Lab 1

Fundamentals of Measurement and Error Analysis

1.1 Overview

This first laboratory exercise introduces key concepts and statistical and plotting tools that are used throughout the entire sequence of exercises. It also focuses on the idea of productive scientific experimentation – how to design, conduct, and evaluate an experiment. There are four primary activities: (1) planning and conducting a short experiment with common household items, (2) examining existing data to uncover a basic connection between seasonal changes and the height of the Sun in the sky at noon, (3) analyzing data, including error estimates, and (4) making appropriate conclusions based on evidence.

Students begin by choosing a simple experiment that can be conducted with common household supplies. We provide a list of ten sample experiments and encourage students to select one from the list or to propose a similar experiment to their instructors. The experiments have been designed so that they are non-threatening (it is difficult to be intimidated while tossing marshmallows, or to feel that this is something that one should have paid more attention to in high school), easy to visualize (they involve measuring lengths of time or distance, weights, or counting items), and can produce significantly different results when conducted by different individuals in different environments (the average width of books drawn from a shelf of children's picture books will differ from that for a set of novels). These results of these experiments cannot be "looked up" on Wikipedia or found recorded in scientific texts. Each experiment is described by only a single short sentence (such as "what is the average distance between the pupils of people's eyes?"), leaving students to fill in the blanks and construct an experimental plan of action.

Each student designs their experiment, and submits an experimental design within their laboratory report. They take a small amount of preliminary data, and then evaluate their protocol and their data to determine whether they need to vary the plan (strong winds might make marshmallow tossing untenable outside; a scale designed to weigh humans will not work well for individual pinto beans). They then revise their design, and take a set of 30 measurements. A sample experiment is discussed in the project chapter to provide guidance, and a web-application allows students to vary the number of data points and the precision of measurements in order to determine the effect on derived quantities such as mean values and standard deviations.

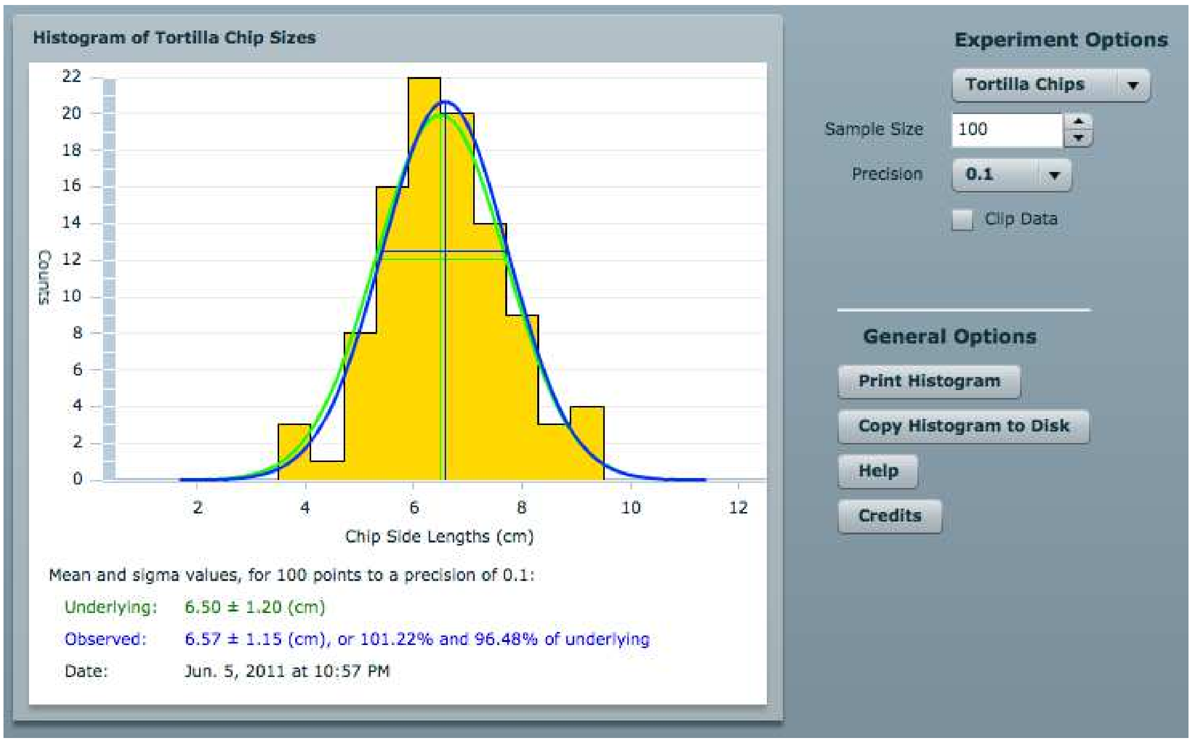

Figure 1.1: A web-application which demonstrates the basic properties of histograms of measurements. Students can choose from a variety of samples drawn from simple hands-on experiments and define a sample size and the precision of modeled measurements. A Monte Carlo simulation presents a set of representative data in the form of a histogram; each simulation run is generated on the fly and differs from previous and future runs (as would real data samples). Mean values and standard deviations are presented for both the underlying complete sample and the simulated finite sample, with differences that illustrate the effect of varying the sample size and the precision of the measurements.

Figure 1.1 is a screen capture of the primary window for this histogram-based web-application. This tool allows students to simulate collecting data for various physical samples such as the lengths of tortilla chips or the weight of pinto beans. By selecting different values for the sample size and the measurement precision, they are able to study the relationship between these factors and the intrinsic scatter in the measured quantity. For example, one might measure the distance that a marshmallow can be thrown to the nearest inch, while needing significantly higher precision to measure the length of a tortilla chip. By repeatedly creating small samples, students can observe how random scatter and sample selection can produce changes in the measured mean values and standard deviations.

All of our web-applications are supported with help screens explaining tool usage and options to either print figures or to save them (in PNG format) as files on local disks. Figures are all date- and time-stamped to the nearest minute. These flash-based applications are fully supported under Windows, Mac iOS, and the Linux operating system, and run within all major browsers.

After completing their experiments, students trade experiment designs with another student. Each student then attempts to reproduce the experiment conducted by their partner, working from just the submitted description of the experiment. The pairs of students are then connected via e-mail. Students use plotting tools to create histograms of their own data. They work together to determine whether their experiments were in agreement or not. Our rule of thumb is that values within two standard deviations of each other do not disagree, differences between two and five standard deviations suggest that it would be good to take additional data and verify experimental set-ups, and differences of more than five standard deviations suggest that there is a significant difference between the two experiments that should be identified. Deficiencies in the descriptions of the experiments come to light quickly as students are frequently forced to guess about critical factors left undefined by their partners and students need to agree on a single set of units in order to calculate and compare mean values and standard deviations.

We have elected to keep this statistical test as simple as possible, to keep it accessible to students with no experience and no intuition for judging data. We make the point separately that larger data sets generally produce better constrained mean values and standard deviations.

Success in the partnering exercise is not defined by whether the students were able to measure similar mean values for quantities, but instead by whether they could discuss and identify differences or agreement in derived quantities and whether they could suggest probable causes for any differences. This activity also prompts of students to work cooperatively in order to succeed, and establishes peer partnerships that students are able to develop further throughout the semester. It also illustrates the need for scientists to communicate their ideas clearly, and to write well enough to describe experiments adequately to their peers.

We introduce three types of "errors" (causes for differences between repeated measurements of a given quantity) to students: (1) natural variation in a population, that produces a range of values that can be measured due to the intrinsic width of a distribution of a property, (2) measurement errors, tied to the precision with which a value is measured, and (3) systematic errors, in which a bias is introduced into a data set (such as a systematic overestimate due to measuring people's heights while leaving their shoes on). These concepts are developed throughout the entire sequence of projects, to build familiarity and comfort.

Additional tools for analyzing data include error bars and fits of particular models to data. We illustrate linear fits, and also show how in some cases a higher-order function can produce a better fit to xy data. The accuracy of various fits is expressed in terms of root mean square deviations from the models. We discuss several sample data sets (using illustrative data sets that will be explored more deeply throughout the semester), and then students work through a plotting and analysis exercise for themselves with worldwide data on surface temperatures and peak solar altitudes at high noon. Because many of our students are unfamiliar with the basic idea of creating and analyzing a plot of data in order to test a hypothesis, we emphasize first the basic skill of reading plots, and then develop the idea of fitting simple models to data.

1.2 Learning Objectives

After completing this laboratory exercise, the student should be able to do the following:

- Design and carry out simple scientific experiments, while guarding against strong biases in sampled populations, selecting a sufficiently precise measurement technique, and sampling a sufficiently large number of times to achieve a desired outcome.

- Comprehend the meaning of natural variance, and measurement and systematic error, and gauge their effects on data.

- Understand how to plot a histogram from repeated measurements, and the meaning and proper use of mean and standard deviation values.

- Explain how to plot two variables against each other, fit a line to them with a slope and y-intercept, and interpret the results in view of associated errors.

- Relate an rms deviation for a fit to a set of (x, y) data to a standard deviation for repeated measurements of a single quantity.

- Connect correlation coefficients to relationships between two variables, primarily in a qualitative sense.

- Compare multiple measurements of a particular quantity, with associated errors, and determine whether the results are in agreement or disagree.

1.3 Teaching Toffees

Teaching toffees are practical tips based on the hard-won experience of fellow instructors.

For all GEAS lab exercises:

Please have students access the list of components for the lab exercise (lab chapter, video tutorial, web applications, report template) directly from the GEAS website so that they use the current version of all tools.

By having them save and share the lab report template, they will have their work backed up automatically and instructors will be able to observe their progress 24/7. This allows instructors to check for appropriate progress towards intermediate goals over a two-week period, and for students and instructors to interact asynchronously through threaded comments placed in the report margins at exactly the point where a question occurs. Note that the Lab #1 video tutorial describes how to work with Google resources, and there is a two-page handout on the lab website as well.

Students sometimes try to complete the lab exercise by skipping over the lab chapter entirely and working from just the report template. This is not an effective learning strategy, nor does it save time. The report template contains only abbreviated forms of the questions with spaces for answers, but none of the context or framework for the questions.

We strongly encourage you to strongly encourage all of your students to watch the video tutorial before beginning to work on the exercise. If a picture is worth a thousand words then a video might be worth a million words. Each tutorial is designed to introduce the scientific concepts that motivate the companion exercise. Because students may be working remotely on their own, it also focuses on the most potentially challenging aspects of each project and walks them through key steps by showing them what they look like in practice.

For Lab #1 in particular:

This lab #1 (Fundamentals of Measurement and Error Analysis) can serve as an effective introductory exercise for laboratory work in any of the scientific fields. It is not specific to astronomy.

The video tutorial for this exercise explains how to create a Google login (for those who may not already have one) and how to copy and save a copy of the lab report template. Students may use the provided Microsoft 365 report template instead if you wish; what is important is that they copy a current report template from the GEAS website and that they share it with you before beginning to work on the lab exercise so that you can observe their incremental progress and intervene if they go off track.

The video also describes how to approach conducting a simple experiment with repeated measurements, emphasizing general guidelines that apply to any experiment. The concepts of mean values and standard deviations for a set of measurements are introduced, and students are shown how to enter measurements into a plotting tool on the GEAS website to create an annotated histogram of their results. Students are then shown how to work with another student to reproduce each other's experiments, and to evaluate their results in doing so. Natural variation, and measurement and systematic errors, are discussed.

The next video topic is the plotting of one variable against another to determine a linear trend, with error bars attached to the dependent variable. A GEAS plotting tool is again used to create an annotated figure. Students are also shown how to save a plot as an image file and then insert it into a lab report.

A common challenge in conducting this laborator exercise is that it is generally placed at the beginning of the semester, and so students who add the class late will need to catch up quickly. Because students perform their own simple experiment and then reproduce another student's experiment, it is important that everyone completes their experiment on time so that they can be paired up and have time to do a second experiment (though note they will need to make only 7 measurements, not another 30) and compare results.

It is helpful to pair students who measure a length-scale for their own experiment with someone who measured something different (such as a weight) for variety, but it is not essential. If anyone weighs light objects which require a kitchen or jeweler's scale, make sure that their partner has access to similar equipment. Sometimes a particular experiment (throwing marshmallows, or measuring tortilla chip lengths) is very popular, and so one has to be careful to not end up with two or more students with the same experiment at the end of the pairing procedure.

We have been impressed with how supportive students generally are towards their partners; this exercise can be a good start to guiding students towards more peer-peer discussions throughout the semester. Student sometimes leave out a critical detail of their experiment design when describing it, and the impact that this has on their partner's attempts to reproduce their experiment bring home the importance of being thorough.

1.4 Keywords

Altitude – The altitude of an object in the sky is the number of degrees which it lies above the horizon. At local noon the Sun could have an altitude of 90 degrees (lying directly overhead), and as it sets the altitude falls to a value of zero.

Correlation coefficient – The correlation coefficient R is a measure of the strength of the relationship between two variables x and y. It ranges from -1 to 1, where +1 indicates the strongest possible positive correlation (as x increases, so does y), zero indicates no predictive relationship between quantities, and -1 indicates the strongest possible negative correlation (as x increases, y decreases). Correlation coefficients are well-suited for determining zero-point offsets in periodic relationships (such as syncing sine waves to remove phase offsets).

Data set – A data set is a collection of measurements made within an experiment.

Error bar – An error bar is a symbol attached to a point on a plot, which shows the associated error (how much the point have might shifted in position due the way in which it was measured). It often resembles a small bar (or line) placed on one side or another of the point value.

Histogram – A histogram is a plot which shows the number of measurements of a particular quantity which fall within bins defined to extend over the range of measured values. The bin size should be selected so that the bins with the largest number of measurements within them hold a statistically meaningful number of measurements, and should also not be smaller than the precision (the resolution) of the measurements.

JPG format – Images are often stored on computer disks in JPG-format files, a format which allows the files to be stored and transferred from computer to computer without loss of information. A JPG-format file should have a file name which ends with the extension ".jpg," so that the image analysis and display packages can recognize its contents.

Mean value – The mean value μ of a set of N repeated measurements m_i is defined to be the unweighted average, or

$$\mu = \frac{1}{N} (m_1 + m_2 + m_3 + \ldots + m_N) = \frac{1}{N} \sum_{i=1}^{N} m_i.$$

Measurement error – Measurement error refers to the precision with which a set of measurements were made (to how many decimal places the measured values were recorded).

Model – A model fit is a mathematical expression which attempts to reproduce the relationship between two or more variables.

Mu – The Greek letter "m" (μ), often associated with the average value of a set of measurements.

Natural variation – Natural variation refers to the intrinsic width of a distribution of a measured property.

Precision – The precision of a measurement is defined as the smallest change in its value which can be observed with a given experimental technique. A ruler with markings every millimeter (mm), for example, could carry a precision of ± 0.5 mm.

RMS (root mean square) deviation – The rms deviation is the square-root of the average square of the offsets in y between a set of N data points and a fit function. For a linear fit, where $y = mx + b$,

$$\text{rms} = \sqrt{\frac{1}{N} \sum_{i=1}^{N} [y_i - (mx_i + b)]^2}.$$

Scatter plot – See xy plot.

Sigma – The Greek letter "s" (σ), often associated with a measurement of a standard deviation.

Slope – The slope m of a line is the change in y divided by the change in x, or for two points along the line with coordinates (x_1, y_1) and (x_2, y_2),

$$m = \frac{\Delta y}{\Delta x} = \frac{y_2 - y_1}{x_2 - x_1}.$$

Standard deviation – The standard deviation σ, also called the spread, of a set of N repeated measurements m_i with an mean (average) value μ is defined as

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

Systematic error – A systematic error is one which biases all of a set of measurements in the same fashion (as opposed to making some smaller and some larger).

xy plot – A plot which shows the relationship between two variables by plotting one along an x-axis and the other along a y-axis is commonly called an xy, or scatter, plot.

y-intercept – The y-intercept b of a line is the y coordinate of the point on the line for which $x = 0$, or for two points along the line with slope m and coordinates (x_1, y_1) and (x_2, y_2),

$$b = y_1 - mx_1 = y_1 - \left(\frac{y_2 - y_1}{x_2 - x_1}\right) x_1.$$

1.5 Relevant Lecture Chapters

This laboratory exercises draws upon the material in Chapter 2: Scientific Notation and Chapter 9: The Scientific Method. There are peripheral connections to various astronomical topics discussed throughout the course. Questions in the self-review database for Chapter 9 review the statistical material and data fitting techniques for this exercise.

1.6 References and Notes

1. Data for summer temperatures at various locations around the globe were extracted from records at wunderground.com.

2. Figure 1.1 through Figure 1.8 are shown courtesy of Nicole Vogt.