Physics 290 Lab 4
Error Analysis – Height of my Office

An error is the more dangerous in proportion to the degree of truth which it contains.
Henri Frederic Amiel

The only thing that makes life possible is permanent, intolerable uncertainty; not knowing what comes next. ***Ursula K. Le Guin***

Happiness makes up in height for what it lacks in length. ***Robert Frost***

Introduction

Measurements in physics are limited by uncertainties, colloquially called "errors". In this lab we will explore in a bit more detail the nature of statistical and systematic errors and how to properly deal with them when making measurements. For some background information on error propagation, the error primer handout from last quarter is probably useful. Ask your TA for a copy if you do not already have one.

In this lab we will make a measurement, the height of my office off the atrium floor, which will be limited by the accuracy of your equipment in addition to other systematic errors. The goal is to work through the error analysis for a complete measurement, and then explore a bit what these errors really mean.

Laboratory 4 Objectives

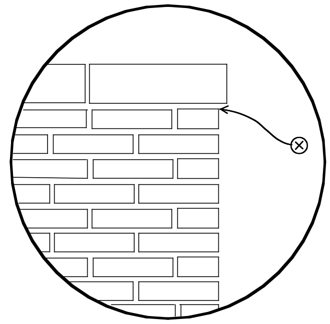
- Measure the height of the atrium by triangulation.
- Analyze errors in the measurement and evaluate the quality of the result.
- Explore the characteristics of a statistically-distributed sample of values.
- Apply these same statistical tools to the class data for the atrium height.

Recommended Schedule

You should be able to take all of the atrium data (Task 1) during the first week. The analysis of this data (Task 2) and the coin flipping analysis (Task 3) should be completed by the end of the second week, and the analysis of the class atrium height data (Task 4) can be finished during the final week.

IMPORTANT NOTE: In order to complete this lab properly, the results from the second task must be completed by everyone and shared before the final week! You should email your height results to me before January 30th. Please do not leave this to the last minute!

Task 1: Height of the Atrium

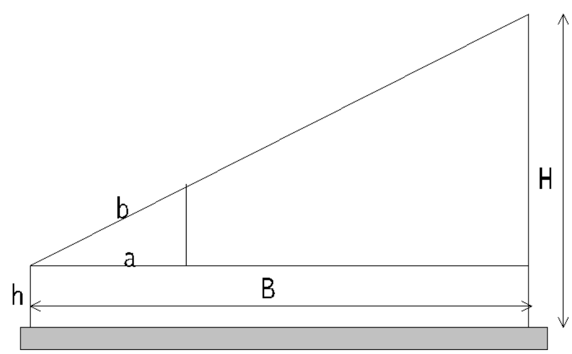


The task at hand is to measure the height of a spot on the interior wall of the atrium in Willamette Hall just under Prof. Torrence's office. See the figures to the right. What you see is the west wall just inside the 13th Avenue entrance. Your assignment is to make the best possible measurement of the height of the spot marked ⊗ (the top corner of the highest red brick in the column which stretches up to the 4th floor) above the floor of the atrium.

You will use the triangulation method sketched in the figure below. Tools at your disposal are: (a) a tape measure to determine the base distance ***B***, and the apparatus height ***h***, (b) a pair of adjustable wooden scissors, (c) a level, (d) lab stands, (e) a short string to construct a plumb line. These tools will either be in room 17 or in the undergraduate reading room in the atrium.

The triangulation procedure is as follows:

- Set up the lab stand containing the wooden scissors. Using the level make sure that the fixed wooden leg is horizontal. Drop the plumb line from the pivot of the scissors, and make a mark on the floor. (Use tape, don't scratch the floor…)
- Practice sighting the target using the free leg of the scissors. Determine the ratio b/a for a similar triangle constructed on the scissors. A measurement set requires moving the previously set free arm and resetting it. **You must not move the bases of the lab stand while you are resetting.** Practice this repeatedly, taking turns. When you are confident, take the data as follows.



- Each student should take at least three sightings. A sighting consists of the student using the free leg of the scissors to line up the target, and determining ***a*** and ***b*** using a plumb line. The plumb line should be lined up to a given value of ***b*** (each student's choice) and the corresponding ***a*** determined. Each student does this three times without disturbing the apparatus base. Each of these three sets of readings should be shouted out (yeah!) in order that partners can record them in their lab notebooks as well. The final step is to drop a plumb line from the intersection point of two legs of the scissors to the floor, and check for the mark. Each student should take a turn at these sighting tasks (3 separate settings of the moveable leg and checking the plumb line.) The group will therefore obtain a total of 6 or 9 sets of measurements of ***a*** and ***b*** (depending upon whether it is a group of 2 or 3 students). **Note:** If, at any time during these sightings, the plumb line does not coincide with the original mark on the floor, ***all sightings must be restarted***.

- Measure the distance from the floor to the pivot of the scissors, again 3 times, one measurement per student. Move the lab stand out of the way. Make the base line distance measurement from the mark on the floor to the point below the ⊗ mark. (How can you determine you are vertically below the ⊗ mark?) Each student should take the lead getting help from partners. Make at least one measurement per student, giving the group 3 measurements.
- A consideration of systematic errors should be made at this time. (a) How good is the level? Here is one way to find out. Set up a meter stick on its edge with the level along the edge. Insert enough sheets of paper under one end to cause the bubble to move. This should be the sensitivity of the level. Record these measurements. What happens if you flip the level over. Does it still read level? If not, there is probably an offset to the sight glass. This must be taken into account as a possible uncertainty. (b) How level is the floor of the atrium? (c) How perpendicular are the walls of the building to the floor? These are only some of the possible uncertainties. Try to estimate numerical values for each of these and explain how you arrived at these estimates.

Task 2: Analysis

- Calculate the ***mean*** and the ***average deviation*** from the mean of each set of readings of ***a*** and ***b***. Are the deviations consistent with your expectations? Write down your thoughts in the lab notebook.

- Calculate the height of the mark ⊗. Calculate the mean of three datasets (each belonging to a student; i.e., three measurements of ***a*** and ***b***, and the group measurement of ***h*** and ***B***). Calculate the average deviation from the mean.

- Carry out the error propagation analysis. Using the average deviations in step 1 as source errors, propagate the errors as you calculate the height in step 2. The propagated error should be "consistent" with the deviation observed in step 2. Include any significant systematic errors you measured in the final result.

- Determine which measurement contains the largest source of error. Discuss how you might reduce this and other sources of error.

Discussion

- Discuss possible systematic errors. Estimate or measure their relative importance.
- Discuss further improvements needed to obtain more certain values than what has been obtained in your measurements. What would ultimately limit the accuracy of this technique?

Task 3: Understanding Random Errors

All measurements contain some amount of random (or statistical) uncertainty. While we often cannot predict the nature of these uncertainties, we can characterize the distribution of values by taking repeated measurements. While the true distribution may well never be known, Physicists will often make the assumption that their data follows a **normal** (also known as **Gaussian**) distribution. In this task, we will explore how standard statistical measures like the **mean** and **standard deviation** are related to a Gaussian distribution and how they are used to characterize a specific random process: coin flipping.

A quick primer on statistics

If you have a perfectly symmetric coin, there is a 50% chance that the coin will land as heads when flipped. If you flip the coin 10 times, it is not surprising that on average the coin will come up heads 5 times out of 10. If you want to know how often you will get three heads or less, however, you need to know something about how probable each outcome is. This is known as the expected frequency distribution. For a binary system with two exclusive results (like a coin flip) this can be computed exactly and follows what is called the Binomial distribution. The probability of getting a certain result ***r*** times in ***N*** trials is given by

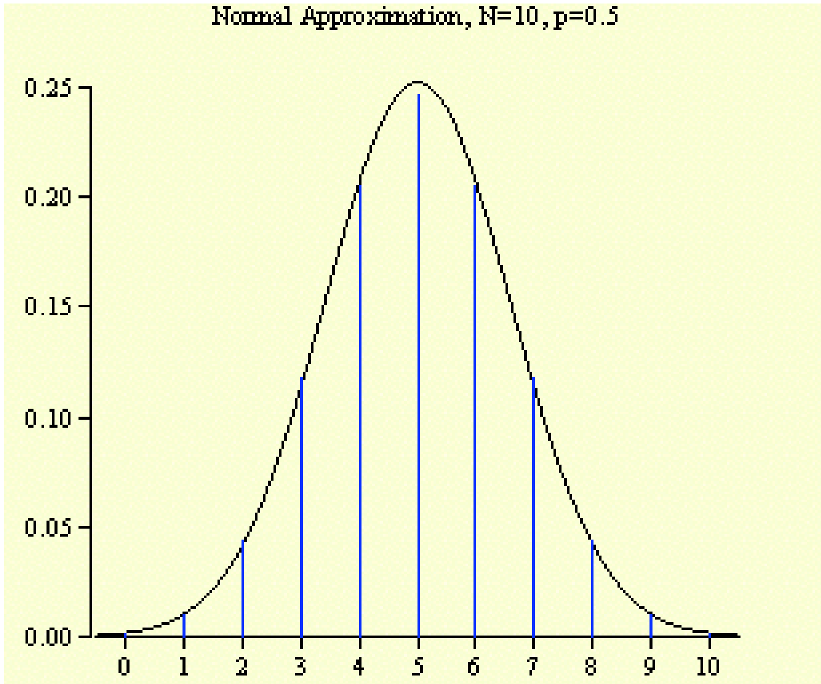
$$P(r) = \frac{N!}{r!(N-r)!}p^r(1-p)^{N-r}$$

where ***p*** is the probability of the result on any given trial (50% for heads with a perfect coin). As long as the number of trials N is large enough, this exact distribution is well estimated by the analytic Gaussian distribution

$$P(x) = \frac{1}{\sqrt{2\pi\sigma}}e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

The factor in front is only needed to normalize the distribution such that the indefinite integral of ***P(x)*** equals 1. The part in the exponent controls the shape of the distribution, and only depends on two parameters: the mean μ and the width σ.

The difference between the exact binomial distribution (vertical lines) and the Gaussian approximation (smooth curve) for a coin flipped 10 times is shown below. As you can see, the two distributions are almost identical.



Sample Estimation

If you have a finite number of measurements, you cannot determine exactly the **parent distribution** (i.e.: the truth) but you can make an **estimation** from the data. This is what we do when we make measurements. We are trying to make an estimate of the true value. To describe a Gaussian distribution you only need two values: the mean and width. The distribution mean μ is approximated by the sample mean $\overline{x} = \frac{1}{N}\sum x_i$ which is also simply the average of the values measured. The distribution width σ, meanwhile, is approximated by the ***standard deviation*** sometimes written as σ_x and given by $\sigma_x = \sqrt{\frac{1}{N}\sum(x_i - \overline{x})^2}$. This is similar to the average deviation, but is actually more like the quadrature sum of the deviation from the mean. For obscure theoretical reasons, it is actually better to replace N with (N-1) in the denominator of the standard deviation formula, but we will ignore that fact for now. As long as N is not small, it doesn't really matter.

With a little algebra, it can be shown that the standard deviation can be computed from the mean of x^2 and the mean of x as follows:

$$\sigma_x^2 = \frac{1}{N}\sum x_i^2 - (\frac{1}{N}\sum x_i)^2$$

Note that this last expression is actually the square of the standard deviation.

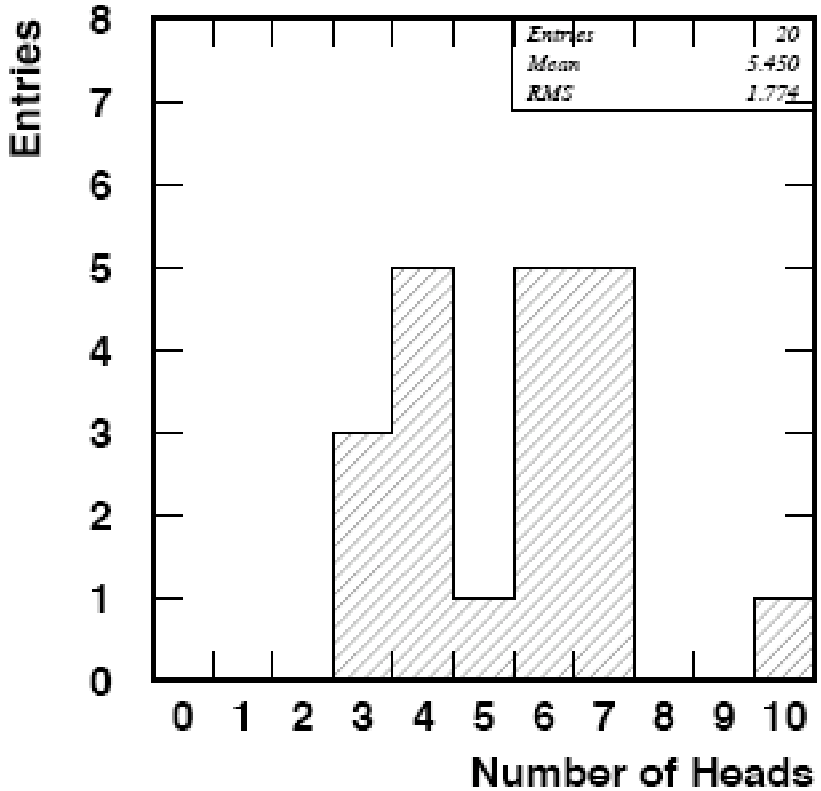
The standard deviation is a useful quantity because it provides a measure of the uncertainty expected in any individual measurement. In fact, in Physics it is customary to quote the standard deviation as the measurement error assuming a Gaussian hypothesis. If our data is truly distributed according to a Gaussian distribution, we can integrate over the range **[$\mu-\sigma$, $\mu+\sigma$]** and we find that 68% of the time an individual measurement should fall in this range. This is the "one sigma" probability, and errors in Physics are always assumed to be ***one sigma errors*** unless stated otherwise. A range of ***+/– 2σ*** should contains 95% of the measured values, while ***3σ*** contains 99.7%. Of course, any particular set of measurements will have fluctuations (and these fluctuations can also be characterized in terms of statistical probabilities) but in the limit of a large number of measurements, the distribution of measured values should follow these values.

Now, the actual task

Suppose you want to know whether a given coin will actually come up heads 50% of the time. To find out, you will need to measure the probability of obtaining the outcome of heads. We want to flip a coin some large number of time (say 10) and record how many times the coin comes up heads. We then repeat this process some large number of times (again, at least 10). To make your life easier in the modern age, there are now web sites that will do this for you, although the best method still seems to be to put 10 coins in a box and shake to flip them all at once. The more data you collect, the better, but don't go crazy. 20 sets of 10 flips is more than enough. Each time you flip the 10 coins, write down how many turned up heads.

To analyze your data, first make a histogram showing how many times you observed a particular outcome. This can be done in Excel, although it is just as easy to do this by hand. An example is shown in the figure on the next page. This shows an actual experimental distribution of coin flip data I made myself. The histogram shows that I obtained 3 heads 3 times, 4 heads 5 times, and once I had all ten coins come up heads.

In order to estimate the parent distribution (and hence measure the probability of obtaining heads experimentally), calculate the ***mean*** and ***standard deviation*** of your data. This is probably easiest to do in a spreadsheet. For my data, the mean is 5.45 and the standard deviation is 1.77 for 20 trials. The standard deviation for a Binomial distribution is expected to be $\sigma_r = \sqrt{Np(1-p)}$. ***How does this compare to what you found in your data?***



Sample Uncertainties

In my data shown above, I have found heads 54.5% of the time. Is this consistent with the expected 50%? To answer that question we need to know the uncertainty on our measured mean value. The standard deviation calculated above gives the uncertainty on any specific measurement. Not surprisingly, the mean value for any set of measurements also follows some probability distribution, and the uncertainty on the mean can be defined as the ***standard deviation of the mean***. This quantity has a rather simple form given by $\sigma_{\bar{x}} = \sigma_x/\sqrt{N}$. That is to say, the standard deviation of the mean is simply the standard deviation of the sample divided by the square root of the number of measurements.

Using my data above, then, my error on the mean is $\sigma_x/\sqrt{N} = 1.77/\sqrt{20}$ or 0.4. In terms of percentage, then, I have measured a (54 +/- 4)% chance of getting heads, which is consistent with 50%. If I wanted a more precise measurement, I would need to flip more coins. The equation above tells me that the uncertainty in the mean improves as the square root of the number of measurements taken, so to make a measurement which is twice as precise would take four times as many data trials (or 80 sets of 10 flips in this case).

Calculate the uncertainty on the mean for your data and write down your measured probability for obtaining an outcome of "heads". Is this consistent with 50%? Don't be too surprised if it is not, as for a Gaussian distribution, 32% of the time, it should not be. Are you within 2 sigma? If there are 20 lab groups in the class, how many would we expect to have a 2 sigma deviation or larger? In particle physics, we only really get excited if a measurement is at least 3 sigma from its expected value.

Task 4: Analyzing Class Data

We will now perform the same statistical analysis from Task 3 useing the class data on the height of the atrium. These will be posted on the class website by the start of week 4. Unlike flipping a coin, we do not know *a priori* how this data will be distributed, but we will assume a Gaussian distribution and quote errors accordingly.

The analysis steps are as follows:

- Make a histogram of the class height data. Don't forget the units.
- Find the mean and standard deviation for this data sample.
- Taking all of the class data together, what is the combined measurement of the atrium height? What is the uncertainty on this measurement?

You may be wondering what we should be doing with the quoted errors on the individual measurements. Here we have ignored them. To properly take these into account, we would have to perform some kind of weighted combination which gives more emphasis to the more precise measurements. This can be done, but it is beyond the scope of this current lab.

One test which we can do is to use the observed distribution of measured values to check whether the error estimates made by the groups are accurate. The error quoted on each group's measured value is supposed to reflect the uncertainty of that measurement. The standard deviation of an ensemble of measurements is also an estimate of the uncertainty of each measurement. So how do these two estimates compare? Since each group probably has rather different errors, calculate the ***mean error*** quoted by all of the lab groups. Compare this to the standard deviation of the atrium height values. Which is larger? What does this tell you about the error estimates made by each group?

Finally, the variation of class data gives a measure of certain class of uncertainties, but there can also be common systematic uncertainties that are the same for all of the groups. One example might be if the floor of the atrium isn't flat. Another would be if the long measuring tapes are not accurate. In your final combined measurement of the atrium height, how large would you expect these common systematic uncertainties to be in comparison with the "statistical" uncertainty you have calculated directly from the variations observed in the data? Describe why measuring the height using a completely different technique would be useful to make a more accurate measurement.