**Measurement and Uncertainty**

The process of doing an experiment in physics involves collecting and understanding data. Data are obtained through measurements with various instruments. With every measurement comes some uncertainty.

There are two general categories of uncertainty: systematic and statistical. ***Systematic uncertainties*** cause measurements to be consistently too large or too small. These can result from various measurement and physical influences such as instrument calibration or variables we inappropriately assumed were negligible. If all systematic uncertainties are identified and accounted for, we can eliminate systematic uncertainty, however, this is usually impractical and, possibly, impossible. ***Statistical uncertainties*** are variations in measurements that occur with any measurement. These uncertainties may arise from limitations in measuring devices or fluctuations in the physical system. We can reduce, but never eliminate, statistical uncertainties from our experiments. All measurements have uncertainty, and they may have both statistical and systematic uncertainties from a variety of sources.

Measurements are often also described as being either ***precise*** or ***accurate***. Precision and accuracy are independent of one another. If you're throwing a ball at a target, you would be very accurate if the ball hit around the target each time. If you hit the target each time, then you would be precise and accurate. However, if you were aiming for the target and hit a nearby tree at almost exactly the same spot each time, you would be precise but not accurate. Measurements that are very ***precise*** are ones with small uncertainty***.*** Measurements that are close to a predicted or standard value are referred to as ***accurate measurements***. It is necessary to understand how precise (or uncertain) our measurements are before making any claims about our accuracy. It's not helpful to average around the target if our throws are all over the place.

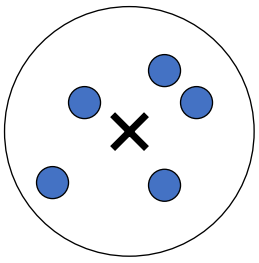


Fig. 1 Accurate

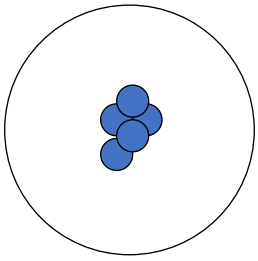


Fig. 2 Accurate & Precise

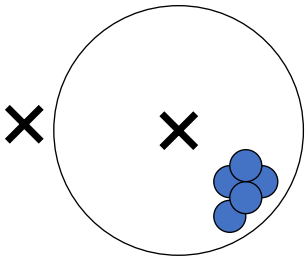


Fig. 3 Precise

In high school science classes you may have encountered ***significant figures*** (often called "sig figs"). For each measurement, the final digit communicates to a reader about the precision of the measurement. Instrumental precision and significant figures are related because the final significant figure is an estimate based on the instrument's precision. Often other statistical uncertainties are much greater than the instrumental precision, so it is important to consider the greatest contributing factor to each measurement's uncertainty.

**Statistics for Repeated Measurements with Statistical Variation**

In some situations you will be able to repeat a measurement many times in order to assess the uncertainty caused by statistical sources of variation. If you take sufficient measurements, you can draw a histogram that displays how frequently the measurement lands in different ranges of values. Three things can be seen in such a histogram. First, the data is often clustered in some way and mathematically we estimate the value they are clustered around by calculating the mean. If $x_i$ are the measured values and if you took $N$ different measurements ($i = 1, \ldots N$), the mean is calculated using

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

Note that the mean $\bar{x}$ is still an estimate based on a finite number of measurements and if you kept taking more measurements the mean would change somewhat.

The second thing to notice in the histogram is that the values are distributed on either side of the mean value, with some characteristic width. The width is usually expressed in terms of the standard deviation $\sigma$

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

About 68% of the measurements will lie within the range from $\bar{x} - \sigma$ to $\bar{x} + \sigma$. However, if you have gone through the trouble of taking $N$ different measurements, the uncertainty in the mean value $\bar{x}$ is quite a bit better than the standard deviation. The uncertainty in the mean of all your measurements is called the standard uncertainty $\delta\bar{x}$ given by

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

This acts to reward you for taking more data, so that the uncertainty in the collection of many measurements (the mean) is smaller than the uncertainty in any single measurement. If you do one hundred times as many measurements, the uncertainty only decreases by a factor of ten – hard work!

**Making Comparisons (Mini-lecture: youtu.be/J8E9jgsB8Zs)**

The quantity that we use to make comparisons between numbers is known as $\boldsymbol{t}'$. If we have two numbers with uncertainties, $A \pm \delta A$ and $B \pm \delta B$. Then the $\boldsymbol{t}'$ value is

$$t' = \frac{A - B}{\sqrt{(\delta A)^2 + (\delta B)^2}}$$

After calculating $t'$ using two measurements, you can evaluate their similarity (or distinguishability) through the following interpretation:

- $|t'| \leq 1$: $A$ and $B$ are indistinguishable compared to the uncertainty (very likely the same). Remember though, if you do a better measurement to decrease the uncertainties, you might later uncover a difference between $A$ and $B$. That is, poor precision may be hiding a subtle difference!

- $1 < |t'| \leq 3$: It is possible that $A$ and $B$ are the same, but they are not clearly indistinguishable.

- $|t'| > 3$: It is very unlikely that $A$ and $B$ are the same.

For $|t'| \leq 1$, you may need to consider:
- Improving your measurements
- Decreasing uncertainties
- Checking that you've appropriately accounted for uncertainty

For $|t'| > 3$, you may need to consider:
- Retaking or improving your measurements
- Evaluating your model and possibly revising it

**Fitting by the Method of Weighted Least Squares (Mini-lecture: youtu.be/UX4Jm33TM3E)**

When fitting any function to a set of data, one tries to find a model with parameters that can minimize the 'distance' between the data points and the function.

Consider a function $f(x)$ and a set of measurements consisting of $x_i$ and $y_i$ values and an uncertainty $\delta y_i$ in each measured $y_i$ . The 'distance' between a data point and the function is given by the residual $y_i - f(x_i)$ .

A measure of how close the function is to the data overall is given by the weighted $\chi^2$

$$\chi^2 = \frac{1}{N}\sum_{i=1}^{N} \frac{\left(y_i - f(x_i)\right)^2}{(\delta y_i)^2}$$

The method of least squares means finding a function with parameters that minimize $\chi^2$. For instance, if you are fitting a straight line with the form $f(x) = mx + b$ , you would vary the

parameters $m$ and $b$ until you find the minimum $\chi^2$ . The fit can be done iteratively, by trial and error (or, for one or two parameter models, analytically minimized).

A benchmark for a good fit would be a $\chi^2$ value close to one. In that case, the scatter in the data, on average, would be similar to the estimated uncertainties. A $\chi^2$ value much greater than one suggests either a bad fit or underestimated uncertainties. A $\chi^2$ value much less than one probably means that you have over-estimated your uncertainty.

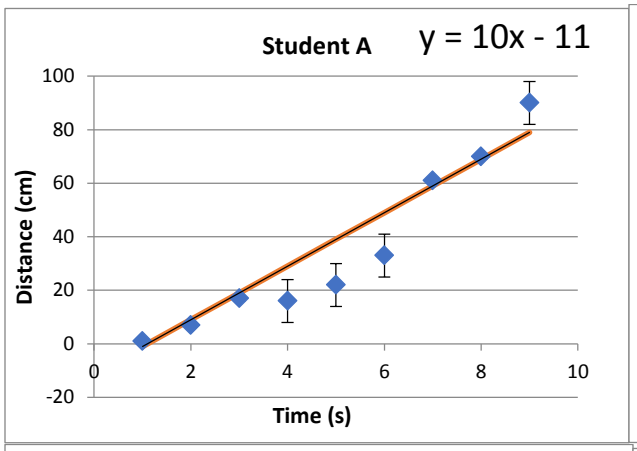
For a large $\chi^2$ value, you may need to consider:

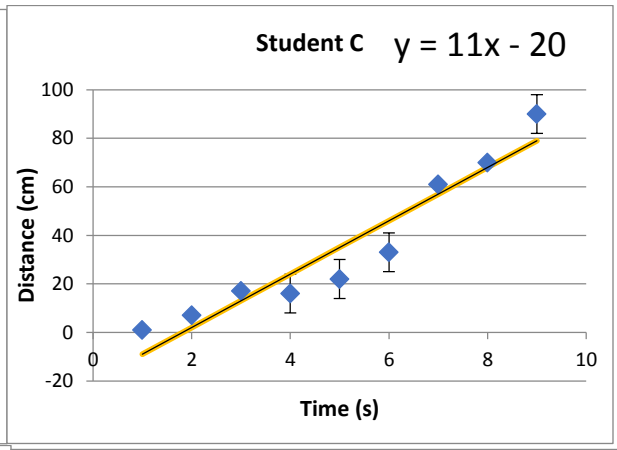
- Retaking or improving your measurements.
- Checking that your uncertainty makes sense.
- Evaluating your model and possibly revising it.

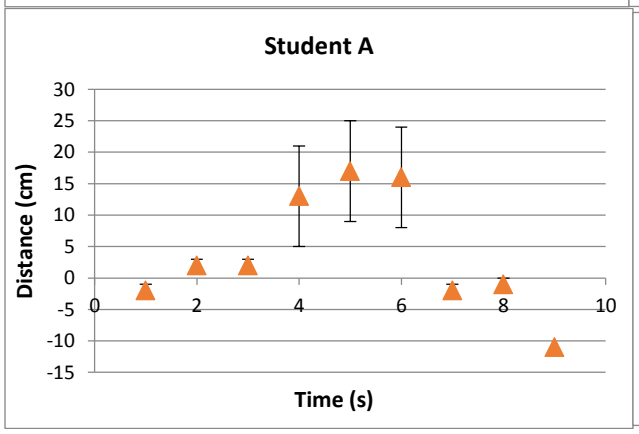
For a small $\chi^2$ value, you may need to consider:

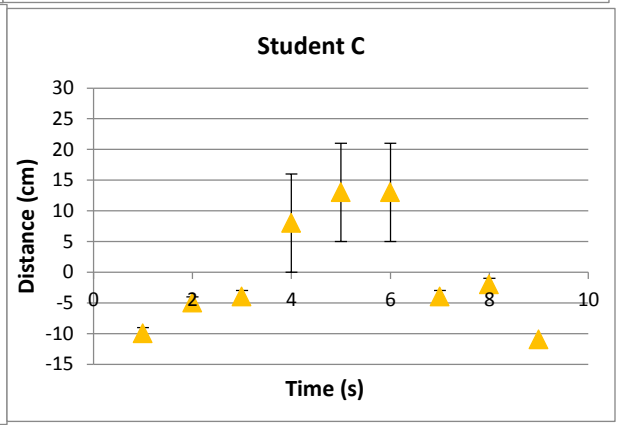
- Checking that you've appropriately accounted for uncertainty

Note that $\chi^2$ uses the uncertainty in each measured value. If you search online, you may notice that many different $\chi^2$ exist, including those that don't take uncertainties into account. During the Terminal Velocity lab, you compared fits that four students made for a set of data (see below).



Student A recognized that more precise measurements, as indicated by uncertainty, should matter more in the fitting process. This is a *weighted* fit using the method of *weighted least squares linear regression* to calculate $\chi^2$. Because you are making measurements and finding evidence for which model appropriately describes your data, the better (more precise) measurements "count" more when fitting your data.

Student B decided that the precision of the measurement does not matter and produced an *unweighted* fit using the method of *unweighted least squares regression*. While this may look visually appealing, Student B has not considered that several measurements have much better precision than the four measurements with larger uncertainties. If you think back to Pendulum for Pros, obtaining more precise measurements allows you to uncover the small angle approximation, so the precision of measurements should matter when determining which model describes your data.

| | **Weighted $\chi^2$** | Unweighted $\chi^2$ |
|---|---|---|
| Student A | **3.3** | 94.7 |
| Student C | **18.8** | 76 |

**Propagating Uncertainty through Multiplication and Power Calculations**

If a quantity that you're calculating has more than one source of uncertainty (i.e., you're multiplying two measurements together), then there is uncertainty in the final quantity. For example, you have recorded two measurements, $A \pm \delta A$ and $B \pm \delta B$, and are determining $R = AB$ with uncertainty $\delta R$.

$$\delta R = |R|\sqrt{\left(\frac{\delta A}{A}\right)^2 + \left(\frac{\delta B}{B}\right)^2} = |AB|\sqrt{\left(\frac{\delta A}{A}\right)^2 + \left(\frac{\delta B}{B}\right)^2}$$

This works for multiplication and division of any number of quantities that have uncertainty. For example, if $R = \frac{AB}{C}$ then $\delta R = \left|\frac{AB}{C}\right|\sqrt{\left(\frac{\delta A}{A}\right)^2 + \left(\frac{\delta B}{B}\right)^2 + \left(\frac{\delta C}{C}\right)^2}$.

If you're raising a quantity with uncertainty, $X \pm \delta X$, to a power, $R = X^n$, then the uncertainty is $\delta R = |n|\frac{\delta X}{|X|}|R|$. Complete propagation of uncertainty through powers before completing multiplication for calculations such as kinetic energy.