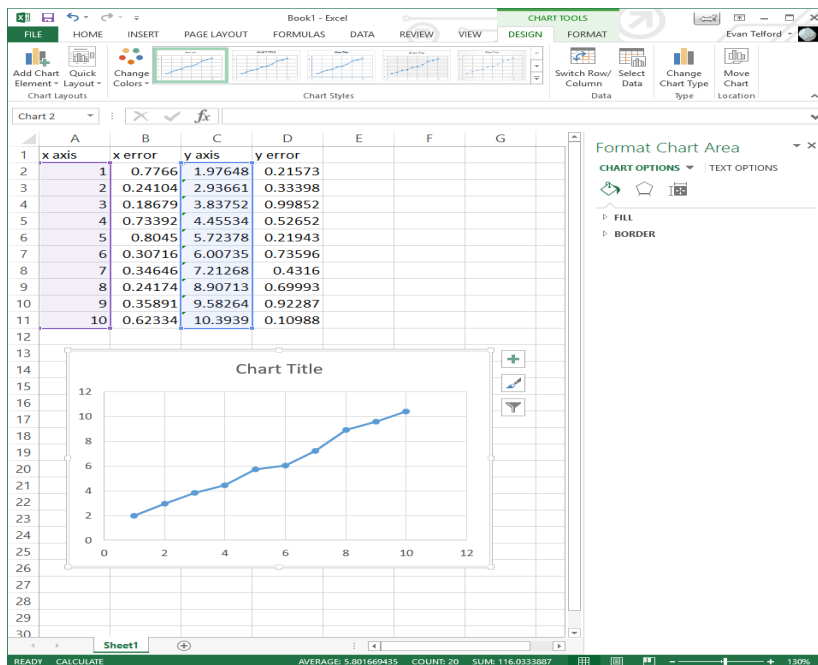


Figure 6: Selecting x and y data and creating a lined scatter plot in excel.

- To include error bars select your chart, then click the “plus” marker on the top right of the chart. Check the box titled “error bars”. Now some basic error bars should appear on the plot. These are not based on the error bar data in your excel document, they are standard error bars.
- To change them so they match your error bar data, select the x axis error bars on your chart, and format the error bars by clicking the “Custom” selection, then “specify value”.
- It should now prompt you for positive error values and negative error values. Delete “{1}” from the two boxes, and select your error bars using the cursor. Your chart will now have the correct error bars (see figure 7).
- Repeat steps 5-6 for your y data.
- To linear fit your data, right click on your data in the plot and select “Add Trendline”. Check the “Linear”, “Display Equation on chart”, and “Display R-square value on chart” boxes.

9. Now the slope, intercept, and R squared values will be displayed on your chart. R squared is a measure of how well the line fits your data. It should be close to 1 and at the very least greater than 0.9.
10. Finally, you must always include a chart and axis titles (with units where applicable). To do this, first double click the chart title text box included by default and type out a descriptive title for your chart.
11. Then, select your chart and click on the “plus” marker on the top right of the chart, and check the box titled “Axis Titles”. Again, double click on each axis title and type out a label describing the content of each axis.

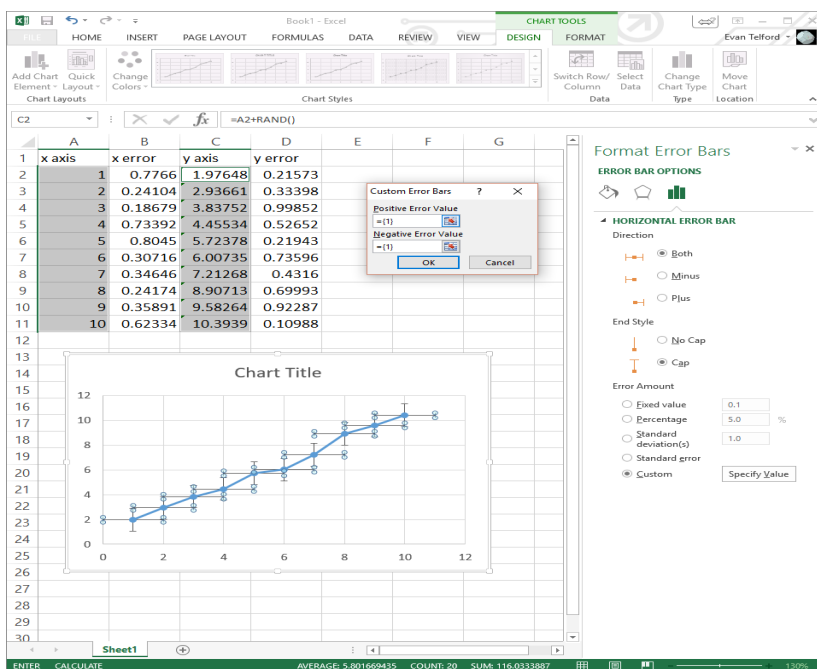


Figure 7: Using your own data set to create x and y error bars in excel.

### 3.2 Finding Error in Slope

The previous steps will help plot your data, but in order to draw conclusions, you must add uncertainty. Excel has a built in function called LINEST which finds the standard error for a linear fit. Using the same columns of data from before, the walkthrough below will help you calculate the error so you can propagate it further in the experiment.

1. The LINEST function is an array-type function, meaning it will output more than one number. Start, by highlighting a 2-by-3 section of empty cells (two columns, three rows).

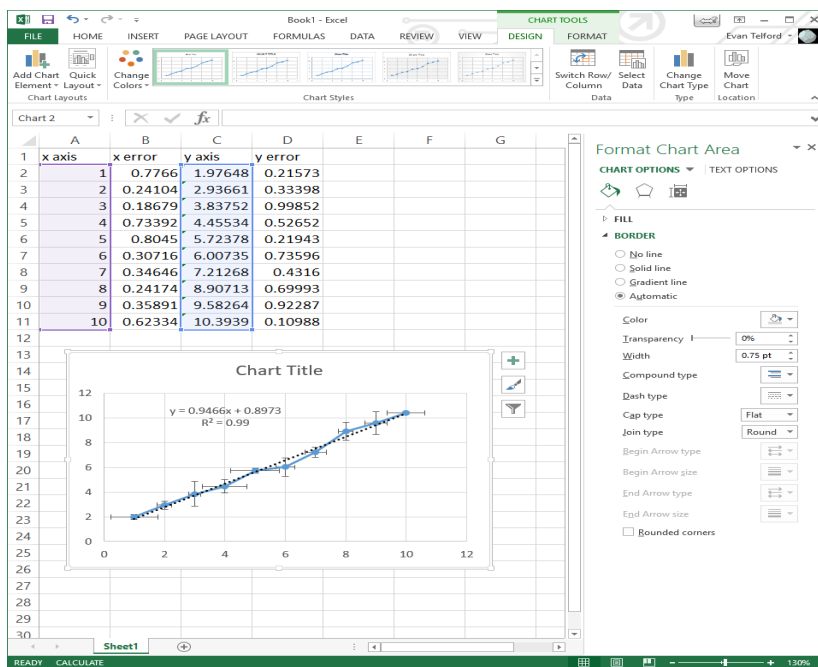


Figure 8: Using excel to perform a linear fit and return the intercept and slope.

2. With these six cells highlighted, in the input box at the top of the screen type “= LINEST(” and add the proper arguments. The arguments should be the list of y-values, the list of x-values, TRUE, and TRUE. For example: = LINEST(C2:C11, A2:A11, TRUE, TRUE).
3. Press CONTROL+SHIFT+ENTER.  
*Note that on a Mac, this may be COMMAND+SHIFT+ENTER.*

Your results should have filled in that 2-by-3 section in the following way:

Slope	Y Intercept
Error of Slope	Error of Y Intercept
$R^2$	Error in Y

Table 1: LINEST function output in Excel.

### 3.3 Helpful Commands

Your TA will guide you through the relevant excel commands necessary for data analysis, however a list of some relevant excel commands are listed below. A list of all excel



commands can be found on the Microsoft Office website<sup>10</sup>

ABS	Returns the absolute value of a number
AVERAGE	Computes the average of the selected data set
COS	Calculates cosine of a number
DEGREES	Converts radians to degrees
EXP	Returns e raised to the power of a given number
LN	Returns the natural logarithm of a number
MEDIAN	Finds the median of a data set
MODE.SNGL	Finds the most commonly occurring number in a data set
PI	Returns the value of pi
POWER	Returns the result of a number raised to a power
SIN	Calculates the sine of a number
SQRT	Calculates the square root of a number
STDEV.P	Calculates the standard deviation based on the entire population
STDEV.S	Estimates the standard deviation based on a sample
SUM	Calculates the sum of a data set
TAN	Calculates the tangent of a number

<sup>10</sup><https://support.office.com/en-us/article/Excel-functions-alphabetical-b3944572-255d-4efb-bb96-c6d90033e188>

## 4. Practice Questions

### Estimation of Uncertainty:

1. You have the following distribution of measured values:

1	(1)
2, 2, 2	(3)
3, 3, 3, 3, 3, 3	(6)
4, 4, 4, 4, 4, 4, 4, 4	(9)
5, 5, 5, 5, 5, 5, 5, 5, 5, 5, 5, 5, 5, 5, 5	(15)
6, 6, 6, 6, 6, 6, 6, 6, 6, 6, 6	(11)
7, 7, 7, 7, 7, 7, 7, 7	(8)
8, 8, 8, 8	(4)
9, 9	(2)
10	(1)

Estimate the uncertainty using the 2/3 estimate.

2. Estimate the mean and uncertainty of the following group of values:

1.6 s, 1.3 s, 1.7 s, 1.4 s, 1.6 s, 1.2 s,

3. In a radioactive decay you get 16 counts. What is the absolute uncertainty of the number of counts? What is the relative uncertainty?

4. In a radioactive decay you get 1600 counts. What is the absolute uncertainty of the number of counts? What is the relative uncertainty?

**5. How many counts should you get so that the relative uncertainty is 1% or less?**

### Significant digits:

6. How many significant digits has  $l = 0.0254$  m?

7. Write  $t = 1.25578 \pm 0.1247$  s with two significant digits (in the uncertainty).

### Propagation of Uncertainty:

**8. For a pendulum with  $l = 1.0 \pm 0.1$  m, you measure a period of  $t = 2.0 \pm 0.2$  s. What is the value of the earth's acceleration g?**

**9. You measure the volume of a box by measuring the length of the single sides. For the lengths of the single sides you get:**

$$a = 10.0 \pm 0.1 \text{ cm} \quad b = 5.0 \pm 0.2 \text{ cm} \quad c = 7.5 \pm 0.3 \text{ cm}$$

What is the volume of the box (including uncertainty and units) in  $\text{cm}^3$ ?  
 What is the volume of the box in  $\text{m}^3$ ?

10. You measure the following quantities:

$$A = 1.0 \pm 0.2 \text{ m} \quad B = 2.0 \pm 0.2 \text{ m} \quad C = 2.5 \pm 0.5 \text{ m/s}$$

$$D = 0.10 \pm 0.01 \text{ s} \quad E = 100 \pm 10 \text{ m/s}^2$$

Calculate the mean and uncertainty of:

- a)  $A + B =$
- b)  $A - B =$
- c)  $C \cdot D =$
- d)  $C/D =$
- e)  $C \cdot D + A =$
- f)  $\frac{1}{2}E \cdot D + C =$
- g)  $A \cdot B/(A - B) =$

Include units! For e)–g) perform it step by step.

Relative and Absolute Uncertainty:

11. What is the relative uncertainty for  $v = 12.25 \pm 0.25 \text{ m/s}$ ?

12. What is the absolute uncertainty if the mean value is 120 s and the relative uncertainty is 5%?

**13. Given the following measurements, which one has the highest absolute uncertainty and which one has the highest relative uncertainty?**

$$l = 10.0 \pm 0.2 \text{ m} \quad l = 10.0 \text{ m} (1.00 \pm 0.03)$$

$$l = 7.24 \text{ m} (1.00 \pm 0.04) \quad l = 12.5 \pm 0.25 \text{ m}$$

14. Given the following measurements which one has the highest absolute uncertainty and which one has the highest relative uncertainty?

$$l = 10.0 \pm 0.2 \text{ m} \quad t = 7.5 \pm 0.2 \text{ s}$$

$$d = 5.6 \text{ cm} (1.00 \pm 0.04) \quad v = 6.4 \times 10^6 \text{ m/s} (1.00 \pm 0.03)$$

*Caution:* Don't get tricked!

15. Consider a measurement you made in which you counted the number of cars that passed in 1 hour on various highways in the united states. The data set containing the number of cars per hour is given in the following table.

Trial	1	2	3	4	5
$N_{\text{cars}}$	523	143	496	1045	869

1. Calculate the absolute error and the relative error in each measurement.
2. Which measurement has the lowest absolute error?
3. Which measurement has the lowest relative error?
4. Which measurement would you say is the most precise?
5. Calculate the rate in cars *per second* for each measurement with error.

## Appendix

The rules we've given for propagating uncertainties through a calculation are essentially correct, and intuitively make sense. When adding two quantities together, if one has an uncertainty of  $\Delta x$  and another has an uncertainty of  $\Delta y$ , the sum could indeed range from  $(x + y) - (\Delta x + \Delta y)$  to  $(x + y) + (\Delta x + \Delta y)$ . This implies that the proper way to find the uncertainty of  $(x + y)$  is to add their respective absolute uncertainties.

There is, however, a small problem – this overestimates the uncertainty! Since  $x$  and  $y$  are equally likely to be wrong by either a *positive* amount or a *negative* amount, there's a good chance that the respective errors of each variable will partly cancel one another out. To account for this, a more accurate way to estimate uncertainty turns out to be to add uncertainties *in quadrature*. This means:

### Adding/Subtracting Quantities

$$(A \pm \Delta A) + (B \pm \Delta B) = (A + B) \pm \sqrt{(\Delta A)^2 + (\Delta B)^2} \quad (17)$$

### Multiplying/Dividing Quantities

$$A \left(1 \pm \frac{\Delta A}{A}\right) \times B \left(1 \pm \frac{\Delta B}{B}\right) = (A \times B) \left(1 \pm \sqrt{\left(\frac{\Delta A}{A}\right)^2 + \left(\frac{\Delta B}{B}\right)^2}\right) \quad (18)$$

While this method gives a closer approximation to what the true propagated uncertainty should be, it is clearly a more complex calculation. In the limited time available to complete your experiment and lab report, you may use the simpler, earlier uncertainty calculation method provided, and avoid this complicated calculation. But do remember that the simpler method *overestimates* the total uncertainty.