UNCERTAINTY PROPAGATION AND ERROR ANALYSIS

Uncertainty

In Physics 0030, students test physical theories by taking measurements and comparing the theoretical predictions with the measured values. However, the measurements taken (e.g. with meter sticks and timers) are never absolutely exact; instead they are approximations of the exact value. In fact, no measurement is absolutely exact; all measurements have a 

measurement uncertainty

associated with them. In order to test a theory, uncertainty in measurements must be quantified in order to determine whether or not an experimental result is consistent with the theory.

Example 1:

Suppose Alice and Bob were testing the theory: “The speed of light is exactly 1 foot per nanosecond.” Each of them performs one measurement of the speed of light. Alice states her measurement of the speed of light as:

0.98 ± .03 feet per nanosecond,

while Bob states his measurement of the speed of light as:

0.98 ± .005 feet per nanosecond.

Is either of these measurements consistent with the theory “The speed of light is exactly 1 foot per nanosecond”?

As Alice’s range of possible values can be written as (.95 ft/ns, 1.01 ft/ns), her result includes the value 1 foot per nanosecond as a possible value for the speed of light, so this experimental result is likely consistent with the theory. However, since Bob’s range of values is (.975 ft/ns, .985 ft/ns), which does not include the value 1 ft/ns, Bob’s experimental result is very unlikely to be consistent with the theory.

Example 1 shows the importance of uncertainties in testing scientific hypotheses. Without stating the uncertainty of an experimental result, it is impossible to make conclusions about the result’s consistency with a theory, and one cannot judge the experimental result’s usefulness (one would not know whether a value has an uncertainty of 1%, 10%, 100%, 1000%, etc., and one typically avoids using values with 1000% uncertainty).

Measurement uncertainty arises from the limitations of the measurement instrument used. For example, standard meter sticks with millimeter graduations are not suited for making length measurements to nanometer precision. To quantify the measurement uncertainty associated with a given measurement, the experimenter must study the measurement device and must gauge her own confidence in the measurement method. Unlike most aspects of science, two different
experimenters using the exact same measurement device may arrive at different measurement uncertainties, completely legitimately. Because the experimenter must judge her own confidence in the measurement to determine the measurement’s uncertainty, there is a certain subjective element to estimating measurement uncertainty.

**Example 2:**

Suppose Bob and Alice are measuring the length of a metal rod by holding it up against a meter stick with half-centimeter graduations.

Bob sees that the smallest increment shown on the meter stick is a half-centimeter. He believes that the most precise measurement he can reliably take is to align one end of the rod up with the 0 cm mark on the meter stick, and then he can determine which graduation mark is closest to the other end. Since the end of the rod is closer to 4.5 cm than to 4.0 cm, his reported measurement is 4.5 cm. The range of rod lengths that are closest to the 4.5 cm mark is 4.25 cm to 4.75 cm, so Bob reports his measurement as 4.5 ± .25 cm.

Alice, on the other hand, has very sharp eyesight and believes she can determine if the end of the rod is closer to the midpoint between two graduation lines than to the nearest graduation line. Examining the rod, she believes its end is closer to the midpoint between 4 cm and 4.5 cm, so she reports her measurement as 4.25 cm. The range of rod lengths that would be closest to that midpoint would be 4.125 to 4.375, so Alice reports her measurement as 4.25 ± .125 cm.

In Example 2, both experimenters did a good job in thinking about their uncertainty and making a quantitative estimate of it. This shows that uncertainty involves the experimenter’s judgment of her own confidence in the measurement, and thus uncertainty may be subjective.

Often, we are interested in quantities that are not directly measured, but instead must be calculated using the measured values. Because the measured values have an uncertainty, if we calculate a quantity using the measured values, then the calculated quantity will also have uncertainty. The process of determining the uncertainty in a calculated value is called *uncertainty propagation*. Here we illustrate two methods for uncertainty propagation—the first is a simple, intuitive method, while the second is less intuitive but more mathematically correct.
Example 3:

Suppose Claire was interested piece of printer paper, and unbeknownst to her, the piece of paper was made very carefully in a papermill so that its dimensions were 8.5” x 11”, within one thousandth of an inch. Claire measures the width and height of the piece of paper using a meter stick with half-centimeter graduations, and reports her measured values as:

\[
\text{Width} = 21.5 \pm .25 \text{ cm} \\
\text{Height} = 28.0 \pm .25 \text{ cm}
\]

Claire would like to calculate the area of the sheet of paper, so to get her best estimate of the sheet’s area, she multiplies the width by the height:

\[
\text{Area} = (21.5 \text{ cm}) \times (28.0 \text{ cm}) = 602 \text{ cm}^2
\]

To quickly approximate the uncertainty in her stated value for the area, Claire first does the following to get a rough estimate of the area’s uncertainty. In order to find the largest possible area from her possible values of width and height, Claire calculates:

\[
\text{Largest area} = (21.5 + .25 \text{ cm}) \times (28.0 + .25 \text{ cm}) = (21.75 \text{ cm}) \times (28.25 \text{ cm}) = 614 \text{ cm}^2
\]

and to find the smallest possible area consistent with her width and height measurements, she calculates:

\[
\text{Smallest area} = (21.5 - .25 \text{ cm}) \times (28.0 - .25 \text{ cm}) = (21.25 \text{ cm}) \times (27.75 \text{ cm}) = 590 \text{ cm}^2
\]

Using these as the bounds for her range of possible values of area, Claire reports her experimental result as 602 ± 12 cm². Claire’s Physics 0030 TA is very happy with her method of estimating uncertainty and gives her full marks. Claire, however, would like to use her result in a more advanced course and recalls that there is a more mathematically correct way to propagate uncertainty. There are two non-calculus-based equations that are helpful for uncertainty propagation:

\[
Z = X + Y \text{ or } Z = X - Y \\
\Delta Z = \sqrt{\Delta X^2 + \Delta Y^2}
\]

This is known as combining absolute errors in quadrature: adding the squares, and then taking the square root.

\[
Z = XY \text{ or } Z = X/Y \\
\frac{\Delta Z}{Z} = \sqrt{\left(\frac{\Delta X}{X}\right)^2 + \left(\frac{\Delta Y}{Y}\right)^2}
\]

Note that we are combining errors in their fractional or relative form.

\[
140928
\]
Since she calculates the area $A$ as a product of the width $W$ and the height $H$, Claire applies the second formula as follows:

$$\frac{\Delta A}{A} = \sqrt{\left(\frac{\Delta W}{W}\right)^2 + \left(\frac{\Delta H}{H}\right)^2} \implies \Delta A = A\sqrt{\left(\frac{\Delta W}{W}\right)^2 + \left(\frac{\Delta H}{H}\right)^2}$$

$$\implies \Delta A = (602 \text{ cm}^2)\sqrt{\left(\frac{.25 \text{ cm}}{21.5 \text{ cm}}\right)^2 + \left(\frac{.25 \text{ cm}}{28.0 \text{ cm}}\right)^2} \approx 9 \text{ cm}^2$$

So Claire reports this result as $602 \pm 9 \text{ cm}^2$.

Example 3 shows that rigorous uncertainty propagation generally leads to a slightly narrower uncertainty interval than the intuitive method. Both methods are valid; however, a better estimate of the error arises from the second technique, using error propagation, and this is the strongly preferred method for you to aspire to.

One concept that may not be obvious from the above examples is that measurement uncertainties arise from the measurement method, and thus an experimenter can always determine the measurement uncertainty even before the measurement has been taken. In other words, the data, i.e. the measured values, do not determine the measurement uncertainty, and thus the measured values should never be used to retroactively determine the uncertainty in the measured values. However, retrospective analysis of the data can be used as a consistency check with propagated uncertainty, or to characterize imperfections in the experiment arising from other factors besides measurement uncertainty, as discussed in the next section.

**Statistics**

When performing an experiment to determine a physical quantity, it is often the case that many experiments (or repetitions of a single experiment) are performed, and there is some variation between the results from each experiment. This variation in the data can be due to measurement uncertainty, or it could be due to variation between the actual value of the physical quantity in each experiment. Once we have collected the set of experimentally determined values, we can do further arithmetical operations on these values to calculate numbers that help to characterize the data—these characterizing numbers are called statistics.

Suppose we have performed $N$ experiments to measure some quantity $h$, and the resulting value from each experiment is labeled as $h_1$, $h_2$, $h_3$, … , $h_N$. If we wanted to predict what the most probable value would be if we performed one more experiment, then a good unbiased guess would be the arithmetical average of all the $N$ measurements. This commonly used statistic, the arithmetic average $\overline{h}$ is obtained from all the individual measurements $h_i$ as follows:
Suppose the following experiment was being performed: NBA players are picked at random, each has their height measured using a ruler with uncertainty ±0.5 inches. Assume that after testing a sample of \( N \) players (but not all the players), we observe their average height to be 6’6”. Even though the uncertainty in this measurement is ±.5”, we observe some players who are nearly 7’ and others who are barely 6’. In this case, our data is spread out by intrinsic variation in the measured quantity between each experiment. However, we can still characterize the spread in the players’ heights using a statistic called the sample standard deviation, generically denoted with \( s \). The heights’ sample standard deviation \( s_h \) is given by:

\[
\bar{h} = \frac{1}{N} \left( h_1 + h_2 + h_3 + \ldots + h_N \right) = \frac{1}{N} \sum_{i=1}^{N} h_i
\]

(1)

The sample standard deviation will be a larger number if the spread in the data is large (e.g. if there are many NBA players who are much taller and many who are much shorter than average) and it will be smaller if the spread is small (e.g. if most NBA players are very close to the average height.) Since we are only testing a subset of the entire population of NBA players, it could turn out that the actual average height of players in the NBA is not exactly the same as the average from our sample of \( N \) players. In order to characterize how close the sample average is to the true average, we calculate a statistic called the standard error, generically denoted by the symbol \( SE \). The standard error of the average height is given by:

\[
s_h = \sqrt{\frac{1}{N-1} \sum_{i=1}^{N} (h_i - \overline{h})^2}
\]

(2)

A smaller standard error indicates that the sample’s average is probably close to the overall population’s average, while a larger standard error indicates that the sample’s average may differ significantly from the overall population’s average.

Characterization of experimental results using statistics can be very useful alongside uncertainty propagation. For example, if it is observed that the measurement uncertainty is very small compared to the standard deviation, then one may conclude that intrinsic variation exists in the population. On the other hand, if one believes that there is no variation in the actual values between experiments, then it is good to check that the standard deviation is accounted for by measurement uncertainty, and the standard error in the average value is accounted for by the propagated uncertainty in the mean.
**Error** is defined as the difference between the experimentally determined value and the expected/actual value. If a well-controlled experiment is performed carefully, the student should ideally be able to use the tools of uncertainty propagation to account for any observed error. However, there are often unforeseen or uncontrollable complications in experiments that lead to additional error that is larger than measurement uncertainty. Statistics are useful in characterizing these additional sources of error. We distinguish between two kinds of error: *random error* and *systematic error*. Random errors increase variation in the data but do not influence the data’s mean value. Systematic errors, on the other hand, can change the data’s mean.

**Example 4:**

Suppose Claire was attempting to precisely measure the direction of the Earth’s gravitational field in her vicinity. To do this, Claire hangs a plumb bob by a string and precisely measures the angle at which it hangs, waits a few moments, and then repeats the process several times. She performs this technique in three different nearby places:

Claire’s first experiment is performed in a closed room built to isolate it from outside vibrations.

Claire’s second experiment is performed in a laboratory she shares with other students who are stomping around causing vibrations and turbulent air flows in the room.

Claire’s third experiment is performed outside with a strong breeze continually blowing in the same direction.

Claire expects that because of vibrations and wind currents, her measured angles in the second experiment probably have more variation than in the first experiment—thus the sample standard deviation for the second experiment would be larger than the sample standard deviation for the first experiment. However, since these vibrations and wind currents probably caused the bob to swing in all directions roughly equally, then the average angle from the second experiment is approximately equal to the average angle from the first experiment. The second experiment’s average will tend to be a closer approximation to first experiment’s average as Claire increases the number of measurements performed in the second experiment. This is an example of random error—and the effects of random error on the average value are reduced when more trials are performed. In the third experiment, however, Claire expects the average angle to differ from other experiments’ average because of the wind consistently pushing the bob in one direction. This is an example of systematic error—performing more trials and averaging will not reduce this effect.

Example 4 shows that the effects of random error can often be reduced by making additional measurements and computing the average, but that systematic errors are independent of the number of measurements. If you see any evidence of such problems in your data, you should
mention it in your lab report and try to suggest specific causes. Quantitative investigation of error using the basic statistics outlined above can help to determine the causes, effects, and significance of various sources of error, and is warranted when error is observed to exceed measurement uncertainty.

The uncertainties and experimental errors discussed above do not include mistakes in reading or setting instruments, sometimes referred to as “human error”. The student’s job in the lab is to perform the experiment without introducing human error, and attributing error to human error indicates poor performance in the lab (and thus is not an acceptable way to account for error). Unavoidable sources of human error can be incorporated into the estimates of measurement uncertainty.

Although we have presented the mathematical rules and formulas for error analysis it is important to have an intuitive understanding of the concepts involved. In this section we explain the difference between accuracy and precision and give some examples of error analysis. In an experiment, there is a difference between accuracy and precision. **Accuracy** means how close a measurement is to the known (true) value. **Precision** means how close your data points are to each other. It is possible for data to be accurate and not precise and vice versa. For example:

![Diagram](image)

The average in (a) is accurate (all measurements average to the bulls eye) but not precise, whereas in (b) the average is precise but not very accurate.

Usually, if an experiment is accurate but not precise, this is due to random errors such as human error or imprecise equipment. If an experiment is precise but not very accurate, there is systematic error. A systematic error is an error with the equipment or an error in technique that is repeated consistently.

Is it possible to have a small percent error in a value and rotten precision? YES! This could be indicative of a large amount of random error (and a lot of good luck). Is it possible to have a small standard deviation and low accuracy? YES! This could be indicative of a systematic error.
For example if one were to measure the acceleration of gravity and measured the value to be $2.0 \pm 0.2 \text{m/s}^2$, this would be a very precise answer but it does not mean the measurement gave a good result. Although the percent error is low, the true value of $9.8 \text{ m/s}^2$ is 6 standard deviations away. The probability of using the experimental apparatus to measure the correct value of $9.8 \text{ m/s}^2$ is near 0. This is indicative of large systematic error. Likewise if one were to measure that the acceleration due to gravity is $2.7 \pm 0.7 \text{m/s}^2$, the percent error could be large but the true value for the acceleration due to gravity of $9.8 \text{ m/s}^2$ is within the standard deviation. This would be indicative of a lot of random error to make the standard deviation so large.