

✓ Tutorial 2: Uncertainty

Developed by Rebeckah Fussell for Cornell Physics Labs. Adapted by Cristina Schlesier for Spring 2025.

✓ Introduction

Experimental physics is all about the active creation of scientific knowledge. If you've studied physics before, you've probably mostly learned about existing scientific knowledge. You've studied the models and frameworks that explain the universe reasonably well, e.g. Newton's Law of Gravitation. You've practiced putting these models to use with mathematics. But where do the models come from? Often we are taught about the creation of scientific knowledge as if it comes from thin air. Some genius had an idea in the bathtub and wrote it down and that's the end of that.

These myths about the creation of knowledge obscure the fundamental essence of science as we understand it today. In science, we conspire with the physical world to create knowledge. We design experiments and make measurements that allow the physical world to communicate its true nature to us. Science is an active, iterative, never-ending process of asking questions and developing new ways of understanding the physical world. This process is mediated by experiment and measurement. At its core, science is about measurement.

This tutorial is concerned with a crucial aspect of measurement known as **uncertainty**. (You may also see the less precise term 'error' used interchangeably [but we'd like to avoid this](#). The word 'error' implies a mistake, something that should be fixed and avoided. Uncertainty, on the other hand, is a fundamental, necessary, and generally unavoidable component of measurement.)

Say you'd like to measure the period of a pendulum. You use a stopwatch with high instrumental precision (many decimal places) and you are confident in your reaction time. You might get a measurement of 2.500 s. How confident are you in that measurement? Perhaps you try again just to check. This time you measure 2.721 s. What value do you quote as the best estimate for the period of the pendulum? And how confident will you be in that estimate?

A measurement is only as good as its uncertainty. If I tell you I've measured the acceleration due to gravity to be 9.610 m/s^2 , it is possible that I've made a world-shattering discovery. After all, the average value measured by professional physicists is 9.8067 m/s^2 . These two values are clearly distinct. Or are they?

Perhaps my experiment wasn't very precise and the uncertainty on my measurement is $\pm 0.2 \text{ m/s}^2$, in which case these values *are* consistent. But what if my uncertainty was $\pm 0.002 \text{ m/s}^2$?

Statistical and Systematic Uncertainties

There are two fundamentally different sorts of uncertainties associated with any measurement procedure, namely **statistical** (or random) and **systematic** uncertainties. Statistical uncertainties come from random fluctuations that occur during the measurement process. The example of the pendulum period given above illustrates the presence of statistical uncertainties in the measurement. I may measure the period many times and always get slightly different values.

On the other hand, systematic uncertainties arise from all sorts of sources and basically entail everything that is not a random, statistical fluctuation.

When we make a series of repeated measurements, statistical uncertainties cause us to measure a spread of values around the "true"¹ value. In contrast, systematic uncertainties may cause the measurement to be offset from the "true" value, even if the individual measurements are close to each other.

While we will develop tools for understanding and mitigating statistical uncertainty in this tutorial, we will leave the discussion of systematic uncertainty for another time. In general there are no simple prescriptions for eliminating systematic uncertainties. To a large extent this task requires experience and in-depth knowledge of the experimental apparatus.

✓ Distributions

Although measurements come with the possibility of statistical uncertainty, we still expect our measurements of a single quantity to cluster around a single value². The presence of random uncertainty means that any value is possible in principle, but that different values may have different probabilities of occurring. A **Gaussian** (or normal) **distribution** illustrates this behavior perfectly. It is often safe to assume our measured values will follow a Gaussian distribution.

A nice way of visualizing this behavior is by plotting our measurements in a **histogram**. A histogram tells us the frequency of a value in our dataset. If I measure the period of the pendulum 10 times, my stopwatch may provide me with a set of values $T = (1.62, 1.76, 1.80, 1.53, 2.0, 1.7, 1.75, 1.81, 1.88, 1.67)$ in seconds. Run the code below to load the histogram of this data.

```
1 import numpy as np
2 import matplotlib.pyplot as plt
3
4 period = [1.62, 1.76, 1.80, 1.53, 2.0, 1.7, 1.75, 1.81, 1.88, 1.67]
5
6 fig, ax = plt.subplots(1)
7 ax.hist(period)
```

```
8 plt.xlabel("Period T (in s)")
9 plt.ylabel("Number of measurements")
```

With this histogram, we get a general sense of where the average value may lie and how spread out our measurements are -- which is tied to how confident we are in our estimate of the quantity. Running the code multiple times should also drive home the point that no single measurement would be sufficient to perfectly extract the pendulum's period from our experiment.

✓ Determining (statistical) uncertainty

Now that we've motivated the idea of uncertainty, we can finally begin to explore the statistical tools to help us understand repeated measurements. We will continue to use the example of measuring the period of a pendulum.

Important note: the code cells throughout this tutorial will build off of previous code cells. It is important to **run every code cell (SHIFT + ENTER) in the document up to the point where you are working** every time you open this tutorial. If you get an error message (particularly one that says that a particular variable is not defined) after attempting to run a code cell, first make sure that you have run every previous code cell. If you see a warning like 'Warning: This notebook was not authored by Google', just click 'Run anyway'.

The code below generates a histogram showing the results of 25 trials measuring the period T of a pendulum. The precision of the measuring instrument, also known as *instrumental precision*, was ± 0.005 s because the instrument (a stopwatch) displayed values to the nearest 0.01 s. We will estimate instrumental precision as half of the last known digit.

✓ Q1:

Run the code cell below and observe the histogram. Do not worry about the details of the code inside the `pendulum_hist()` function. Run the cell a few more times so that different histograms are generated.

Hint: If you see an error that `np` is not defined, make sure you imported the packages above. It is good practice to make sure that you have run every code cell up to the point in the document you are working in every time you open the tutorial.

```

1 def pendulum_hist(N):
2     # Generate a normal distribution, center at x=0 and y=5
3     mean = 1.5
4     sd = 0.1
5     data = np.random.normal(mean, sd, N)
6
7     #Create a histogram of [this] *these* data
8     fig, ax = plt.subplots(1)
9     ax.hist(data,label = "hi");
10    plt.xlim(xmin=1.15, xmax = 1.85)
11    plt.title(r"Pendulum Period Measurements ($\pm$ 0.005 s)")
12    ax.errorbar(x=[mean+sd, mean -sd,mean], y = [N/14, N/14, N/14], yerr = [N/200, N/200, N/200])
13    ax.errorbar(x=[mean+sd/np.sqrt(N), mean -sd/np.sqrt(N),mean], y = [N/7, N/7, N/7], yerr = [N/200, N/200, N/200])
14    plt.xlabel("Period T (in s)")
15    plt.ylabel("Number of measurements")
16    plt.text(1.46, N/8, r"$\Delta \bar{x}$ (standard uncertainty of the mean)")
17    plt.text(1.46, N/17, r"$\sigma$ (standard deviation)")
18    plt.show()
19    return data
20
21 data=pendulum_hist(25);

```

Q. How does the histogram change each time you generate a new one? What stays the same?

[Your answer here -- keep your responses in boldface so we can easily distinguish your work from the main text of the tutorial]

When we take repeated measurements, it is useful to calculate the mean $\bar{x}$ of all of the measurements, as a measure of the center of our collection of measurements. The mean is defined as:

$$\bar{x} = \frac{1}{N} \sum_{i=1}^N x_i$$

where N is the number of measurements and x_i is the i^{th} value of the measurement x . In Python, we can calculate the mean using a built-in function in the *numpy* package (nicknamed *np* here). Run the code cell below to try it on the data from our histogram above.

```

1 # an example of calculating mean
2 m = np.mean(data)
3 print('the mean period of the pendulum is ', m, 'seconds')

```

When we take the same measurement over and over, we will often find some natural variation in the data. If the variation is a result of statistical uncertainty, we can characterize the width of statistical variation with the **standard deviation**, a statistical measure of the uncertainty in a single

measurement. The standard deviation characterizes the spread in the distribution of individual values due to statistical uncertainty in the measurement process. In the histogram above, the standard deviation is labeled with σ .

Some of the desired properties of standard deviation are:

1. We want to incorporate all available points into the calculation.
2. The standard deviation should stay approximately the same as we add more points with the same level of variation.
3. Standard deviation should increase as we measure more points further away from the mean.
4. The units of standard deviation should be the same as the units of the measurement.
5. The standard deviation should treat values on either side of the mean the same.

Standard deviation is calculated with the following formula:

$$\sigma = \sqrt{\frac{1}{N-1} \sum_{i=1}^N (x_i - \bar{x})^2}$$

where N is the number of measurements³, x_i is the i^{th} value of the measurement x , and $\bar{x}$ is the mean.

✓ Q2:

Q. Explain how the formula for σ takes into account the five properties above.

[Your answer here]

Run the code cells below in sequence to define and use a function for standard deviation on your data above. Remember, you must define a function before you can use it.

```
1 #define a function that calculate the standard deviation
2 def standard_deviation(data):
3     N = len(data)
4     return np.sqrt(np.sum((data - np.mean(data))**2)/(N-1))
```

```
1 # calculate standard deviation of the data displayed in the histogram above
2 print('the standard deviation of the period of the pendulum measurement is ', standard_c
```

✓ Q3:

Q. Look back at the histogram and the formula for standard deviation. Why can the standard deviation σ also be viewed as an estimate of the uncertainty in any single measured value? *Hint:*

consider the averaging behavior of the equation for standard deviation.

[Your answer here]

✓ Q4:

Q. If we added 10 more measurements to our data set, the standard deviation σ would not necessarily get larger. Why not?

[Your answer here]

Although we calculate standard deviation from our measurements, we use it as a representation of the physical phenomenon. After taking repeated measurements of the same phenomenon, we expect that around 68% of our measurements will fall within one standard deviation of the mean, if the measurements are approximately *normally* distributed (or distributed according to a *Gaussian* distribution).

A single measurement is informative, but we gain much more information from the mean of a set of multiple measurements. The uncertainty in the mean of multiple measurements, therefore, should have a smaller uncertainty than that of a single measurement (i.e. the standard deviation). We define the uncertainty in the mean as:

$$\delta\bar{x} = \frac{\sigma}{\sqrt{N}}$$

where σ is the standard deviation and N is the number of measurements.

This definition serves to reward you for taking more data, so that the uncertainty in the mean of a collection of many measurements is smaller than the uncertainty in any single measurement.

Note: Read the symbol δ as "uncertainty in". As discussed above, $\bar{x}$ is the symbol for the mean, so $\delta\bar{x}$ should be read as the mathematical notation for "uncertainty in the mean". It does *not* mean the "uncertainty times the mean".

✓ Q5:

Q. Say you have taken m measurements (where m is some arbitrary number). If you do one hundred times as many measurements (i.e., $100m$), how much smaller is the uncertainty in the mean compared to when you had only had m measurements?

[Your answer here]

The standard uncertainty of the mean is shown in green in the histogram for Q1. The uncertainty itself is just an estimate, so we usually round it to one or at most two digits.

Below we define a function for calculating the uncertainty in the mean.

Run the code cells below in sequence to calculate the standard uncertainty of the mean of our histogram data.

```
1 def standard_unc_of_mean(data):
2     N = len(data)
3     return standard_deviation(data) / np.sqrt(N)
```

```
1 # an example of calculating standard uncertainty of the mean
2 print('the standard uncertainty in the mean is', standard_unc_of_mean(data), 's')
```

✓ Q6:

Q. Is this measurement of the uncertainty larger or smaller than the standard deviation? Explain whether or not that makes sense.

[Your answer here]

Let's go back to generating histograms of the results of 25 pendulum measurement trials.

✓ Q7:

Run the cell below and notice that the mean is printed out below the histogram. Copy this mean into the array `all_means` in the second code cell. Generate at least nine more datasets by re-running the first code cell, and store each mean as a new entry in `all_means`, with each mean separated by a comma.

```
1 hist_data = pendulum_hist(25)
2 print("mean is " + str("%.4f" % np.mean(hist_data)) + " s") #for the curious, "%.4f" short
```

```
1 all_means = np.array([copy,your,means,here,...,...]) #at least 10 means for 10 different
2
3 print("standard deviation of our means is " + str("%.4f" % standard_deviation(all_means))
```

✓ Q8:

Q. How does the standard deviation of our list of means compare to the standard uncertainty of the mean calculated in Q6 (round each to 1 significant digit), and why does that make sense?

[Your answer here]

Because the uncertainty of the mean $\delta\bar{x}$ is actually less than the uncertainty in each individual measured value σ by a factor $1/\sqrt{N}$, we can reduce the uncertainty in our best estimate of a measured quantity by taking more and more repeated measurements.

Run the code cell below, which will give you a histogram for 900 repeated trials measuring the period T of our same pendulum.

```
1 big_data = pendulum_hist(900);
2 print("mean is " + str("%.4f" % np.mean(big_data)) + " s")
3 print("standard deviation is " + str("%.4f" % standard_deviation(big_data)) + " s")
4 print("standard uncertainty of the mean is " + str("%.4f" % standard_unc_of_mean(big_data)) + " s")
```

✓ Q9:

Q. How do the mean, standard deviation, and standard uncertainty of the mean of this data set compare to the data set with 25 repeated trials? Explain in your own words why these results make sense.

[Your answer here]

In comparing this histogram with the original histogram for 25 trials, three main features are apparent:

1. The mean appears to be roughly the same as it was for 25 trials.
2. The standard deviation σ appears to be roughly the same as it was for 25 trials. The additional trials have simply "filled out" the distribution more smoothly so that the histogram has a more consistent shape.
3. The uncertainty of the mean $\delta\bar{x}$ is noticeably less than it was for 25 trials.

Recall also that the instrumental precision of these measurements was limited to ± 0.005 s by the digital display of the stopwatch.

✓ Q10:

Q. How does the uncertainty of the mean of our data set with 900 measurements compare with the instrumental precision of the measuring instrument?

[Your answer here]

✓ Q11:

Q. What does this say about the possibility of "beating instrumental precision" by taking repeated measurements?

[Your answer here]

✓ Q12:

Q. Write the best estimate for the pendulum's period T from these 900 trials in the form $T = \bar{T} \pm \delta$. Use the template given below.

Double click this cell to edit it!

$T = [\textit{InsertValueHere}] \pm [\textit{InsertValueHere}]$.

✓ Q13:

Q. If the standard deviation for repeated measurements of the period of this pendulum is around 0.1 s, roughly how many repeated measurements of the pendulum period would be needed to reduce the uncertainty of the mean to $1/2$ of the instrumental precision of 0.005 s (i.e., as small as 0.0025 s)?

[Your answer here]

Submit a PDF of this document, with all code cells run and displaying output, to the Gradescope assignment.

Footnotes

1 We use the word "true" with some caution here. The idea that there exists a true value or that we can ever know it is debatable. As it stands, when we perform our experiment we certainly do not know the so-called true value. Otherwise we wouldn't bother with the measurement at all! This means we must determine our confidence in the measurement entirely from the data and experimental conditions – not by comparison to any *seemingly-true-but-actually-not-true* value, such as a theoretical value or previous measurement. [back up](#)

2 This is sufficiently true for most simple, classically-modeled phenomena. It may not be true for more complex systems and quantum mechanically-modeled phenomena. [back up](#)

3 Don't worry too much about the fact that we subtract 1 from N in the denominator. We can think of it as making sure we can't calculate the standard deviation of one measurement. It really comes from using one degree of freedom in our data to calculate the mean, which we need for the standard deviation. [back up](#)

References

A Practical Guide To Data Analysis for Physical Science Students by Louis Lyons