ASTR 121 – Spring 2016

Uncertainties and Error Analysis

What is Uncertainty?

Whether in science class or a real laboratory, a crucial part of understanding experimental results is analyzing uncertainty, or error. In scientific context, 'error' does not refer to a mistake made during an experiment. Rather, it describes the unavoidable imprecision in measurements. Though there are exceptions (for example, counting the number of people in a room), most measurements will have some amount of uncertainty, and quantifying this uncertainty is a necessary part of interpreting experimental results.

In the context of this lab, we will strive to quantify experimental uncertainties whenever possible. Additionally, we will focus on identifying and understanding sources of error, as well as trying to minimize them.

Types of Uncertainty

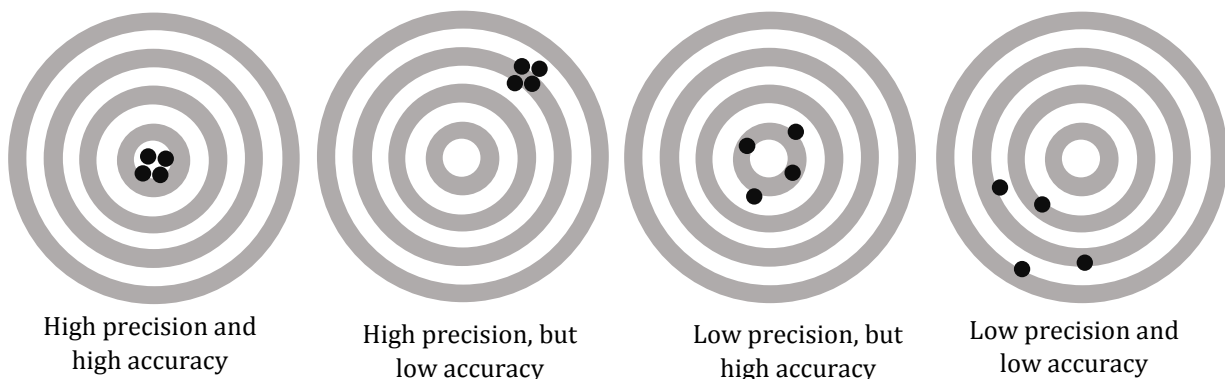There are two distinct types of error: random and systematic.

Random errors (sometimes called statistical errors) are caused by unpredictable fluctuations in an experimental setup, either from the environment or the instrumentation itself. Some common random errors encountered in astronomy are the background noise signal in CCD electronics or the resolution of a spectrometer. Because random errors often have a normal Gaussian distribution, they can be minimized by using a larger data set.

Systematic errors, on the other hand are caused by inherent flaws in an experimental setup that, unlike random errors, will affect all measurements in a similar way. This effect could either be caused by bad calibration or some scale factor. In astronomy, some sources of systematic errors could be cosmological redshift or the miscalibration of a photometer. Because systematic errors cannot be corrected simply by taking more data, scientists must either correct their experiment for the source of error, or observe enough of a trend to quantify the error.

Accuracy versus Precision

Though they're often used interchangeably in everyday conversation, there is an important distinction between accuracy and precision in science. Simply put, the accuracy of a measurement is how close it is to the actual value, while precision is how close multiple measurements are to each other.

Consider, for example, the common analogy of throwing darts. You can hit the bull's eye with high accuracy (hitting close to the target), high precision (hitting close to the same spot repeatedly), neither, or both. All four examples are illustrated in the figure below. In science, as in darts, we strive for both.

| High precision and high accuracy | High precision, but low accuracy | Low precision, but high accuracy | Low precision and low accuracy |

A useful way of thinking of this distinction is noticing that the precision of a measurement is limited by the random errors, while the accuracy of a measurement is limited by the systematic errors.

Intrinsic Uncertainties

One common source of random error is the intrinsic precision of a measuring device. No matter how sophisticated an instrument, there is some uncertainty associated with its ability to resolve small measurement differences. For example, if a ruler's smallest markings are 1 millimeter apart, the ruler cannot distinguish more than about 0.5 millimeters. A common standard is to assume that the intrinsic precision of an instrument is half of the smallest increment the device reports. So if a scale reads 4.36 kilograms, a conservative estimate of the intrinsic uncertainty of the scale is 0.005 kilograms. The measurement in this case would be written (4.360±0.005) kg.

Uncertainty Propagation Theory

Imagine you are conducting an experiment, make some measurements x and y, and find the uncertainty of each. However, you are not interested in x or y; rather, you are concerned with a value z which is the product of the two. How do you quantify the uncertainty in z?

For situations like this, propagation of uncertainty is an invaluable tool. Propagation of uncertainty (or propagation of error) is the combined effect of the uncertainty of measurements on the uncertainty of a function of those measurements. It is used to find the uncertainty of a calculated value where the values in the calculation have their own errors.

In general, the equation for propagation of uncertainty of a function σ_f is:

$$\sigma_f = \sqrt{\sum_{i=1}^{n} \left(\frac{\partial f}{\partial x_i}\right)^2 \sigma_i^2}$$

where f is a function of a set of variables x_i with uncertainties σ_i. In this example, we would report the quantity f as $(f \pm \sigma_f)$.

As this lab does not require knowledge of partial derivatives, you will not be responsible for deriving error propagation equations (but if you are familiar with them, feel free to practice!). Below is a list of some common functions and their propagated uncertainty equations that may be used in this class.

For each function f, A and B have uncertainties σ_A and σ_B, respectively, and a and b are exactly known non-zero constants (they have no uncertainty).

Function	Uncertainty
$f = aA$	$\sigma_f = \|a\sigma_A\|$
$f = aA \pm bB$	$\sigma_f = \sqrt{a^2\sigma_A{}^2 + b^2\sigma_B{}^2}$
$f = AB, f = \dfrac{A}{B}$	$\sigma_f = f\sqrt{\left(\dfrac{\sigma_A}{A}\right)^2 + \left(\dfrac{\sigma_B}{B}\right)^2}$
$f = aA^b$	$\sigma_f = \left\|\dfrac{fb\sigma_A}{A}\right\|$
$f = a\ln(bA)$	$\sigma_f = \dfrac{a\sigma_A}{A}$
$f = ae^{bA}$	$\sigma_f = \|fb\sigma_A\|$

As an example, consider the distance to a star using parallax. You measure the parallax angle to be $p = (0.7471 \pm 0.0012)$ arcseconds. The calculation of the distance is simple enough:

$$d = \frac{1}{p} = 1.338509 \text{ pc}$$

Now to determine the uncertainty in this calculated value of *d,* you propagate the uncertainty in *p,* either using the general formula or the given equation for division:

$$\sigma_d = d \sqrt{\left(\frac{\sigma_P}{P}\right)^2} = 0.00215$$

Now you have calculated the distance as well as the uncertainty in that calculation. Using the convention described in the end of this handout, our calculated distance is reported as:

$$d = (1.338 \pm 0.002) \text{ pc}$$

Though we will often use uncertainty propagation, it is important to realize the assumptions that it requires. Uncertainty propagation assumes standard Gaussian uncertainties, that is, uncertainties that are symmetrical Gaussian functions. Though this is thankfully the case much of the time, there are exceptions, in which other methods of quantifying uncertainty must be used (e.g., using Poisson statistics with a photometer). Additionally, this form of uncertainty propagation assumes that the variables are uncorrelated. In the scope of this lab, both of these assumptions will usually be valid.

Descriptive Statistics

One way of describing a set of data is by using measures of central tendency. Though there are many different measures of central tendency, the most common are arithmetic mean, median, and mode.

Arithmetic mean is the sum of a set of values, divided by the number of values. Formally, the mean μ of a set of N measurements of x_i is written as:
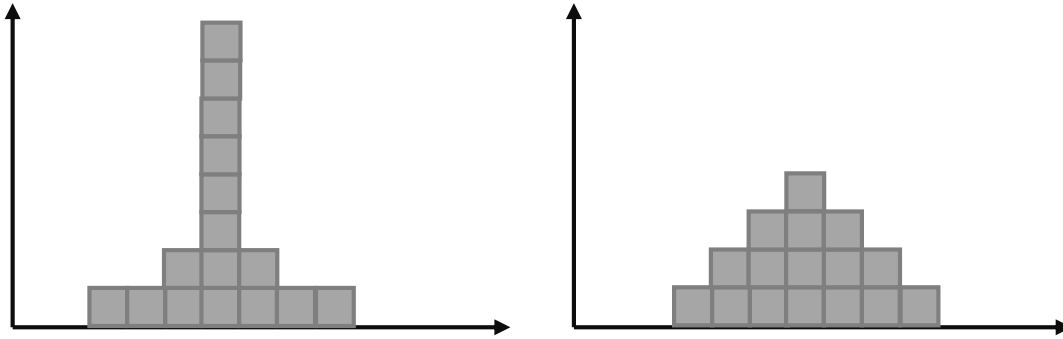
$$\mu = \frac{1}{N} \sum_{i=1}^{N} x_i$$

The mean is undoubtedly the most common measure of central tendency, though it does have its disadvantages, most significantly, how easily it is affected by outliers.

The **median** is a measure of central tendency that isn't as susceptible to influence by outliers. If a set of values is arranged in order, the middle value (or mean of two middle values, for a set with an even number of values) is the median.

Finally, the **mode** is the most frequent value in a set. Most useful for discrete sets of data, the mode is simply the value that is measured most often.

Though useful, measures of central tendency alone can't fully describe a set of data. Consider the two plots below: they both have the same number of data points, range of measurements, mean, median and mode. They are, however, clearly different. The plot on the left has data many points at a single value, while the plot on the right has more evenly spread-out points.

This is where we need to use measures of dispersion. Measures of dispersion describe how scattered a set of data is. The two most common are variance and standard deviation.

The **variance** is the mean of the squares of each easements deviation from the average, formally written as:

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

The **standard deviation** is simply the square root of the variance:

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

The standard deviation basically describes how far a typical measurement is from the average value. Additionally, it can be used to find the uncertainty associated with the mean. For a set of N data points with a standard deviation σ, the uncertainty of the mean is:

$$\sigma_\mu = \frac{\sigma}{\sqrt{N}}$$

Significant Figures

Besides correctly calculating and attributing an uncertainty to a given value, it is important to report the value correctly, as well. Imagine a student measures the diameter of a circle with a ruler and finds it to be 3.7 cm. They then use the diameter to calculate the circumference:

$$C = \pi d = 11.6238928216 \text{ cm}$$

Does it make sense that a measurement accurate to one decimal place could yield a calculation accurate to ten? For this reason, we consider something called **significant figures**. A significant figure (or 'sig fig') is digit of a number that contributes meaning to the measurement. All digits are significant except for leading zeros (any zeros that come before the first nonzero digit) and trailing zeros (zeros that come after the last nonzero digit) in a number not containing a decimal. Consider the following examples:

7.41 - Three sig figs	7041 - Four sig figs	7.4100 - Five sig figs
74100 - Three sig figs	704100 - Four sig figs	7.4156 - Five sigs figs

In this lab, the convention for reporting values will be to round the uncertainty to one significant figure and keep the same number of decimal places in the corresponding value. Some examples:

Raw value: $x = (7.314 \pm 0.2483)$ cm   Reported value: $x = (7.3 \pm 0.2)$ cm

Raw value: $x = (93.733 \pm 11.32)$ pc   Reported value: $x = (90 \pm 10)$ pc

The same rule applies to values in which you use scientific notation, though it's a can be a bit trickier:

$$\text{Raw value: } x = (4.448 \times 10^9 \pm 3.18 \times 10^8) \text{ yr}$$

$$\text{Reported value: } x = (4.4 \pm 0.3) \times 10^9 \text{ yr}$$