ESA21
Environmental Science Activities for the 21st Century

Table of Contents:

Measurement: An Introduction

In general, you are quite familiar with measurements, as almost any occupation requires measurements of some kind. Carpenters measure boards for cutting, nurses measure blood pressure in patients, tailors measure fabric for garments, and advertising executives measure the public's acceptance of their sales pitches. You will therefore undoubtedly be utilizing measurement in your chosen career, regardless of the field you enter.

Measurement plays a particularly large role in science. In their studies, scientists gather data, and to do this they use measurements. Scientists measure the concentration of gases in the atmosphere, the growth of organisms under varying conditions, the rate of biochemical reactions, the distance of stars from the earth, and an innumerable number of other things. As measurements form the basis of scientific inquiry, they are deserving of in-depth analysis in lab. In a scientific experiment, the investigator examines the effects of variations in the independent variable on the dependent variable through measurements. For example, let's assume a biologist is studying the effect of temperature on plant growth. She sets up several different temperature conditions, and grows groups of plants from seedlings in each condition. When the experiment ends, she must compare plant growth in the plants from different temperatures. But how should she do this? Should she just look at the plants and decide which grew the best? Should she pick up the plants and "feel" which ones have the greatest mass? Of course not. She would use some sort of quantitative measurement, such as measuring the height of each plant's stem in centimeters or determining the total plant biomass in grams. Whichever measurement she chooses, she would need to utilize an instrument to make it.

Instruments

Take a minute and look around you at the variety of objects surrounding you. There's probably a few pens or pencils, a notebook or two, assorted computer equipment, and other things. While you may be able to easily distinguish the differences between some objects (e.g., your notebook is longer than your pen), other differences are more difficult to discern. You may not be able to easily determine, for example, whether your computer keyboard or your course textbook has greater mass. Even when we can distinguish differences, it is not always easy to determine the extent of those differences. You may have noted that your computer monitor is heavier than your notebook, but is it twice as heavy or three times as heavy? Using only our senses, we cannot be certain about the answer, so we must take measurements.

To take measurements we need instruments. Instruments include simple things like rulers and graduated cylinders, and complicated electronics like pH sensors and mass spectrometers. All of these instruments provide us what our senses cannot - a quantified measure of the properties of an object. In today's lab you will use get an introduction to measurement by using various instruments to measure the linear dimensions, volume, and mass of objects. As all scientific measurements utilize the metric system, we must first say a few things about it before proceeding on.

The Metric System

If you are measuring something, you need "units" to describe the object. In formal terms, a scale of measurement is the assignment of numbers or symbols to measure an attribute. In the past, natural units of measurement, such as a "foot", were commonly used. Unfortunately, these units were somewhat arbitrary. In Roman times, for example, a "foot" in England was 29.6 centimeters. When the Saxons took over, the size of a "foot" grew to 33.5 cm. Five centuries later, it was reduced to 30.5 cm. Finally, in 1959, the "International Foot" was defined as 30.48 cm. Even today, a "foot" in England is different from a "fod" in Denmark (31.41 cm), a "fod" in Sweden (29.69 cm), and a "fuss" in Germany (31.61 cm). With the increase in international trade during the 18th century, merchants needed to standardize units of measurement. This resulted in the development and nearly universal adoption of the metric system around the world. Of course, the United States is a notable exception to this worldwide trend, as we continue to use the English system of measurement. We buy our gas in gallons, measure our weight in pounds, and gauge driving distances in miles. The metric system has crept into our society somewhat (e.g., the two-liter soda bottle), but universal acceptance of this system of measurement anytime soon is unlikely.

In science, use of the metric system is unquestioned. Because of its international familiarity and ease of use, scientific studies utilize metric measurements. All of the measurements you make in this course must therefore be in metric units, and this is something you should keep in mind throughout the semester. Now as we've all dealt with the metric system during high school, a review of the system will not be provided here. If you need a refresher, please visit the web sites below for additional information.

The NIST Reference on Constants, Units, and Uncertainty
Physics Laboratory at NIST
http://physics.nist.gov/cuu/Units/index.html

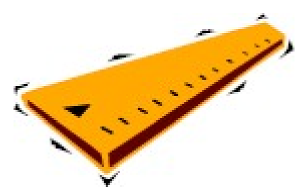
The Metric System
Gordon Speer
http://www.essex1.com/people/speer/metric.html

U.S. Metric Association, Inc.
U.S. Metric Association, Inc.
http://lamar.colostate.edu/~hillger/

How Many? A Dictionary of Units of Measurement
Russ Rowlett, University of North Carolina
http://www.unc.edu/~rowlett/units/index.html

Procedures: Linear Measurements

Common linear measurements in the metric system include the kilometer (km), meter (m), centimeter (cm), and millimeter (mm). Naturally, each is used to measure objects at different scales. The distance between cities would best be measured in kilometers, while the distance between your toes is best measured in millimeters. You will first gain some experience with linear measurements, and then move on to other types of measurements.



Station 1A: Measuring Linear Dimensions
(1.) From Station 1A, obtain three objects and a ruler.
(2.) Measure the length, width and height of each block (or height and diameter for cylinders) and record your results to the nearest tenth of a centimeter in Table 1 on the Activity Sheet. You will complete the volume column at Station 1C, so leave it blank for now.
(3.) Return the ruler to Station 1A, keep the three blocks, and move on to the next station.

Procedures: Measuring Mass

The mass of an object is the amount of matter it contains. This quantity does not change without making changes in the object itself. It is not to be confused with weight, which may differ from one place to another when there is a change in gravity. For example, you would weigh less on the moon than you do on the earth, but your mass is the same. Mass is measured on a balance as opposed to a spring scale, which measures weight. The basic unit of mass is a gram (g). A kilogram (kg), which is a thousand grams, is used to measure large objects. Small quantities are measured in milligrams (mg), which are a thousandth of a gram.

Station 1B: Measuring Mass
(1.) Take your blocks to one of the triple-beam balances at Station 1B. Before massing each block, make sure the instrument is balanced. Move the markers on each of the three beams to the last position on the left. Note that the markers for the top two beams fit into notches on the beam. The pointer for the bottom beam should point to zero. With the markers in this position, the arrow on the right should point to the balance line. If it does not, ask the instructor on duty to balance the instrument.
(2.) Put one of the blocks on the pan and move the marker on the top beam to 10, making sure it slips into the correct notch. If the arrow on the right is above the line, the block's mass is greater than 10 grams, and you should move the marker to the next highest mark. If the arrow on the right is below the line, the block's mass is less than 10 grams, and you should move the marker to the next lowest gram mark. Continue in this fashion, using the other beams until the arrow on the right moves to the same level as the line. Once it's level, you've found the object's mass. Mass to the nearest hundredth of a gram and record it in Table 1.
(3.) Mass each of the other two blocks, record them, and move on to the next station.

Procedures: Measuring Volume

The volume of an object is the amount of three-dimensional space the object occupies. The basic unit of volume in the metric system is the liter (l). Its most common subdivision is the milliliter (ml), which is a thousandth of a liter. You should be roughly familiar with a liter, given that two-liter bottles of soda are quite common. There are several ways to calculate the volume of an object, and we will investigate them at this station.

Station 1C: Measuring the Volume of a Liquid
(1.) From Station 1C, obtain a film can or other small container and a graduated cylinder. Notice that the graduated cylinder has calibration markings on the side, listed in milliliters.
(2.) Put some water into the graduated cylinder from the faucet. To read the volume of a liquid in a graduated cylinder properly, hold the graduated cylinder so that the surface of the liquid is at eye level. Note that the surface of the liquid is concave. Read the volume at the bottom of the curve in the middle.
(3.) Empty the graduated cylinder, carefully fill the film can or other small container, and pour the water into the graduated cylinder. Determine the volume of the can/container.

Calculating the Volume of an Object

The volume of objects with a regular, geometric shape such as a cube or a cylinder can be determined by applying the appropriate formula. For example, the volume of a block can be calculated by multiplying the length times the width times the height. The volume of a cylinder is equal to π (3.14) times the height times the radius squared.

$$V \text{ (of cube)} = lwh$$

$$V \text{ (of cylinder)} = \pi * r^2 h$$

Calculate the volume of each of your blocks and record the answers in Table 1. Note that you measured the volume of a liquid in milliliters, but calculated the volume of the block in cubic centimeters (cc). One cubic centimeter is the same volume as one milliliter though, so this is not a problem (1 cc = 1 ml).

Measuring the Volume of an Object
Calculating the volume of an object from its dimensions is not always possible. How would you calculate the volume of your head? What about rock? The volume of an irregularly shaped object can be found by measuring the amount of water it displaces. One way to do this is to place the object in a known quantity of water and measure the volume of the two combined. From this new volume, subtract the original volume of the water. The answer is the volume of the object.

Station 1C: Calculating the Volume of an Object
(1.) Put 15 ml of water into your graduated cylinder.
(2.) Add one of the pebbles from Station 1C and record the new volume in Table 2.
(3.) To find the volume of the pebble, subtract 15 ml from the new volume.
(4.) Repeat the procedure for two additional pebbles of differing sizes and enter the data in Table 2.
(5.) Return your blocks to Station 1A and all other materials to Station 1C.

Procedures: Calculating Density
Now that you are familiar with mass and volume, let's move on to density. The density of a material is the amount of mass per unit volume. Dense materials have large mass in small volume, while materials with low density have the opposite relationship. For example, a 1 cc lead block is denser than a 1 cc block made of cork, because the lead block contains more mass.

Density = mass/volume

As the formula indicates, you can determine the density of an object if you've measured its mass and volume. Use these measurements to calculate the density of your three objects and enter the results in Table 1.

Did you ever wonder why some objects float and others sink when placed in water? The answer concerns density. If an object has a density less than or equal to the liquid in which it is placed, it will float. If its density is greater than that of the liquid, it will sink. Water has a density of 1 g/cc, indicating that substances with densities less than 1 g/cc will float when placed in water, and those with densities greater than 1 g/cc will sink. Hence, density can be an important measure of an item's properties under the proper circumstances.

Inferential Statistics: Introduction
Up to this point, we've discussed the proper methods for taking measurements, and you've gained some experience with simple instruments. In this section, we will follow up on that by introducing a new group of statistics - inferential statistics. Before I define inferential statistics, let me show you why they are useful. In the previous section, I described a research study that sought to determine the effects of temperature on plant biomass. Upon completion of her experiment, the experimenter would have sets of biomass measurements for each group. What does she do then? Should she take the mean of the values for each group and then make conclusions based on this statistic alone? What if the mean biomass for one group is only *slightly* higher than that for another group - is the difference sufficient for her to make a solid

conclusion? Inferential statistics allow you to make comparisons in scientific studies and determine with confidence if differences in treatment groups truly exist.

Inferential Statistics: An Example
Inferential statistics are used to make comparisons between data sets and infer whether the two data sets are significantly different from one another. It is important to realize that when dealing with statistics and probability, chance always plays a role. When we compare means from two groups in an experiment, we are attempting to determine if the two means truly differ from one another, or if the difference in the means of the groups is simply due to random chance. The best way to explain this concept is with an example.

Chance and "significant" differences: A Case Study

After losing a close game in overtime, a local high school football coach accuses the officials of using a "loaded" coin during the pre-overtime coin toss. He claims that the coin was altered to come up heads when flipped, his opponents knew this, won the coin toss, and consequently won the game on their first possession in overtime. He wants the local high school athletic association to investigate the matter. You are assigned the task of determining if the coach's accusation stands up to scrutiny. Well, you know that a "fair" coin should land on heads 50% of the time, and on tails 50% of the time. So how can you test if the coin in question is doctored? If you flip it ten times and it comes up heads six times, does that validate the accusation? What if it comes up heads seven times? What about eight times? To make a conclusion, you need to know the probability of these occurrences.

To examine the potential outcomes of coin flipping, we will use a Binomial Distribution. This distribution describes the probabilities for events when you have two possible outcomes (heads or tails) and independent trials (one flip of the coin does not influence the next flip). The distribution for ten flips of a fair coin is shown in Figure 1.

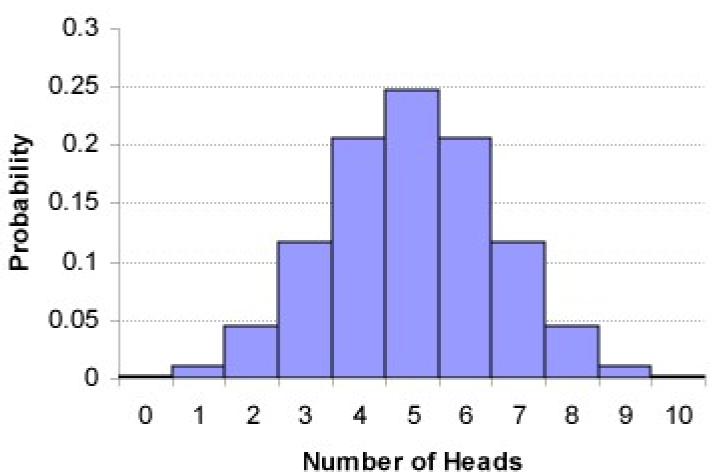


Figure 1. Binomial distribution for fair coin with ten flips

Note that ratio of 5 heads:5 tails is the most probable, and the probabilities of other combinations decline as you approach greater numbers of heads or tails. The figure demonstrates two important points. One, it shows that the expected outcome is the most probable - in this case a 5:5 ratio of heads to tails. Two, it shows that unlikely events can happen due solely to random chance (e.g., getting 0 heads and 10 tails), but that they have a very low probability of occurring.

Also note that the binomial distribution is rather "jagged" when only ten coin flips are performed. As the number of trials (coin flips) increases, the shape of the distribution begins to smooth out and resemble a normal curve. Note how the shape of the curve with 50 trials is much smoother than the curve for 10 trials, and more representative of a normal curve (Figure 2).

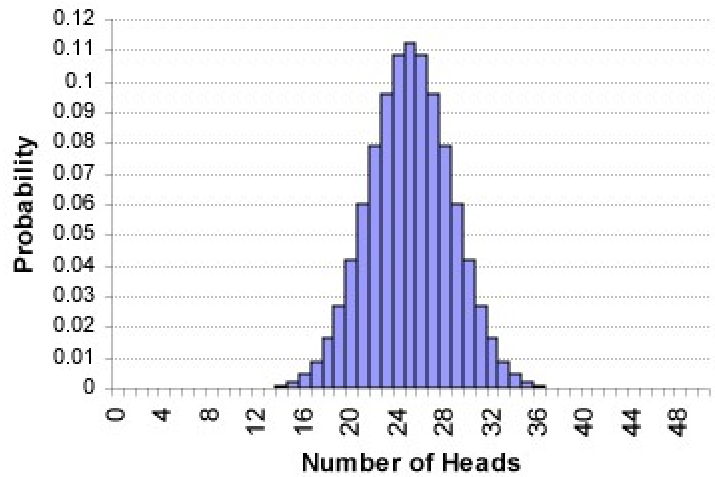


Figure 2. Binomial distribution for fair coin with 50 flips

Inferential Statistics: Probability
Normal curves are useful because they allow us to make statistical conclusions about the likelihood of being a certain distance from the center (mean) of the distribution. In a normal distribution, there are probabilities associated with differing distances from the mean. Recall that 68% of the values in a distribution are within one standard deviation of the mean, 95% of values are within two standard deviations of the mean, and 99% of the values are within three standard deviations of the mean.

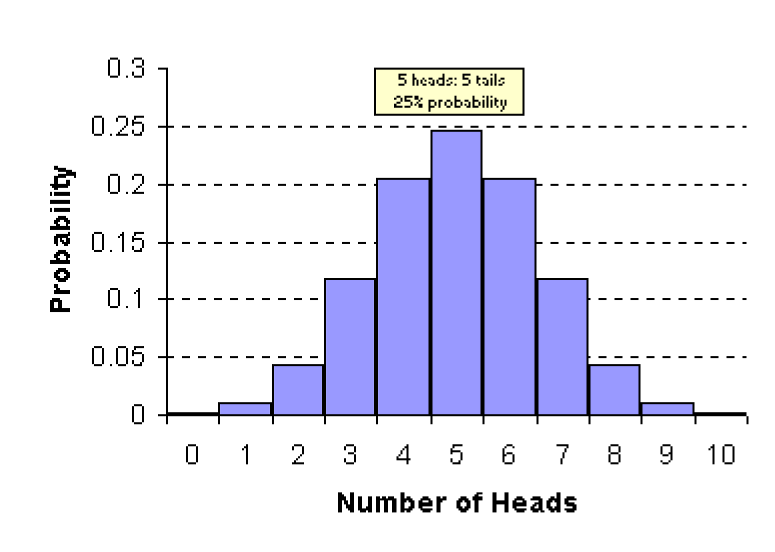


Figure 3. Probabilities for results of 10 coin flips [animate]

The difficulty with working with probabilities is knowing when to conclude that an occurrence is *not* due to random chance. Values far from the mean in a distribution can occur, but will occur with low probability (Figure 3). We are therefore essentially testing the hypothesis that the observed data fits a particular distribution. In the coin flip example, we're testing to see if our results fit those expected from the distribution of a fair coin. So we need to come up with a point at which we can conclude our results are definitely not part of the distribution we are testing. So when do you determine that a given data set no longer fits a distribution when random chance will always play a role? Well, you've got to make an arbitrary decision, and statisticians set precedent long ago. Given that 95% of the values in a distribution fall within two standard deviations of the mean, statisticians have decided that if a result falls outside of this range, you can determine that your data does not fit the distribution you are testing. This essentially says that if your result has equal to or less than a 5% chance of belonging to a particular distribution, then you can conclude it is not a part of that distribution. As probabilities are listed as proportions, this means that a result is "statistically significant" if its occurrence is equal to or less than 0.05. This leads to our statistical "rule of thumb" - whenever a statistical test returns a probability value (or "p-value") equal to or less than 0.05, we reject the hypothesis that our results fit the distribution we are testing. The standard practice in such comparisons is to use a null hypothesis (written as "Ho"), which states that the data fits the distribution.

Ho: The data fit the assigned distribution

To practice your interpretation of p-values, decide if each of the p-values below indicates that you should reject your null hypothesis. Answers are provided at the end of the exercise.

PRACTICE PROBLEM #1

$p = 0.11$	Reject or Do Not Reject Ho?
$p = 0.56$	Reject or Do Not Reject Ho?
$p = 0.99$	Reject or Do Not Reject Ho?
$p = 0.01$	Reject or Do Not Reject Ho?
$p < 0.005$	Reject or Do Not Reject Ho?

So our coin test is comparing our result to the distribution of a fair coin. To test the coin, you opt to flip it 50 times, tally the number of heads and tails, and compare your results to the fair coin distribution. You obtain the results listed below.

Heads	33
Tails	17

So what does this mean? Referencing the distribution (Figure 2 below), we see that a ratio of 33 heads to 17 tails would only occur about 1% of the time if the coin were indeed fair. As this is less than 5% ($p < 0.05$), we can reject our hypothesis that the result fits the distribution. We were testing the distribution of a fair coin, so this suggests the coin was not fair, and the coach's accusation has merit. If we look at a distribution for a rigged coin that comes up heads 70% of the time instead of 50% of the time (Figure 4), we notice that our result fits quite well into this distribution. This indicates that further tests should be conducted, and the number of trials (coin flips) increased so a more definitive conclusion could be reached. Man, I love a good controversy...

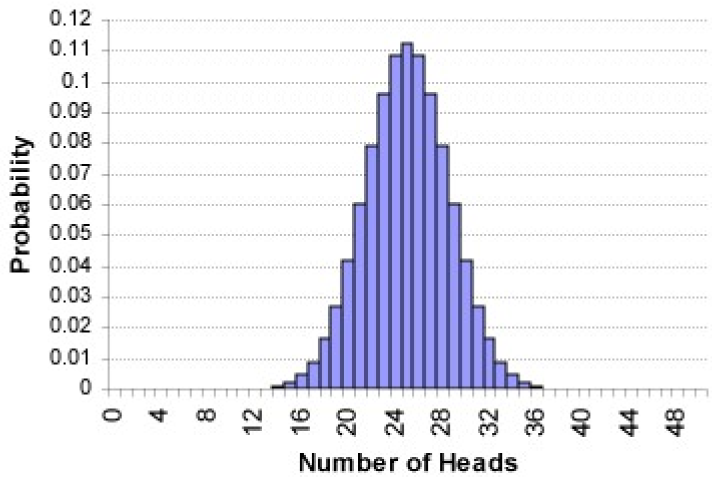


Figure 2 (again). Binomial distribution for fair coin with 50 flips

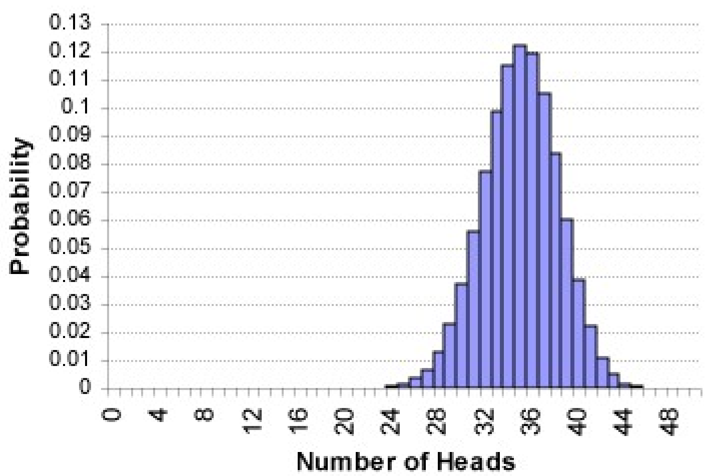


Figure 4. Binomial distribution for coin rigged towards "heads"

Inferential Statistics: Using Inferential Statistics

In the coin flip example, we tested to see if the results of our tests fit the distribution of a fair coin based on the predicted probabilities of the various outcomes. Most statistical tests in the sciences do not work this way, however. In most cases, scientists manipulate variables in experiments, and then make comparisons between groups receiving different treatments. For example, if you are investigating the effects of different brands of fertilizer on tomato plant growth, you would be interested in comparing the growth of the plants in the different treatments to one another. In a situation like this, we are interested in comparing two groups to one another, rather than comparing one group to an existing distribution. To do this, we will use a statistical test that compares the distributions of two data sets to one another - the t-test.

The t-test

The t-test is an inferential statistic that enables you to compare the means of two groups and determine if they are statistically different from one another. In essence, the test compares the distributions of the two data sets to one another, and tests the hypothesis that the two data sets belong to the same distribution. If there is a low likelihood that the two data sets belong to the same distribution (probability less than or equal to 5%), then we can conclude that true differences in the means of the two groups exist, and the two groups are significantly different from one another. The t-test accomplishes this task by looking at both the mean and the dispersion of the data in the two groups. Figures 5 and 6 will help to illustrate how a t-test works: Let's say we've got data from two groups that we wish to compare. To help visualize things, we can graph the distributions of the two data sets on one graph so we can see the mean and dispersion for the two groups (Figure 5).

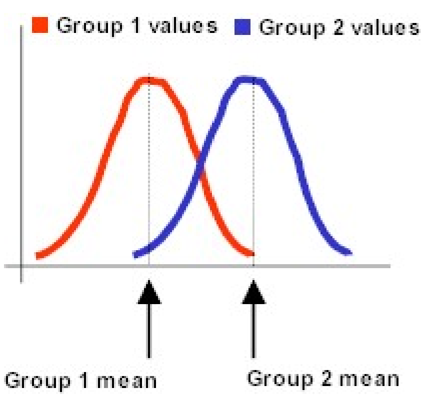


Figure 5. The t-test

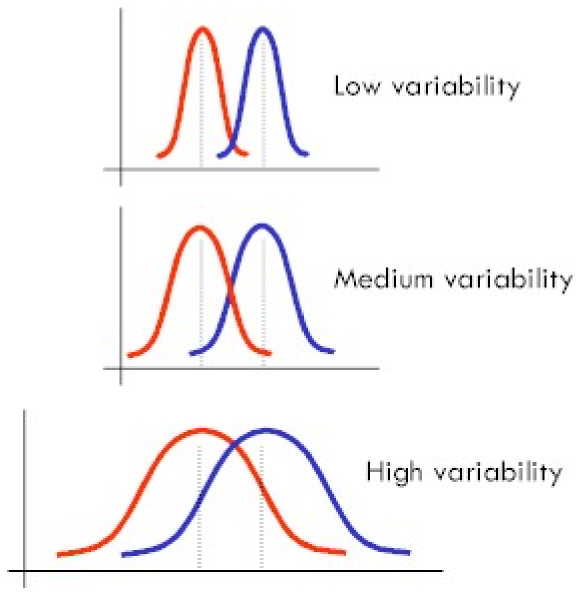


Figure 6. t-test comparison of distributions

A t-test compares both the means and distributions of data sets in order to determine if they are different from one another. Why does it do this? Why not just look at the means and make conclusions based on that? As illustrated here, the dispersion of the data tells you a great deal about the data set. Note that in each of the three scenarios shown (Figure 6), the means for the two groups are the same but their distributions are very different. In the low variability example, the distributions for the two groups are very narrow and overlap only slightly. In the medium variability example the distributions overlap much more, and in the high variability example the two distributions overlap almost entirely. A t-test looks at the ratio of the difference in group means to variability, in essence taking the ratio of the "signal" (the means) versus the "static" (the variance). Click *animate* below to see the formula.

Animated t-test formula [animate]

This ratio is the t-statistic, and the value of this statistic is used to determine the p-value for your test. The t-statistic is referenced to a statistical table and this determines the probability of that result being due to chance alone. Recall that a p-value equal to or less than 5% (0.05) indicates that the two groups are significantly different from one another. Our goal for this course is for you to gain experience with the t-test and to realize its usefulness in making conclusions when comparing groups of data. While we are not stressing the complexities and underlying mechanics of the t-test, you should be able to understand the test, use it to compare data sets, and correctly interpret the results the test gives you. That having been said, let's continue on and do an example to show you the usefulness of the t-test.

Inferential Statistics: The t-test - An Example

Let's assume that a researcher is attempting to determine the effects of pesticide pollution on the hatching success of fish eggs. He has noticed that fish eggs in streams near agricultural fields fare poorly, while those is undisturbed areas hatch successfully. The researcher sets up a lab experiment in which ten groups of fish eggs are allowed to develop in unmanipulated stream water, and ten groups of eggs develop in stream water with the addition of pesticide. The proportion of eggs hatching in each group is tallied (Table 1), and descriptive statistics for the two groups compared.

Statistic	No Pesticide	Pesticide
Mean proportion eggs hatching	0.87	0.59
Standard deviation	0.07	0.09

Group	Proportion eggs hatching	
	No Pesticide	Pesticide
1	0.80	0.50
2	0.76	0.45
3	0.81	0.68
4	0.90	0.77
5	0.95	0.64
6	0.84	0.60
7	0.88	0.54
8	0.99	0.57
9	0.86	0.62
10	0.93	0.57

Table 1. Effects of pesticide pollution on hatching success of fish eggs.

While the results indicate an adverse effect of pesticide exposure on egg hatching success, how can we be sure that the difference in the hatching success is "significant", and didn't just occur by chance? After all, the two data sets do overlap as the lowest value in the no pesticide groups was 76% (0.76) and the highest value in the pesticide groups was 77% hatching success. To compare these two data sets, we must first state our hypothesis. Our null hypothesis would be that the two groups are not different, and that both data sets belong to the same distribution.

Ho: The mean proportion eggs hatching in the two groups is not different

OR

Ho: The two experimental groups are part of the same distribution

We then conduct a t-test on the data set, which will examine the differences in mean and dispersion in the two groups, and provide a probability that the two groups are part of the same distribution.

As the p-value returned by the test was less than 0.05, we can reject our null hypothesis, conclude that the two groups are indeed from different distributions, and that they are significantly different from one another. The researcher can therefore conclude that pesticide exposure reduces the hatching success of eggs of this species. All of the comparisons you will be making in laboratory exercises this semester will mimic this example, so you should take special care to ensure you understand the operation and usefulness the t-test for comparing data sets.

Comparison	t-statistic	p-value
t-test	7.56	$p < 0.0005$

Performing t-tests

To perform t-tests on data sets, we suggest using an online t-test calculator from Graphpad.com. It can be accessed at the address below, and is exceptionally easy to use. WebCrunch, the program we are using to calculate descriptive statistics and create graphs, also has a t-test function, but we suggest you use the one below as it is tailored to the needs of a non-majors science course.

Graph Pad Statistical Application
GraphPad Software Inc.
http://www.graphpad.com/quickcalcs/index.cfm

(1.) Select the "Continuous data" option, then the "Continue" button.
(2.) Select the "*t* test to compare two means" option, then hit the "Continue" button.
(3.) Simply enter your data in the columns by group, select "Unpaired *t*-test", and then hit the "Calculate now" button. Your p-value and *t* statistic will be listed on the results page.

t-tests and Statistical Assumptions - A Word of Caution

Those of you familiar with the t-test have likely noticed that we are omitting a step in our use of the t-test - the testing of assumptions. For a given data set to be suitable for analysis with a t-test, it must meet two assumptions: (1.) the variance in the two groups being compared cannot be significantly different from one another, and (2.) the data must roughly fit a normal distribution. When statisticians and scientists conduct a t-test, they first verify these assumptions with statistical tests, and only proceed once these assumptions have been satisfied. If the variances in the two groups differ appreciably, the data can be mathematically "transformed" to bring variances closer together. If the data are not normally distributed, they can be transformed for normality, or an alternative test that does not require a normal distribution can be used.
As these steps appreciably increase the statistical complexity of t-test analysis, we will not be testing data sets for assumptions in this course. You must therefore realize that the statistical rigor of your results may not be comparable to that in published scientific studies, and that we are consciously avoiding the use of assumption tests to simplify the statistical analyses used in this exercise.

Answers to Practice Problems

PRACTICE PROBLEM #1

$p = 0.11$	Reject or **Do Not Reject** Ho?
$p = 0.56$	Reject or **Do Not Reject** Ho?
$p = 0.99$	Reject or **Do Not Reject** Ho?
$p = 0.01$	**Reject** or Do Not Reject Ho?
$p < 0.005$	**Reject** or Do Not Reject Ho?

Sample Problems
The link below will take you to a PDF file with sample problems. Complete the problems assigned by your instructor. If none were assigned, complete problems #1-4.

Sample Problems
http://esa21.kennesaw.edu/activities/stats/problems.pdf

ESA 21: Environmental Science Activities

Activity Sheet
Measurement

Name:

Instructor:

Measurement:
Table 1: Measurements and Calculations Made on Three Blocks

Block Shape	Length (cm)	Width (cm)	Height (cm)	Volume (ml or cc)	Mass (g)	Density (g/cc)

Calculating the Volume of Irregular Objects:
Table 2: Calculating the Volume of Pebbles

Object	Initial Volume	New Volume (ml or cc)	Volume of Pebble (ml or cc)
Pebble #1	15 ml		
Pebble #2	15 ml		
Pebble #3	15 ml		

Inferential statistics:
List the null hypothesis of the study, and fill in the table for the t-test results. Complete the problems assigned by your instructor – tables for four problems have been provided.

Problem #:

H_o:

t-statistic	p-value	Do you reject the H_o ?

Problem #:

H_o:

t-statistic	p-value	Do you reject the H_o ?

Problem #:

H_o:

t-statistic	p-value	Do you reject the H_o ?

Problem #:

H_o:

t-statistic	p-value	Do you reject the H_o ?