INTRODUCTION TO MEASUREMENT AND ERROR

Purpose

To make some simple measurements and to estimate their best values and uncertainties.

Theory

The word "**error**" suggests a kind of mistake or blunder. But in science that is not the case. All scientific measurements have an **uncertainty** in them. No physical quantity, such as a time or length or mass, etc. can be measured *exactly*. "Error" means this uncertainty. These errors are not blunders or mistakes and they cannot be completely eliminated by simply being very careful. But we try to make these errors as small as reasonably possible and we try to determine a good estimate of how big they are. We will use "error" and "uncertainty" interchangeably.

The theory relating to this part of the lab is that of statistical analysis and uncertainty. You may have had to perform statistical analyses in other classes at the College. There is a shared purpose visiting the topic of statistical analysis in several different contexts: measured experimental data can depend on many factors and may not be exactly repeatable - hence the idea of uncertainty. In order to estimate not only the value of a quantity but the uncertainty in its measurement, we perform several measurements and then calculate the spread of the data. Any result of a measurement of a quantity "X" can be presented as a range Δx around the **best estimated value** x and is written as: (measured value of *X*) = $x \pm \Delta x$. Here, x is normally the **average** of several measured values. Δx is the estimate of **uncertainty** in the quantity X. For example, when you measure the time of ticks of 5 cycles of a pendulum during this lab, you will find that the values of this quantity vary somewhat. So what value should be stated as the time? The answer is that we quote the average value *and* the estimate of the uncertainty. Uncertainty comes from several sources, but there are two general forms of uncertainty, or error.

Firstly, there are "systematic" errors that bias a measurement in one direction. For example, if a mass balance is not reset to zero, then all masses measured on it thereafter will be incorrect by the same amount. Secondly, there are "random" errors, which can arise from measurement difficulties, fluctuations in temperature, pressure and so on. We try to minimize both types of error as much as possible. In your lab reports this semester, you should not only state what kind of error is present in your experiment, but also the *source* of the error. In this way, your lab reports will be as particular to your experiment as possible.

From *multiple* measurements of the same quantity, we may find the *average* or *mean* value. For example, if the above quantity "X" is measured *N* times, the average value, x, of the quantity is

$$x = (\Sigma_i x_i)/N = (x_1 + x_2 + x_3 + \ldots + x_N)/N \quad (1)$$

So this is our *best estimated value*. (Sometimes the average quantity is denoted by a line over the symbol: $\bar{x}$.)

Our method for finding the estimate of the uncertainty now follows: Find ***N/6*** rounded off to the *nearest integer***.** Eliminate that number of the largest x_i's and the same number of the smallest x_i's. *From the remaining* (approximately 2/3 of the) x_i's, find ½ the difference between the largest remaining value and the smallest remaining value. That is our estimate of Δx. *In general, the greater the spread of results of a measurement, the larger is the uncertainty* Δx. For large *N* it represents a probability of about 2/3 (67%) that a random measurement will lie within Δx of the mean value x. The result of this analysis is we may quote our estimate of X and the uncertainty in it by stating that

$$(\text{measured value of } X) = x \pm \Delta x.$$

In Part 1 of this lab, you will make measurements of time that will lead to a statement of average time and uncertainty. *Note: the choice of letter (X, x etc.) here is just for demonstration purposes.*

Two rules regarding uncertainties and best values:
Rule 1. *Experimental uncertainties should* ***almost*** *always be rounded to one significant figure.*
Rule 2. *The last significant figure in any stated best value should* ***usually*** *be of the same order of magnitude (in the decimal position) as the uncertainty.*

Running the experiment

Note: The data sheet for the experiment is on page 4

Part 1: Statistical analysis for time measurement

1) Open the simulator https://phet.colorado.edu/sims/html/pendulum-lab/latest/pendulum-lab_en.html
Click 'Intro'. Check the option of 'Stopwatch' at the bottom left. This simulator shows you a mass attached to a string. This is called a pendulum. It is like the one used in old clocks. Keep all default settings.

2) Click and drag the pendulum to 10^o to the right and release it. The pendulum starts swinging.

3) One cycle of the pendulum could be, from the extreme left to the extreme right and back to the extreme left. As the pendulum is swinging and as it hits the extreme left start the stop watch. Count five cycles then stop the stopwatch. Record this time in your data sheet in trial 1 in table 1. Reset the stopwatch by clicking reset on the stopwatch. Keep the pendulum swinging.

4) As the pendulum is swinging (with the same settings) repeat step 3 for 11 more trials, so you will have a total of 12 trials. Record your values in table 1. Pause the simulator by clicking .

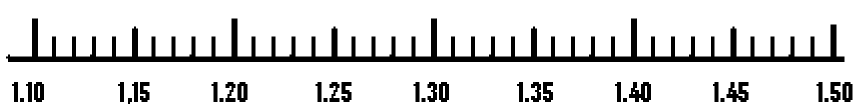
5) Calculate the average time which is the best estimate as mentioned in the introduction and calculate the uncertainty by the method mentioned in the introduction theory. Record your result in the data sheet.

Computations

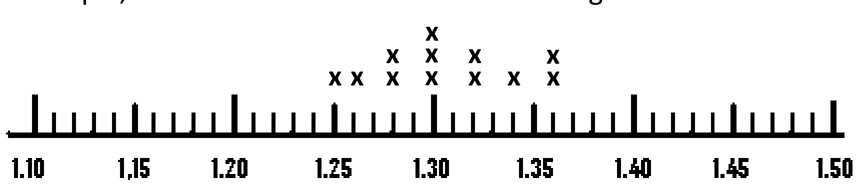
Students should imagine that they have all just measured some deep fundamental constant of nature, not just the time of 5 cycles of a pendulum, and that they did this at, say, Brooklyn National Laboratory (BNL). The world is waiting for their results. Copenhagen calls and wants to know. What do we tell them? Do we give them all 12 numbers? No! Copenhagen wants only *two* numbers: Our *best value*, *t*, and the *uncertainty*, Δ*t*, associated with it. This is how experimental results are communicated. So our task is now to reduce these 12 numbers to these two crucial quantities.
Then we will report what we have accomplished as *Best Value* ± *Uncertainty*, i.e.,

(measured value of time) = $t \pm \Delta t$.

Students should use graph paper. For example, if the largest number on the list is 1.36 sec and the smallest is 1.25 sec, the student can draw:



student record the data as he marks an **x** on the scale for each number, producing a visual plot of the distribution of measured values. For example, the final result for all 12 numbers might look like this:



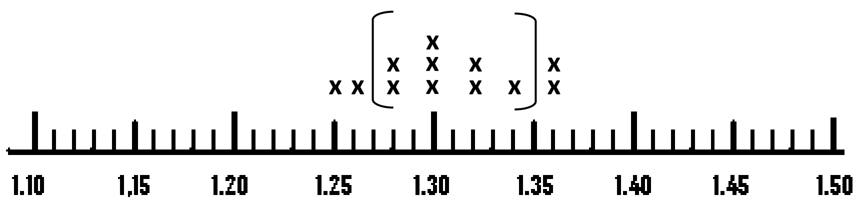
This picture of the results is called a *histogram*. **(*You, the student, should be constructing it, and keep it in your data*.)** Several things are worth noting about it. First, it is unique to this specific set of measurements, i.e., if we were to do 12 more measurements and make a new histogram, chances are it will **not** look exactly like this one. (This effect is more pronounced for a small number of measurements. If we had 12 *million* points to plot rather than just 12 points, the plot would look like a smooth curve, the famous bell-shaped curve.)

But even with just 12 points, we often see some interesting features, leading to our second observation, namely that within the distribution, there are often seen regions of *low* popularity near the edges of the distribution and a region or regions of *higher* popularity near the middle.

Our method for finding the best value and uncertainty of the distribution now follows:

First, the BEST VALUE, *t*, is simply the average of the measurements. What is *N*, the number of measurements? It's 12 in our example. So add up the 12 values of the x's and divide by 12. We get 1.30583 s.

Next, for the UNCERTAINTY, *Δt*: We note that *N*/6 = 12/6 = **2**. (If *N* were not exactly divisible by 6, we would have rounded off *N*/6 to the nearest whole number.) The **2** largest and the **2** smallest measurements are "outliers" and now we attempt to place a set of brackets around *the remaining* x's. This will generally be approximately the MIDDLE TWO-THIRDS of the x's. In our case, it's the middle 8 x's (i.e., 2/3 of 12). An obvious possible set of brackets is this:



The UNCERTAINTY is then the *distance from the midpoint to either bracket*. In practice, we would *subtract the smallest* x *within the bracket from the largest* x *within the bracket and take half of that difference*. So we get ½(1.34 - 1.28) = 0.03. In this case, then, the uncertainty is ***Δt* = 0.03 sec**. So our result is

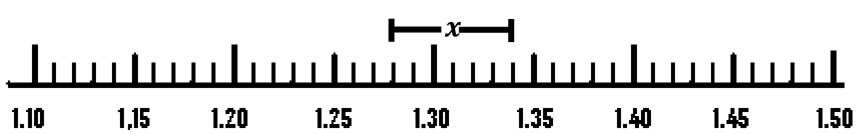
$$\text{\textit{(measured value of time)}} = 1.30583 \pm 0.03\ sec.$$

But hold it. From **Rule 2** we see that we must give our result as

$$\textbf{\textit{(measured value of time)}} = \mathbf{1.31 \pm 0.03}\ \boldsymbol{sec.}$$

*Record the **result** on the data sheet.* Note that this method is not always unambiguous. There are some distributions for which more than one set of brackets may be drawn. However, differences between final results are usually insignificant compared to the overall uncertainty of the data. Students in this lab should learn to think like experimentalists rather than mathematicians! To an experimentalist, there is no difference between 1.30 ± 0.03 and 1.31 ± 0.03.

So our efforts for this measurement are well represented numerically by 1.31 ± 0.03 sec. It should also be noted that we can represent this same information *visually* on our scale as follows:



This is called a *data point.* The $\boldsymbol{x}$ represents the best value, and the bars to the sides (unfortunately called *error bars*) represent the estimated uncertainty in the best value. *The student should draw this on the graph paper.*

Part 2: Measuring Speed

Speed can be calculated as distance travelled divided by time. Since we can measure distance in meters, m and time in seconds, s, the unit for speed is meters per second, m/s.

1) Open the simulator http://physics.bu.edu/~duffy/HTML5/motion_diagrams.html

2) Set the velocity of Car 1 (red) to 2 m/s and set the velocity of Car 2 (blue) to 4 m/s. Keep all other settings unchanged as default (initial position = 0 m and acceleration = 0 m/s^2).

3) Click play. Click pause as car Car 2 (blue) becomes close to 20 m, and fine adjust to exactly 20 m using the step advance icon Step >> or the step rewind icon << Step. Record the time and distance traveled by the red car and that traveled by the blue car in table 2.

4) Click play and repeat the procedure of step 3 to find the position of the red car and the time for the instants when the blue car reaches 40 m, 60 m, 80 m, and 100 m. Record the values in table 2.

5) Calculate $\frac{d}{t}$ (distance traveled/time) for each data point and record in table 2.

6) a) What quantity does $\frac{d}{t}$ represent? b) Which of the two cars has a larger $\frac{d}{t}$?

7) Calculate the best estimate (average) of $\frac{d}{t}$ for both cars. Note that our cases here are constant speed. Since the simulator gives, for repeated trials, exact values for d and t, we will not have an uncertainty in the measurements. So here since the simulator uses a formula in its code to calculate $\frac{d}{t}$, not a physical lab measurement, the uncertainties for $\frac{d}{t}$ are taken to be zero.

Data Sheet

Name: **Group:** **Date experiment performed:**

Part 1: time measurements

Table 1

Trial no.	time for 5 cycles (s)
1	
2	
3	
4	
5	
6	
7	
8	
9	
10	
11	
12	

Part 2: Speed measurements

Table 2

d (red car) (m)	d (blue car) (m)	t (s)	$\frac{d}{t}$ (red car) (m/s)	$\frac{d}{t}$ (blue car) (m/s)
0	0	0		
	20			
	40			
	60			
	80			
	100			

Answers to questions of part 2 step 6:

a) b)

Part 2, step 7:

measured speed of slow car= ± 0 m/s

measured speed of fast car= ± 0 m/s

part 1 step 5: (measured time) = _______±_________s