# What Affects the Period of a Pendulum?

# LiveLab 3: Analysis of the data, Part 1

### Once you have collected your data, it is time to analyze it! In order to analyze data properly, you must consider uncertainty. This guide will walk you through how to treat uncertainty in your lab. For a more complete guide to uncertainty, see the [Error and Uncertainty guide](https://smccd.instructure.com/files/5075072/download?download_frd=1) in Files/Labs.

## Calculating the Period for Each Setup

For most of you, you took multiple measurements with each setup, and each measurement was the time for multiple swings. Your first step of data analysis it to average the multiple measurements, and divide the average by the number of swings, to get a single data point for the period of each setup. Do this in MSExcel or Google Sheets: learning to use these tools is essential to handling data efficiently. Use the [Excel Crash Course](https://smccd.instructure.com/files/5209719/download?download_frd=1) if you are not familiar with using these. (The tutorial works just as well for Google Sheets, except for the plotting part, which you can skip). The product should be a series of two tables (one the data where you vary the length, one for the data where you vary the amplitude) that look something like this (but with more rows):

|  |  |  |
| --- | --- | --- |
| Amplitude (deg) | Length (cm) | Period (s) |
| 15 | 50 | 4.53 |
| 25 | 50 | 4.62 |

1. Report your final measurements of the period in a spreadsheet in the GradeScope document.

## Propagation of Uncertainty

Often, we do calculations with the raw data. When we do this, we must figure out the uncertainty in the calculated quantity is as well. Above, we did two things we need to do with our data: 1) averaged the results of multiple time measurements taken with the same setup and 2) divided the total time measured by the number of swings. Fortunately, this uncertainty propagation is pretty straightforward:

* When we average multiple measurements with the same setup, we expect the average value to usually be more accurate than the individual values. There are exceptions to this rule: for example, if you take multiple measurements with a miscalibrated scale, that will not help you improve accuracy. But if we assume the errors are random (rather than systematic) and follow a typical (normal) distribution, you can estimate the uncertainty in the average by dividing the uncertainty in the individual measurements by the square root of the number of measurements. So, if you estimated each induvial time measurement had an uncertainty of ±2.0 seconds (which, by the way, is probably a bad estimate), and you took 4 measurements with each setup, the uncertainty in the average would be .
* When we divide a measured quantity by a constant x, the uncertainty is also divided by x. So, for example, if you measured the time for three swings, and, the uncertainty in the average measurement was ±1 s, then the uncertainty in the time for single swing would be ±1/3 s.

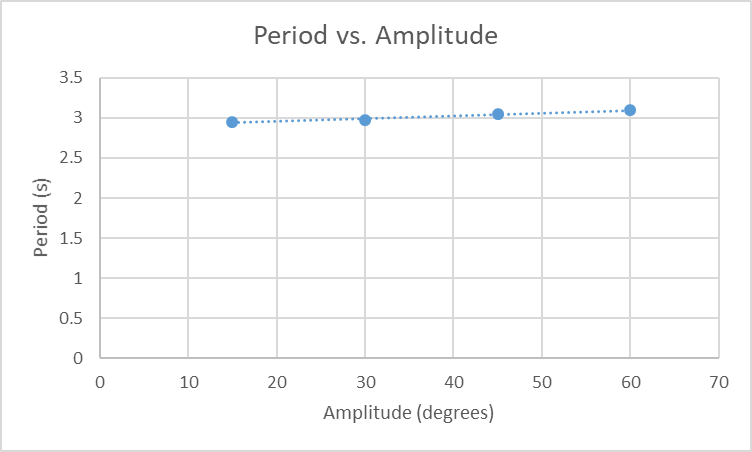
The Error and Uncertainty packet provides more detail on methods of propagation of uncertainty in general.

1. Starting with your uncertainty estimates in the time measurements that you made in the Data Collection lab, use the propagation of uncertainty techniques described above to find a “final” estimate in the uncertainty in your period measurements. Write out your calculations (so that I can understand your process) and report them in the GradeScope document.

## Looking for Significant Effects of Amplitude

With your data on the effect of amplitude (angle) you will likely not observe any dramatic changes in period. However, you may see some more subtle changes, in which case, the question arises: are these changes real effects, or are they just products of experimental error (noise)?

One way to answer this question is to make graphs of the data. The dependent variable (period) should be on the y-axis, and the independent variable (amplitude) should be on the x-axis. We can then use a computer program to make a best-fit line of your data. The result might look something like this:



In the above data we see a slight upward trend, and the computer tells is the slope is 0.0035 s/degrees. But is that slope meaningfully different from zero? That is, is the trend ***significant***? We can answer this question by using our uncertainties for period and amplitude: knowing these uncertainties can help us find the uncertainty in the computer’s estimate of the slope.

The procedure for estimating the uncertainty in a slope is a bit complex, so I have provided you with a Excel spreadsheet, called “Plot and Monte Carlo” that will allow you to do this. The spreadsheet does three things:

1. Makes a plot of the data, with error bars based on your uncertainty estimates
2. Makes fits a line to the data.
3. Uses a procedure called a Monte Carlo simulation\* to estimate the uncertainty in the slope (and intercept) of the best fit line.

The “[Plot and Monte Carlo](https://smccd.instructure.com/files/5323928/download?download_frd=1)” spreadsheet is located in Files/Lab Materials. However, it only works on PCs that have Excel installed. If you do not have access to such a device, use [VMWare](https://smccd.instructure.com/courses/35406/files/5209697/download) to access our Learning Center computers. The Plot and Monte Carlo spreadsheet can be found in the “Simulations” folder on the desktop. Your G number is the password to log into the computers.

Once you have accessed the spreadsheet, clear the fake data that is originally in the spreadsheet and copy/paste your data into the appropriate x and y columns. Also enter your uncertainties. (Be warned that the programming of the spreadsheet is somewhat fragile: do not do calculations or other manipulations within this spreadsheet, and if it stops working, load a fresh copy). When you have entered your data and uncertainties, click the “Find Uncert” Button. The program will then find a best-fit line, whose slope and intercept are displaced in row 2. It will also run the Monte Carlo simulation to find the uncertainty in the best-fit slope and intercept.

We care here about the uncertainty in the slope: say that, I got an uncertainty in the slope of 0.0053 s/deg after having run the Monte Carlo Simulation. My slope itself was 0.0035 s/deg, so I could report my experimental slope as 0.0035 ± 0.0053 s/deg. Here, with the uncertainty displayed, you can see that a slope of 0 s/deg is a possible experimental result. In such as case (where the absolute value of the experimental slope is less than the uncertainty), we can say that there is no significant effect\*\* of the amplitude on the period. If the magnitude of the slope had been bigger than the uncertainty, we could say that our experiment detected a significant effect.

You can also click on the “plot” tab to see a plot of your data with error bars. Change the plot title and axes titles to reflect the actual data. (Don’t forget units on the axes titles!)

1. Follow the instructions above and copy the plot from the plot tab (with proper title and axes labels) into your GradeScope document.
2. Report the slope and uncertainty in the slope (don’t forget units). Using this information, state whether there is or is not a significant effect of the amplitude on the period according to your data.

Don’t forget to save your Excel spreadsheet (and, if using VMWare, to transfer the file to your computer) in case you need to use it in the future.

Next Steps

You should have found that the length had a very obvious effect: this effect is so obvious that there is no real need to do an analysis of whether the effect is significant. Instead, we would like to determine a mathematical model for this effect. The next week’s lab will help you do this.

Footnotes

\* The Monte Carlo simulation works by having the computer randomly generate a new set of data based on the original data and associated uncertainty. For example, if a data value is 11 ± 1, the simulation might generate a new value of 11.5 or 10.3. In this particular simulation, the new values are distributed normally about the original data point, with a standard deviation equal to the uncertainty. The computer does this for each data value, generating a new data set. The computer then fits a line through this new data set, and finds the slope and intercept of the line. Since computers don’t get bored, we then have the computer repeat the process 500 times, generating 500 new data sets (all based on the original data and uncertainties) and associated slopes and intercepts. Finally, the computer looks at how much the values of the slopes (or intercepts) vary from each other by taking the standard deviation of all the slopes (or intercepts). This variation is an estimate of the uncertainty in the slope (or intercept) of the original best fit line.

\*\* Significance tests are much more complex and subtle in actual scientific practice, but the direct comparison of the slope to the uncertainty is good enough for us here.