Significant Figures and an Introduction to the Normal Distribution

**Object:** To become familiar with the proper use of significant figures and to become acquainted with some rudiments of the theory of measurement.

**Apparatus:** Stopwatch, pendulum, table clamp, and meter stick.

**References:** Mathematical Preparation for General Physics, Chapter 2, by Marion & Davidson; Introduction to the Theory of Error by Yardley & Beers.

**Significant Figures:**

Laboratory work involves the recording of various kinds of measurements and combining them to obtain quantities that may be compared to a theoretical result. For such a comparison to be meaningful, the experimenter must have some idea of how accurate their measurements are, and should report this along with the data and conclusions.

One obvious limit to accuracy is the size of the smallest division of the measuring instrument. For example, with an ordinary meter stick, one can measure length to within 1 mm. Thus, the result of such a measurement would be given as, e.g., 0.675 m. This result contains three significant figures. It does not convey the same meaning as 0.6750 m, which implies that a device capable of measuring to 0.0001 m or 0.1 mm was used. If the object were less than 0.1 m long, a measurement might give a result of 0.075 m, which contains only 2 significant figures: 7 and 5. Zeros used to locate or hold a decimal place are not considered significant figures. For an even smaller object, a one significant figure result might be obtained, e.g. 0.006 m. It is important, in any kind of measurement, to judge all the factors affecting the accuracy of that measurement and to record the data using the appropriate number of significant figures.

When addition and subtraction quantities, it is possible to obtain a result with less (or more) significant figures than the original quantities had. For example, the following addition,

\[
\begin{align*}
0.721 \text{ m} \\
+ \quad 0.675 \text{ m} \\
\hline
1.396 \text{ m}
\end{align*}
\]

leads to a result with 4 significant figures. On the other hand, subtraction

\[
\begin{align*}
0.721 \text{ m} \\
- \quad 0.675 \text{ m} \\
\hline
0.046 \text{ m}
\end{align*}
\]

leads to a result where one significant figure has been lost. To add or subtract two quantities of different accuracy, the most accurate must be rounded off. For example, to add 12.3 and 1.57 m, one must first round off 1.57 m to 1.6 m and then add.

When multiplying or dividing quantities, the rule is that the result must contain the same number of significant figures as that of the original quantity with the least number of significant figures. For example, to compute the volume of a piece of sheet metal with dimensions 25.7 cm, 32.6 cm, 0.1 cm, the result is

\[
25.7 \text{ cm} \times 32.6 \text{ cm} \times 0.1 \text{ cm} = 8 \times 10^1 \text{ cm}^3
\]

Notice that the exponential form of expressing the result must be used because to write 80 \(cm^3\) would imply 2 significant figures.
Procedure:

Complete any exercises provided by your laboratory instructor.

Propagation of Uncertainties:

Numerical data obtained during the course of an experiment should be accompanied by an estimate of the uncertainty or error associated with that data point. In the simplest case where only one measurement is involved, the uncertainty should be taken as the smallest unit of the measuring device. For example, the result of a single length measurement with a meter stick might be

\[(0.323 \pm 0.001) \text{ m}\]

If repeated measurements of the same quantity are taken, the proper way to express an uncertainty is to quote what is called the “standard deviation of the mean”. The procedure for doing this (combine different measurements and their uncertainties) to arrive at a single final result is given below.

Suppose that in timing the duration of two phenomena, for whatever reason one cannot measure better than \(\pm\) a few seconds. For example, the result might be stated \((18 \pm 4) \text{ s}\). The question arises as how to properly express the uncertainty associated with the sum or difference of two such measurements, such as,

\[(18 \pm 4) \text{ s} + (23 \pm 3) \text{ s} = ?.\]

It would not seem accurate to add the uncertainties, getting \(\pm 7 \text{ s}\), or to subtract them, getting \(\pm 1 \text{ s}\), so the proper result must be somewhere in between. From a deeper investigation into the theory of measurement, one would find that the correct procedure is to combine the uncertainties as follows:

\[\sqrt{(4s)^2 + (3s)^2} = 5s,\]

So

\[(18\pm 4) \text{ s} + (23\pm 3) \text{ s} = (41\pm 5) \text{ s}.\]

In general, the uncertainty, \(A\), associated with the addition or subtraction of \(N\) quantities with uncertainties \(\pm a_i\) is given by,

\[A = \sqrt{\sum_{i=1}^{N} a_i^2}.\]

To find the combined uncertainty of a product or quotient, the procedure is to first express the individual uncertainties as fractional uncertainties, and then combine them as above, to get a combined fractional uncertainty. Note that fractional uncertainties are dimensionless. This means they carry no unit like \(m\) or \(s\). For example consider the calculation,

\[(15\pm 1) \text{ m} / (7.0\pm 0.4) \text{ s} = (2.1\pm ) \text{ m/s}\]
The fractional uncertainty in the length measurement is written as:

\[ S_l = \frac{1m}{15m} = 0.07 \]

While the fractional uncertainty in the time measurement is given as:

\[ S_t = \frac{0.4s}{15m7.0s} = 0.06 \]

Thus

\[ S = \sqrt{S_l^2 + S_t^2} = \sqrt{(0.07)^2 + (0.06)^2} = 0.09 \]

is the combined fractional uncertainty. Now we can calculate the uncertainty in the finally quantity by simply multiplying \( S \) by the value 2.1 m/s:

\[ 0.09 \times 2.1 \frac{m}{s} = 0.2 \frac{m}{s} \]

So, the final result is given by:

\[ \frac{(15\pm1)m}{(7.0\pm0.4)s} = (2.1 \pm 0.2) \frac{m}{s}. \]

To find the fractional uncertainty in a quantity raised to a power, the fractional uncertainty in the quantity is multiplied by the power to get the fractional uncertainty in the result. Thus, if

\[ s = aT^4 \]

and the fractional uncertainty in \( T \) is 0.1; there will be a fractional uncertainty in \( s \) of 0.4.

**Procedure:**

Complete any exercises provided by your laboratory instructor

**The Normal Distribution:**

In many physics experiments, a single measurement is not adequate, so the measurement is repeated many times to reduce any uncertainty. If the cause of the differences in these individual measurements is random, i.e. varies in an uncontrollable way, then the most probable value for the result is the mean value of all the measurements. (This is not necessarily the case for other distributions where the mean, median and mode do not coincide.)

Suppose that the drop of a ping pong ball from a certain height was being measured with a stop watch. As the experiment is repeated, tiny density fluctuations in air, or tiny breezes, may affect the drop time. Also judging exactly when the ball strikes the floor will vary with each trial. These and other uncontrollable (random) effects may make a single individual measurement either too large or too small for that particular drop, so it doesn’t seem unreasonable to take the mean value as the most probable result. It is assumed that no “systematic errors” present, such as the stopwatch running consistently too slow or too fast.
But how much uncertainty should be attached to the mean value? The theory of measurement asserts that for a large number of repeated measurements with random uncertainties, the measurements form a “normal distribution”, which is illustrated in Fig. 1. For example if one were to make 80 measurements of a length. The value of a measurement is plotted against the number of times it occurs in the sample of 80. The smooth curve in the Figure 1 is the limiting shape of the histogram for a very large number of measurements.

Knowing the ideal distribution curve for a series of measurements, it is possible to define the uncertainty as follows. The amount of uncertainty will be such that the probability of any new measurement falling in the interval from the mean value minus the uncertainty to the mean value plus the uncertainty will be about 67%. This definition of uncertainty is called the “standard deviation of the mean” and is denoted by \( \sigma \), and is shown in Fig. 1.

So with \( N \) measurements the mean, \( \bar{x} \), is found by

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

and to a good approximation the standard deviation is given by,

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

**Pre-lab Questions:** The answers to the following questions are to be included in your pre-lab.

1. How many significant figures are in the following numbers?
   a. 0.00012
   b. 1,000.00
   c. 3.14 \times 10^{23}
2. Using a stopwatch, you measure the period of a grandfather clock. Your results are 0.6, 0.6 and 0.7 sec.
   a. What is the mean value for the period?
   b. What is the standard deviation?
   c. What is your best estimate of the error?
   d. What is your percent error?
   e. What is your fractional error?
   f. Write down your measurement for the period, and express the error in terms of ±.

Procedure:

1. Turn on your computer and open up Capstone. Create a two-column data table by clicking on the Table icon on the right hand tool bar.

2. At the head of each column click on Select Measurement. From this pull down menu first select Create New then select User-Enter Data. Label the first column “Trial Number” and the second column “Period”, do not forget to label the units in seconds.

3. Lab partners should take turns timing the period (T) of the pendulum, and recording the data points, until a total of 40 data points are taken. Rather than attempt to time a single oscillation, time several full oscillations and determine the average period. For example, measure the time it takes to complete five oscillations and divide by five, recording this average period as a single data point in your data table. Data should be recorded to 0.01 s. For the best results try to keep the amplitude of the pendulum swing consistent all the measurements, even if restarting is necessary.

5. To create a histogram of your data, select or highlight you data by clicking and dragging the mouse over the data table. With the data highlighted click on the Histogram icon, found on the right hand tool bar. To create a histogram of your data, similar to Figure 1 click on the Select Measurement tab found on each axis. Use the x-axis to display “period” and the y-axis to display “Trial Number”.
   In order to create a meaningful histogram one needs to “bin” the data. To bin the data, the range of measured time values is divided into at least 10 intervals or “bins”. The number of bins is chosen in order to get reasonable number of data points (n) into each bin or interval. Capstone will do this automatically for you. The number of bins can then be adjusted by clicking on either the + or – bin icon found on the upper tool bar of the histogram, or zooming in or out.

6. Using your 40 collected data points, determine the mean value of the period (T), and the standard deviation of the mean, using the Capstone statistics tools. To analyze your data set, click on the Sigma icon found either on the histogram tool bar or on the Data Table tool bar. Record the Maximum, Minimum, Mean and Standard Deviation in the table below. Make sure to record your answers using the proper number of significant digits, and proper units.

<table>
<thead>
<tr>
<th>Maximum</th>
<th>Minimum</th>
<th>Mean</th>
<th>Standard Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Observe how well your measurements approximate the ideal distribution.

7. The theoretical period (T) of a pendulum can be determined by measuring the length (L) of the pendulum and using the formula.
\[ T = 2\pi \sqrt{\frac{l}{g}} \]

Measure and record the length of the pendulum using a meter stick make sure to use the proper number of significant digits and units. Calculate the record the theoretical period of the pendulum; once again make sure to use the proper number of significant digits.

<table>
<thead>
<tr>
<th>Length of Pendulum</th>
<th>Theoretical Period of Pendulum</th>
</tr>
</thead>
</table>

**Questions:**

1. If you had taken only one measurement for the period of the pendulum, with the same apparatus, what would be the uncertainty of that measurement?

2. In reality, you took 40 measurements. That said, what is the uncertainty in any one of your individual measurements… what value would you give as your uncertainty?

3. What is your best estimate as to the actual period of the pendulum. Give your answer using the proper number of significant digits and properly quoting the uncertainty in your answer.

4. Write in and compare the experimentally determined period of the pendulum to the theoretically determined period of your pendulum. Are they in agreement with one another? Do your measurements support the theoretical equation for the period of the pendulum?

<table>
<thead>
<tr>
<th>Measured Period of Pendulum</th>
<th>Theoretical Period of Pendulum</th>
</tr>
</thead>
</table>

4. In step 1 of the procedures it was suggested that you time several oscillations and determine the average period of the oscillation, rather than attempt to measure the period of a single oscillation, why?

   a. What are the advantages of this method?
   b. How would the results be affected if a large number of periods were timed?

5. Compare your histogram to the normal distribution curve. Does this evidence support the theory that your data is normally distributed? Why or why not?