NOTE: The examples given in this document are for Excel 2007 for Windows. If you are using Excel 2008 for Mac, you should read this document, but refer to the Mac examples that can be downloaded from the Resources section of the Biology web site at:

biology.sewanee.edu/resources/student-guide-to-excellence-in-biology/

Table of Contents:

Scientific Method, Hypotheses, and Experimental Design

Variables and Descriptive Statistics
   Using Excel for Descriptive Statistics

Graphing
   Using Excel to make Scatterplots and Line Graphs
   Using Excel to make Column Graphs
   Using Excel to make Histograms

Analyzing Data and Significance Tests
   Using Excel for T-tests
   Using Excel for Linear Regression

Thirteen Rules for Graphing, Data Analysis, and Reporting Results

Adding-in the Data Analysis module to Excel 2007 for Windows
The Scientific Method

The scientific method enables the progression of scientific knowledge and distinguishes scientific results from casual observation. Although this handout will refer to “the scientist” when explaining the scientific method, the reader should be clear that the method is an approach to knowledge that is available to all and one which they will use through the course of the semester. The scientific method begins with the observation of pattern in the world, which leads the scientist to develop questions about these observations. For example, an individual observes that a certain species of maple occurs in coves, but not on ridges or plateau tops. This leads to the question: Why is this species of maple only found in coves and not on ridges or plateau tops? Based on knowledge gained from previous scientific studies, the individual then formulates a hypothesis about the pattern. Hypotheses differ from questions in that they are testable statements about the observed pattern. For example: I hypothesize that *Acer saccharum* (sugar maple) is more abundant in coves (elevation 1000-1200m) than it is on ridges and plateau tops (elevation 1500-1800m).

A properly worded hypothesis makes testable predictions about the apparent patterns. These predictions are tested through manipulative experiments or field observations (also called natural experiments). The results from these studies are then analyzed to determine whether they support or reject the stated hypothesis. As each individual’s work builds upon the knowledge of previous researchers, the scientific method is self-correcting. If one person through accident or intent comes to invalid conclusions on the cause of a given pattern, subsequent researchers which utilize these conclusions as a foundation for their hypotheses will find their hypotheses not to be supported and will in turn look back to the research their conclusions were based upon for explanation.

Hypotheses that have been repeatedly tested and have not yet been rejected are sometimes called theories; probably the most familiar of these is the Theory of Evolution.
HYPOTHESES, DATA, STATISTICS, AND SIGNIFICANCE TESTS

Scientific use of the word **theory** varies greatly from its use in day-to-day conversation. Many people use the word theory to refer to something based upon personal opinion and/or experience. A *scientific* theory on the other hand is a hypothesis that has been extensively validated by research in many fields and provides the best explanation of experimental results. A theory cannot be proven, but it does provide the best current explanation of the data in-hand.

**Hypotheses**

As stated above, hypotheses differ from questions in that they are testable statements about the nature of the observed pattern. A hypothesis may be a testable statement of the observed pattern or it may seek to elucidate the underlying cause of the observed pattern. An example of the former: I hypothesize that *Acer saccharum* (sugar maple) is more abundant in coves (elevation 1000-1200m) than it is on ridges and plateau tops (elevation 1500-1800m). Example of the latter: I hypothesize that *Acer saccharum* (sugar maple) abundance is positively correlated with the concentration of Nitrogen and Phosphorous in the soil (as the concentration of N and P increases, *Acer saccharum* abundance increases).

In developing a hypothesis, it is very important that it accomplishes the following:
1. Identify the 2 variables being examined (in the above examples the variables were *Acer* abundance and elevation (ex. #1) and *Acer* abundance and soil nutrients (ex. #2).
2. Identify the expected relationship between these variables (ex #1: *Acer saccharum* abundance is greater on in one category (coves) than in the other category (plateaus); ex #2: *Acer* abundance increasing as the N concentration increases).
3. Hypotheses need to be as explicit as possible about the variables being tested.

**Further Examples:**

"Applications of nitrogen fertilizer increase the growth rate of tomatoes" is a testable hypothesis, but it is not clear whether the relationship between the volume of fertilizer application and tomato growth is positive, or how tomato growth is to be measured. Precisely worded hypotheses generally lead to better experiments than vague hypotheses.

"The drug ‘Sniffle-Away’ combats the common cold". Again, this is too general and doesn’t help focus our minds on (a) what hypothesis we are testing, and (b) what variables are being measured, and (c) the expected relationship between the variables is not clear and explicit.

"Adults who take one gram per day of the drug ‘Sniffle-Away’ experience a shorter duration of cold symptoms than those who do not receive the drug.” Notice that this statement is very specific -- it tells us that the hypothesis relates to duration of symptoms, that those adults receiving the drug are being compared with those that do not, and it tells us what the expected relationship is between these two groups.
Experimental Design

Once you have a precisely worded hypothesis, you can move on to designing a study to test this hypothesis. Manipulative and observational studies both have two types of variables: dependent and independent variables. Variables are the things we measure, control or manipulate. The dependent variable is the one that you expect to see a response in and the independent variable is one that causes the response. Thus, we have cause (independent variable) and effect (dependent variable). In the “Sniffle-away” example, the dependent variable is the time an individual has symptoms and the independent variable is the drug treatment (did the patient take the drug or not). In the maple examples, the dependent variable is the abundance of maples and the independent variables are elevation and soil N and P concentrations. You can also think of the relationship between these 2 variables as the value of your dependent variable being determined as a function of the value of your independent variable. This should remind you of y is a function of x statements from math classes: $y = f(x)$. Thinking of things in this way will also help you remember which variable goes on which axis when you graph your data (see Graphing handout).

What would be the dependent and independent variables in studies that measured:
1. Carbon dioxide concentrations in the atmosphere over time
2. The height of males and females in the Philippines
3. Fish swimming speed changed at different water temperatures

Once you’ve formulated your hypothesis, your variables should be self-evident. What remains is designing an appropriate study to measure them. In settling on a design, there are a number of important considerations to remember in all studies (elaborated below):
1. Change one variable at a time
2. Collect enough data (large enough sample size)
3. Beware of sampling bias

Change one variable at a time

Unlike studies in physics and chemistry, the results of ecological studies are often very noisy- that is there is often a great deal of variation in your dependent variable that remains unexplained by your independent variable. This is primarily due to the great complexity of ecological systems. For example, the abundance of a plant at a given location is likely due not only to soil nutrient concentrations, but also sunlight, rainfall, winter and summer temperatures, and past disturbances like tree falls. Each of these has an effect and each varies in their effect at different spatial scales- sunlight availability and deer browse effects change on the scale of several meters due the changing thickness of the canopy and the uneven foraging of deer, and rainfall and temperature changes on the scale of kilometers.
A well designed study seeks to isolate the effect of one of these variables at a time in order to determine the effect of each independently; failure to do so may result in the erroneous assumption that a given independent variable has no effect on the dependent variable. Thus if we are interested in determining whether there is a greater abundance of trilliums (a spring wildflower) in the coves than on the plateau top, we would sample a number of locations (plots) in each habitat while doing our best to keep other variables that might affect trillium abundance, like proximity to streams, trails, and light level, as constant as possible. As ecological systems are complex and often multiple variables will change at the same time (light levels will be different in coves than they are on the plateau top because of the thicker canopy), it is often not possible to hold all variables other than the one of interest constant, but it is something to remember when designing your study.

In manipulative experiments, this concern is addressed in a slightly different manner. In any experiment you need to have a standard with which to judge your results. This helps you to determine whether any other factors might have caused the results of the experiment. For example, if we were testing the ‘Sniffle-Away’ drug we would want to set up our experiment so that one group of people with colds would take the drug, while another group, also with colds, would not. We could then compare the recovery rate of the two groups. The group that is subjected to the manipulation (in this case, the drug) is often called the experimental or treatment group, while the others are termed controls. The control must be treated in exactly the same way as the treatment, except for the factor that we are investigating. In the drug example, the control people would take a “dud” pill (called a placebo) and be told the same things as the people in the treatment group so that they would not know they had not received the drug. If the controls in our drug trial had not been given placebo pills, any difference between the treatment and the control could be attributed to the psychological effects of taking a pill, rather than the drug itself.

A complementary way of reducing the chances that our treatment and control groups differ at the outset of the experiment is to use a technique called pairing. As the name suggests, this method involves placing all the individual units in the experiment into matched pairs. One member of each pair gets the treatment, the other the control. We might, for example, pair people in our cold experiment according to age, sex and where they live (e.g., one pair would be two 20-year-old women from Alaska, another would be two 50-year-old men from Tennessee). Pairing therefore offers a way of reducing any initial differences between our treatment and control groups, thus making us more confident that the results of our experiment were caused by the factor we are investigating. In some paired experiments we don't need "control groups". For example, if we were testing the hypothesis that drinking Coke increases heart rate in humans, we might measure peoples' heart rates, give them a can of Coke, then re-measure the heart rates. We would then compare the heart rates before and after. In this case, there is no control group who never have a drink of Coke -- in a sense the "control" is built into the before/after nature of the experiment. Note that you could design a paired experiment to test this hypothesis. The control group would be given water and the treatment group would be given Coke. Can you think of any advantages or disadvantages of the before-after design compared to the paired design?
The importance of sample size

Imagine flipping a coin 2 times to determine whether the coin had an equal probability of landing on heads or tails. Even if our coin is evenly balanced, we would expect that if we did this 2-flip test 100 times, that only 50 of those tests would give us one head and one tail while 25 times we would get 2 heads and 25 times we would get 2 tails. On the other hand if we flipped that balanced coin 100 times in each test and did this test 100 times, the majority of our tests would have an almost equal number of heads and tails because the odds of getting the same result over and over and over again are very small.

The same principle applies to field studies. If I hypothesized that Bloodroots (another spring wildflower) would be more abundant within 10m of streams than they are 50m away from them, and I only measured their abundance in 2 plots (one within 10m and the other 50m away), I would not be able to make any real conclusions. With only one sample in each location, by random chance I may find fewer Bloodroots next to the stream than I do farther away even if across a larger area they really are more common close to streams. This is because some locations next to the stream will have lower Bloodroot abundance than some locations away from it due to chance alone (sometimes you will get 2 heads in a row). To use a manipulative example, if I tested a new cold drug on just one person, with another person as a control (untreated) and asked whether the drug had an effect, the results would not be very convincing! Any difference between our treatment and our control could be due to all sorts of differences between the two people that had nothing to do with the drug.

In order to get around this problem we try to collect observational data from many samples and perform manipulative experiments on larger numbers of experimental subjects. If we had 200 plots or people in our experiment, instead of just two, there would be a much better chance that the differences detected in our study were due to the effect of the stream or drug. We would have to be very unlucky indeed if all 200 people that we assigned to the control treatment all happened to naturally suffer from particularly bad colds. This example also illustrates the importance of randomly assigning experimental subjects to treatment or control group -- we would want to use a coin toss, or some other unbiased method to assign people to experimental groups. This randomization helps ensure that there are no differences between the treatment and control groups before we even start the experiment.

Removing Biases

One additional factor to think about when planning an experiment is observer bias. Psychological studies have shown that if we are expecting a certain outcome from an experiment, then we can subconsciously bias our experiment towards this outcome. If we actively prefer a particular outcome, the potential for bias is even worse (e.g., if we really want to find a difference between our treatment and control groups). These biases can occur both when we set up the experiment and when we gather the data. In our drug example, we might tend to choose more ‘sickly looking’ people for the control group than the treatment group. When we were interviewing subjects about how quickly they got better (gathering the data) we might tend to interpret the answers of the two groups differently.

There are many solutions to the problem of observer bias, but no “best” method. With observational studies (most ecological studies fall into this category), we can determine beforehand where and how to locate and measure samples to avoid biases. Samples can be located at predetermined set distance intervals (ex. every 50m) or at random distances intervals.
HYPOTHESES, DATA, STATISTICS, AND SIGNIFICANCE TESTS

along a line running through your study area (ex. Plots were located at random intervals along 2 transects (one in the cove forest, one on top of the plateau) which began at a random starting point and continued at a constant elevation). Or plots might occur at certain points in the landscape (paired plots located in the stream 1, 10, and 20m above and below where it was intersected by the trail). With manipulative experiments, we can use rigorously impartial methods for assigning subjects to control or treatment groups: using a coin toss or random numbers table, for example. We can also gather our data in a manner that is ‘blind’ to which treatment we are dealing with. Regardless of our approach, we must have carefully thought out criteria for gathering our data. The criteria should leave no room for subjective ‘judgment-calls’. Remember: the human mind is a sneaky thing to deal with, so be wary!

Although there is no “best” way to organize data collection, our ability to draw conclusions related to our hypotheses is dependent upon the quality of our data. It is therefore essential that you spend time thinking your way through your sampling methodology. Time spent on this will be well worth it.

We have just been through quite a long list of factors that go into a well-designed experiment. Often our experimental subjects, or the resources that we have available, prevent us from doing a perfectly controlled experiment with a huge sample size. That’s OK, but it is important to acknowledge any weaknesses in our experimental design when we discuss our results and we should be careful to temper our conclusions accordingly.
HYPOTHESES, DATA, STATISTICS, AND SIGNIFICANCE TESTS

Variable Types

Variables (both dependent and independent) come in three main types:

I. Continuous data:
These are numbers that can take on any value (i.e. fractions are allowed). For example:
  - Human height measured with a ruler
  - Temperature measured with a thermometer
  - Diameter of hickory nuts measured with calipers
  - Proportion of time a squirrel spends feeding

II. Discrete data:
These are all non-continuous data. There are two main classes:

(i) Categorical data:
These occur in non-overlapping, distinct categories or classes. These categories are not usually numerical. For example:
  - The color of bird feathers classified into red, blue or brown
  - Sex: male or female
  - A bird nest classified as either having successfully produced young birds or not.

(ii) Counts
This type of data is numerical, but not continuous. For example:
  - The number of acorns found in a square meter of forest. Here we cannot have “half acorns”, so the data are not continuous.
  - The number of birds seen on a census

Descriptive statistics

Descriptive statistics convey information about data in an easy to understand and digestible way. We encounter descriptive statistics everyday: Grades and GPA’s are supposed to summarize the results of countless tests and papers, stock market indices (e.g., the Dow Jones Industrial Average) summarize the performance of some aspect of the stock market, “last frost” dates are used by farmers and gardeners to plan their planting, newspapers publish “crime statistics” which summarize criminal activity in our communities. All these summaries are descriptive statistics. Later handouts will cover inferential statistics or "statistical tests" which are designed to test for differences between groups, or look for correlations between variables.

There are many methods of summarizing data, and our choice of method will affect the message that we convey. Both producers and consumers of descriptive statistics should know how the choice of descriptive statistic affects this “message”. Producers of information should be aware so that they can choose the most appropriate (or deceptive, if you are in the spin-doctoring trade!) method. Consumers of information need to be aware so that they can evaluate the information that is given to them. For example, if a college claims that its “average class size is 12”, the question that should jump to the front of your mind is: what kind of average? There
are three kinds of averages that can be use with continuous numbers and counts, and each tells you different things about your data. It is very important when you describe your data that you identify which type of average you are using. These different averages are calculated as follows:

**Mean:** To calculate, add up all the observations and divide by the sample size (number of observations). The mean gives a measure of the data’s “center of gravity” and can be used with continuous or count data. The mean is the average most commonly used in biology.

**Median:** To calculate, line up the numbers from biggest to smallest (i.e. put them in rank). The number in the *middle* is the median. The median is also called the 50th percentile. You can use this with continuous or count data. The median is sometimes used instead of or in conjunction with the mean when the data have unusual outliers or other extremes that have a disproportionate affect on the mean.

**Mode:** This is the most common number (i.e., the number that is most often repeated in the data). This statistic is seldom used in biology.

Another important group of descriptive statistics measures the amount of variation in the data. **Variation** is the measure of the degree of spread in your data; a measure of the degree to which the values of your variables are grouped around the mean. There are three measures of variation that can be used with continuous or count data:

**Range:** This is the difference between the highest and lowest values in a sample

**Standard deviation (s.d.):** To get the s.d. we calculate the difference between the value of each data point and the mean of all the data points. The exact formula need not concern us here: the important thing to note is that the s.d. is larger in more variable datasets than in more uniform datasets. Specifically, the standard deviation is the value that when added and subtracted to your mean will give you the range of values between which 33% of your observations will be contained. For example, say you were recording the diameter at breast height (dbh) of white oaks at the cross. If your mean dbh was 24 and your standard deviation was 3, then 33% of your trees would have a dbh between 21 and 27.

**Standard error (s.e.):** This is the standard deviation divided by the square root of the sample size. This number is also a measure of variation. It is smaller than the standard deviation (and therefore looks better on figures!) and its value is closely related to the results of statistical tests. Therefore, *the standard error is the measure of variation most often used by biologists.*
Descriptive Statistics in Excel 2007 for Windows
(if you are using a Mac, see the Mac instructions instead)

To calculate Descriptive Statistics (including means and standard errors):

• Arrange your data in Excel into columns (example at right). All of the values that you want to analyze as a group need to be in a column together.

• In the Excel menu, choose ‘Data: Data Analysis’. A Data Analysis window will appear (Example below). NOTE: If you do not see the Data Analysis option, refer to the separate instructions on how to Add-in Data Analysis capability.

• From the Data Analysis window, select ‘Descriptive Statistics’. A Descriptive Statistics window will appear.

• In the Descriptive Statistics window: (a) Click in the “Input Range” box, then use your mouse to select the data you want summarized from your Excel sheet (this would be all of the values in your column of interest, (b) check “Summary statistics”, (c) check the “Labels in first row” if your first row of what you select has a label for the column, (d) click in the “Output Range” box, then use your mouse to select a cell in your Excel sheet – this is where Excel will write its results. See the example below.
HYPOTHESES, DATA, STATISTICS, AND SIGNIFICANCE TESTS

- Click ‘OK’ and examine the results (example below)

- Excel will produce a table like the one shown here. The mean and standard error are in the first two rows. You will also see a variety of other statistics: the standard deviation, median, minimum, maximum, range, and count (= number of observations or sample size).

- Some useful Excel shortcuts:
  - You can calculate Descriptive Statistics for more than one set of data at a time. To do so, arrange your samples in adjacent columns and then select multiple columns as your ‘Input Range’. If you do this it helps to label each column to help keep track of the output.
  - **To calculate a mean**: type “=AVERAGE()” then place your cursor in between the parentheses and use the mouse to select the data you wish to use, then hit “return”.
  - **To calculate a standard deviation**: type “=STDEV()” then place your mouse in between the parentheses, then select the data you wish to use, then hit “return”.
  - **There is no shortcut for standard error in Excel**. You have to use the Descriptive Statistics tool.
Graphing Results

Most scientific papers use graphs (generally referred to as figures) to communicate their results. Because of our visual nature, figures can show patterns in your data more easily than tables. A clear figure can save many paragraphs of text. Figures do not analyze your data or determine whether your hypothesis is supported – they simply illustrate your data. Significance tests are necessary to determine whether there is indeed a significant relationship between your independent and dependent variables. Nonetheless, figures are essential companions to any paper or statistical analysis.

Some general points about figures:

- Label each axis and include the units of all measurements.
- Give the figure a figure number and a descriptive legend that is placed below the graph.
- Scientific journals do not publish 3-D figures unless there really are three axes. 3-dimensional graphs may look fancy, but they are hard to read. Use 2-D figures without shadows.
- The dependent variable goes on the y axis and the independent variable goes on the x axis. Remember: y as a function of x.
- A figure should be interpretable without having to refer back to the text. The figure legend should tell the reader what was measured, where and when. It should also include results from statistical analyses and relevant significance tests (see the ‘Analyzing Data’ portion of this document).

![Figure 1: Graphs should include a figure legend placed below the graph. The legend includes information such as the sample size used, the location of the study, and the species being studied. The legend’s job is to help the reader interpret the figure.](image-url)
HYPOTHESES, DATA, STATISTICS, AND SIGNIFICANCE TESTS

The four most common Figure types

There are 4 basic types of Figures that you may use to illustrate your data. The type of Figure utilized is determined by what sort of data you have and what you are trying to show.

<table>
<thead>
<tr>
<th>Figure type</th>
<th>When to use this type:</th>
</tr>
</thead>
<tbody>
<tr>
<td>Scatter plot</td>
<td>For continuous or count data. It shows relationship between two sets of numbers.</td>
</tr>
<tr>
<td>Line graph</td>
<td>For continuous or count data when we want to look at the details of &quot;ups and downs&quot; in the data. You typically only use this when illustrating repeated measures of the same thing (like flow volume in a stream).</td>
</tr>
<tr>
<td>Column graph</td>
<td>Used when data is arranged in categories and you wish to present a summary (mean and standard deviation or error) for each category.</td>
</tr>
<tr>
<td>Histogram</td>
<td>Displays your data’s distribution. Histograms illustrate the data mean, spread (standard error and deviation), and distribution.</td>
</tr>
</tbody>
</table>

Scatter plot

Scatter plots are usually used when we want to show the relationship between two variables that are made up of continuous or count data. We will also use this type of graph when we are illustrating a linear regression between two variables (see the ‘Analyzing Data’ portion of this handout).

For example:

Note: “(n = 8)” means “the sample size was eight individuals” and use of unit abbreviations is OK in the axes labels.

Figure 2: Swimming speed of ten-day-old green frog (Lithobates clamitans) tadpoles (n = 8) plotted against water temperature of lakes from which the tadpoles were sampled

When using this type of Figure to illustrate a linear regression, we will draw a “best-fit” line through the points, specifically the “best-fit straight line”. Be careful though-- only use “best-fit straight lines” when you have good reason to expect a linear relationship. A common mistake is drawing straight lines through data when the relationship between variables is not linear. To see how to fit and interpret this line, see the statistical tests handouts.
HYPOTHESES, DATA, STATISTICS, AND SIGNIFICANCE TESTS

Line graphs

These are essentially scatter plots with the points connected. Line graphs connect all the points, so this type of graph is only suitable when we are interested in the fine details of how our data moves up and down. Line graphs also only work when there is just one y-value (dependent variable) associated with each x-value (independent variable). Typically you only use line graphs when showing repeated measurements of the same thing. In the example below, these are repeated measurements of the atmosphere.

For example:

![Line graph example](image)

**Figure 3: Atmospheric carbon dioxide concentration measured at the Mauna Loa Observatory, Hawaii from 1958 to 2007**

Column graph

Column graphs are used when the independent variable (x axis) is divided into categories. These categories represent the only possible values for that variable. The dependent variable (y axis) is usually continuous though it could also be discrete counts (ex. # of individuals). In the example below we have graphed the heights of ten trees from each of five species within a given sampling area. The height of the columns indicates the mean height of each tree species (the mean height of sampled white oaks is 20m). The “whisker bars” (capital ‘I’ shaped lines at the top of the column) indicate the standard error. The horizontal lines on the error whiskers show the values you get when you add and subtract the mean by the standard error. Shorter whiskers indicate that there is a lower standard error value which results from most trees being close to the same height (see the histogram plot for a further description of standard error). The larger the error value, the longer the whisker. It is important to note that while the graph is showing the
mean heights of these species, the y axis is a range of heights, NOT a range of mean heights and is labeled accordingly.

For example:

![Histogram of tree heights](image)

**Figure 4: Mean and standard error of the height of five tree species (n=10 individuals per species) measured in July 2005 at Green’s View, Sewanee, TN**

**Histograms**

Histograms are used to investigate the distribution of your data in a more detailed manner than is possible when looking solely at the summary statistics (mean, standard error) or at a column graph. With a histogram we visually see the median value within our data (number in the middle of the range of values), the mode (the number that occurs most frequently in your data), and we can often tell approximately what the mean is going to be as well.
HYPOTHESES, DATA, STATISTICS, AND SIGNIFICANCE TESTS

Histograms also show us how the data are distributed. There are two aspects of data distribution that are important to us: the degree to which the data is dispersed and whether or not the distribution is skewed; both of which are described below. How the data are distributed has implications for the processes impacting the variable you are graphing.

Data dispersal

Compare for a moment the two following histograms (legends omitted).

Both data sets have the same mode, median, and mean. However, the left graph has a tighter distribution of values than the right one; the data are less dispersed. We measure this through the standard deviation and standard error. So the left data set would have the same mean as the right one, but a smaller standard error and deviation.

Data Skewedness

Data distributed like that shown in the above histograms is said to be “normally” distributed – the median, mode, and mean are the same or almost exactly the same. This is also referred to as a bell curve (for its shape like a bell) and is a common distribution for many types of data. Human height for example; the mean height for American females is 5’ 3.8”. The
overall distribution of heights approximately follows a normal distribution; the number of people 1” taller than the mean is approximately equal to the number which are 1” shorter than the mean, the number of people 2” shorter is approximately equal to the number 2” taller, etc.

However, not all data are distributed in this manner; some data sets are left or right skewed as we see in the below examples. The figure on the left is right-skewed and the figure on the right is left-skewed. **Note that the direction of the skew is to the side that has the longer tail.**
Creating each Figure Type in Excel 2007 for Windows

Scatter plots and Line graphs

- For both of these graphs you need to have your data arranged in 2 columns. Each row is an individual observation and the two columns are the values of your dependent and independent variables. Excel will use the left hand column as your independent variable (x axis), so it will be easier to have your data arranged accordingly (example at right).

- Click and drag your mouse to select the data you want to use in your graph (example at right).

- To create a graph, select ‘Insert’ on the Excel menu and then select the type of graph you want: either ‘Line’ or ‘Scatter’ (example below)

- A first draft of your graph should appear. The example below is of a scatter plot.

- To format your chart, click on your chart, then choose ‘Layout’ from the Excel menu. You can label your axes using ‘Axis titles’. Use ‘Gridlines’ to remove them as well. Use ‘Legend’ to remove the legend on the right.

- Move the chart title from above the figure to below it (click and drag the title and the graph where you want them), and write a descriptive title. Alternatively you can delete the chart title and copy your entire figure into a Word document before adding your Figure legend below the Figure.
• See the example below for a finished scatterplot.

To add a linear trendline line to a scatter plot

• Click on your scatter plot, then go to ‘Layout: Analysis: Trendline: Linear Trendline’. To add the R-squared value and the equation of the line to the chart, select ‘Layout: Analysis: Trendline: More Trendline Options’ and then check the boxes for ‘Display Equation on Chart’ and ‘Display R-squared value on chart’. Then click and drag the equation and R-squared value wherever you want them on your graph. NOTE: if you include a linear trendline in your figure, you should note that in the figure legend (see example below).
HYPOTHESES, DATA, STATISTICS, AND SIGNIFICANCE TESTS

Column graphs

- Use the “Descriptive Statistics” tool within the Data Analysis menu described in the descriptive statistics handout to calculate the mean and standard error of each group.

- Once you’ve done this, rewrite them in a small table in Excel as shown.

<table>
<thead>
<tr>
<th></th>
<th>Mean</th>
<th>Standard error</th>
</tr>
</thead>
<tbody>
<tr>
<td>White Oak</td>
<td>34.6</td>
<td>7.9</td>
</tr>
<tr>
<td>Chestnut Oak</td>
<td>23.6</td>
<td>5.5</td>
</tr>
</tbody>
</table>

- Then, in your table select the category names (e.g., “white oak”, “chestnut oak”) and the values for the means (not the values for the standard errors) as shown in the table above.

- From the Excel menu, select ‘Insert: Column: 2-D column’. A first draft of your chart will appear (example at left). Highlight the chart by clicking on it once, and then use the Layout menu to modify the chart title, axis titles, legend and gridlines.

- A final version of a column graph could look like this:

Adding standard error bars to your column graph

- Once you have formatted the graph to look the way you want, Click once on the bars in your graph. From the Excel menu, select ‘Layout: Analysis: Error Bars: More error bars options’. A ‘Format Error Bars’ window will appear (example at right).
• From the 'Format Error Bars' window, select 'Both' for the Direction. For Error Amount, select 'Custom' and then hit the 'Specify Value' button (you are going to tell Excel what the standard error values actually are for your data). A 'Custom Error Value' window will appear.

• You need to fill both the Positive and Negative Error Value boxes with the standard error values you calculated earlier, and entered into your Excel sheet. To do this, first empty the contents of the Positive Error Value window, leaving your cursor in the box, and then select BOTH of your s.e. values from your Excel sheet. Repeat this for the Negative error value box (again selecting BOTH s.e. values (example at right). Click OK.

• Your standard error bars should now be on your graph, and they should match the sizes you specified. NOTE: On any figure that has s.e. bars, you should indicate this in the figure legend as shown below.
**HYPOTHESES, DATA, STATISTICS, AND SIGNIFICANCE TESTS**

**Histograms**

- A histogram divides your data into categories (Excel calls these 'bins') and displays the number of observations in each category (bin). For example if you have a list of people's ages, you could break them into categories like 0-10 yrs old, 10-20, 20-30..... The appropriate number and size range of categories will depend upon the data set. To start to get a handle on this, use the Descriptive Statistics tool to find the minimum and maximum values for your data. From here, you want to divide your data so that you do not have too many or too few categories and each contains the same range of values. For instance, if my data had a range from 0-100 I would not want categories of 50 (0-50, 50-100) because the categories would be too large to tell me anything useful about the spread of my data. Similarly I would not want to have categories with a range of 5 because that would result in too many categories and I might not have enough data in each.

- Once you’ve settled on the size range of your bins, you need to create a new column in your data table that will tell Excel what categories you want. The values in this column (labeled “bins” in the example at left), tell Excel the top value of the category (so the category “1” ranges from 0-1, the category “2” ranges from 1.1-2.0).

- From the Excel menu, select ‘Data: Data Analysis’ and then ‘Histogram’ from the window that appears. Put your data values into the ‘Input Range’ box, Put your bins into the ‘Bin Range’ box, check the ‘Output Range’ button and tell Excel where you want it to put the results by selecting a cell from your Excel worksheet (example below).
HYPOTHESES, DATA, STATISTICS, AND SIGNIFICANCE TESTS

- Click ‘OK’. Excel will give you a new data table of results (example at right).

- Before making your figure, you should re-label your Bins so that they clearly indicate what range of value they cover (example below).

<table>
<thead>
<tr>
<th>Bins</th>
<th>Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>0-1</td>
<td>1</td>
</tr>
<tr>
<td>1.1-2</td>
<td>9</td>
</tr>
<tr>
<td>2.1-3</td>
<td>19</td>
</tr>
<tr>
<td>3.1-4</td>
<td>3</td>
</tr>
<tr>
<td>4.1-5</td>
<td>2</td>
</tr>
<tr>
<td>More</td>
<td>0</td>
</tr>
</tbody>
</table>

- On the new data table, select the values in the ‘Bins’ and ‘Frequency’ columns. Do not include the ‘More’ row which will have a frequency of 0 if your bins covered all of your data points. From the Excel menu select ‘Insert: Column: 2-D Column’. A first draft of your figure will appear.

- Click on the graph once and use the ‘Layout’ options to label your axes, remove the legend (‘Series1’) and gridlines, and add a descriptive legend below the figure. An example of a finished histogram is below.
Taking your figures beyond “generic Excel”

The following are two figures from Excel. They present the same data, but the second is better than the first for the following reasons:

1. The figure legend is complete and is correctly positioned.
2. The “legend” on the right has been removed (simply click on it and delete).
3. The axes are correctly labeled and they include units of measurement.
4. The “gridlines” have been removed – these are unnecessary.
5. The gray background and blue bar color have been replaced with a clear background and simple black lines for the bars.
6. The “shadow” effect has been removed from the bars.

**Not good:**

![Figure 1: Graph of oak height](image)

**Better:**

![Figure 1: Mean and standard error of the heights of white oak and chestnut oak (n = 8 trees per species) measured at Green's View, Sewanee, 23rd July 2006](image)

You can adjust your figures by double clicking on the various parts of the figure to change them. To remove something entirely, hit the Delete key when you’ve got that portion highlighted, but make sure that you have the correct thing highlighted. Also remember that most changes can be undone by clicking the Undo arrow along the top toolbar.
Introduction to Significance tests

You have already been introduced to hypotheses and how to create figures which visually represent the trends in your data. However in science it is not sufficient to look at a figure and say that there appears to be a trend in the data and we therefore accept/reject our hypothesis; Statistical tests of significance are necessary to make that determination. These tests work in a way that is not initially intuitive; rather than seeking to accept your hypothesis, they assess whether your data are significantly different from what you’d expect if there was no relationship between the variables.

The hypotheses that we covered in the first handout were all what are called alternative hypotheses (designated by the abbreviation H_A); that is they were all hypotheses that suggested that there would be a relationship between two variables. There is another type of hypothesis called a null hypothesis (designated H_O). Every alternative hypothesis has a correlating null hypothesis that there is no relationship between two variables, and it is this null hypothesis that is tested and then accepted or rejected through the statistical analyses.

For example:
H_A: I hypothesize that *Acer saccharum* (sugar maple) is more abundant in coves (elevation 1000-1200m) than it is on the plateau top (elevation 1500-1800m).

H_O: There is no difference in the abundance of *Acer saccharum* (sugar maple) in coves (1000-1200m) and on the plateau top (1500-1800m).

In science, a difference between groups or a trend is considered significant only if the data are statistically different from the null. Statistical tests provide us with a quantitative evaluation of the probability that the pattern we see in our data is due to chance alone.

There are many different types of statistical tests. The specific test you use depends on the question you are asking and the type of data that you have. In this document, we describe two common statistical tests: a t-test and linear regression. A t-test compares the means of two categories of data, assessing whether the means significantly differ from one and other (e.g., do sugar maples produce more seeds than red maples?). A linear regression looks at whether there is an association or trend between two continuous variables (e.g., as the size of trees increases, does the number of seeds they produce also increase?).

Most significance tests (and all of the ones we will use here) will provide you with a probability (p) value. The p-value is the likelihood that you are rejecting the null hypothesis when in fact it is true. Another way of thinking about the p-value is that it is a measure of the probability that the pattern we see in our data is due to chance alone. Most commonly in science we set our level of acceptance of this error (rejecting the null hypothesis when in fact it is true) at 5% or less. Therefore, if the p-value is less than or equal to 0.05 (=5%) we reject H_O in favor of H_A; we conclude that the difference between the two means is statistically significant. Another way of thinking about this is to say that if p < 0.05, there is less than a 5% chance that the pattern observed is due to chance alone. Thus there is less than a 5% chance that we are erroneously rejecting the null hypothesis.
Conducting a T-Test in Excel

A t-test is used to compare the means of two samples. To run a t-test on your data:

- Arrange the data from your samples in separate columns (example at right). With these data we can test whether men or women are, on average, taller.

- From the Excel menu, click on ‘Data: Data Analysis’. A ‘Data Analysis’ window will appear. (If ‘Data Analysis’ is not an option, check the handout about Adding-in the Data Analysis package.)

- Select ‘t-Test: Two-Sample Assuming Equal Variances’. A t-Test window will appear. Highlight the cells for your first sample for ‘Variable 1 Range’, and the cells of your second for ‘Variable 2 Range’. Check the ‘Labels’ box if you also highlighted your column labels. Check the ‘Output Range’ box to tell Excel where to put your results (example below).

- The analysis results will report the number of observations (= sample size) and mean value for each of your categories. The statistical analysis reports a two-tailed p-value (outlined below), which is what you should report and use to determine whether there is a statistically significant difference between the mean values for your two categories.

- If your p-level is less than 0.05, you found a statistically significant difference between the mean values of your categories. In this case, we found a statistically significant difference between mean male and mean female height.

- This could be reported as: “We found that males are significantly taller than females (t-test, two-tailed p-level = 0.00056). Mean male height was 174.1 cm (n=15), and mean female height was 160.8 cm (n=13).’
Linear Regression

• The purpose of linear regression is to determine if there is a significant linear relationship between an independent variable (X) and a dependent variable (Y). That is, does the value of Y vary in a predictable way as X changes. If so, then the independent variable (X) can be used to predict values of Y.

• The regression calculates a “best fit” line through your plotted data. This is termed a trendline in Excel. The Graphing handout describes how to add a trend line to a scatter plot.

• A trendline with a positive slope indicates that as the value of X increases, so does the value of Y. This is also known as a positive correlation. Conversely, a negatively sloping line indicates that as X increases, Y decreases; a negative correlation. With a regression, your p-value is called “significance F”. Different significance tests use different terms for their p-values (like significance F), but they all mean the same thing.

• To run a regression analysis, first arrange your data in two columns. Each row should present the measurements from an individual sample (example at right for a sample of height and weight from 15 individuals).

• Select ‘Data: Data Analysis’ and then ‘Regression’ from the Data Analysis window. Input your X- and Y- data values. Notice that it asks for your dependent variable (‘Input Y Range’) first. If you highlighted your column labels be sure to check the ‘Labels’ box.

• The results present two values that are of particular importance for interpreting your results: the significance F and the R Square (R²). These are both boxed in the screenshot below. The significance F is a p-value that indicates the probability that the slope of the best fit line is different from zero. If the significance F value is \( \leq 0.05 \), then there is a significant linear relationship between X and Y. Describing the type of relationship will depend on the slope of the line (is it a positive or a negative relationship?)
HYPOTHESES, DATA, STATISTICS, AND SIGNIFICANCE TESTS

- The R Square (or R²) value is the proportion of the variation in your dependent variable that is explained by variation in the independent variable. In other words the R² gives an indication of how well the straight line through your data actually describes the relationship between X and Y. R² values range from 0 and 1. An R² of 1.0 indicates a “perfect fit” (i.e. every data point falls exactly on the best fit line), which almost never happens. Low R² values indicate that your data points are scattered around the line. Keep in mind that the R² value does not tell you anything about whether or not there is a linear relationship between X and Y – that depends on the significance F value. Both statistics (R² and significance F) are important in describing the relationship between X and Y, but they tell you different things.

- In this case, we found a statistically significant relationship between height and weight. If you make a scatterplot of these data you can tell that as height increases, weight does as well.

- This result could be reported as: “Height was significantly related to weight (linear regression, p-level = 0.0021). As height increased, so did weight (R Square = 0.665, n=15).”
HYPOTHESES, DATA, STATISTICS, AND SIGNIFICANCE TESTS

For Figures and Tables:

1. Figure legends go **below** Figures, Table legends go **above** Tables. Because Excel doesn’t give you much space for the legend, you may want to copy your Figure or Table into a Word document and then add the legend above or below it as is appropriate.

2. Figure and Table legends should be detailed enough that the Figure or Table needs no further explanation and is easily interpreted by a reader who has no idea how, when, or where you collected the data.

   • Wrong: “Figure 1: Tree sizes”

   • Right: “Figure 1: Distribution of Chestnut Oak diameters at breast height. Data collected at Green’s View in Sewanee, TN on October 15, 2017.”

3. If you have standard error bars or a trendline or any other sort of marking on your Figure you should state what it is in the legend.

   • Example: “Figure 2: Mean weight (+/- standard error) of Chestnut Oak acorns from cove and plateau habitats. Acorns collected in Sewanee, TN at Green’s View (plateau) and in Shakerag Hollow (cove) on November 4, 2023”

   • Example: “Figure 3: Relationship between White Oak acorn width and weight (with a linear trendline and the equation for that line) for acorns (n = 147) collected at Lake Cheston, Sewanee, TN on October 15, 2009.”

4. If you have statistical results you should include them in the figure legend.

   • Example: “Figure 2: Mean weight (+/- standard error) of Chestnut Oak acorns from cove and plateau habitats. Acorns from the cove (mean = 7.6 g, N = 43) weighed significantly more than those from the plateau (mean = 5.2 g, N = 61) (two-tailed t-test, \( P = 0.01 \)). Acorns collected in Sewanee, TN at Green’s View (plateau) and in Shakerag Hollow (cove) on November 4, 2023.”

   • Example: “Figure 3: Relationship between White Oak acorn width and weight (with a linear trendline and the equation for that line) for acorns (n = 147) collected at Lake Cheston, Sewanee, TN on October 15, 2009 (\( R^2 = 0.54 \), significance \( F = 0.23 \)).”

5. Include correct **units** on each axis.

For Data Analysis:

6. The basic approach: if your independent variable is **categorical** (also called ‘discrete’) (e.g. gender) then you will present your data with a bar graph and compare your categories with a t-test. If your independent variable is **continuous** (also called ‘linear’)

30
(e.g. height) then you will present your data as a scatterplot and compare your variables with a regression.

7. For a categorical hypothesis, be clear on whether your hypothesis is directional ("A will be greater than B") or not ("A will be different than B", which is true if A is greater than B or if A is less than B). For a directional hypothesis you should use the one-tailed P value. For a non-directional hypothesis, you should use the two-tailed P value.

8. The standard threshold value for statistical significance is 0.05. If your P value (from a t-test) or significance F value (from a regression) is less than 0.05, then you have found a statistically significant difference between your categories (for a t-test) or a statistically significant relationship between your variables (for a regression).

For Reporting Results:

9. The basic approach: use normal sentences, and report means and statistical results in parentheses.

10. Reporting results from a categorical hypothesis: include mean values (with units) for your categories, note whether you found a statistically significant difference (or not), and refer to your P value in parentheses at the end of the sentence. Also good to include are sample sizes, a reference to your Figure, and whether or not your supported your hypothesis.

   - Example: We found that Chestnut Oak acorns from the cove (mean = 7.6 g, N = 43) weighed significantly more than those from the plateau (mean = 5.2 g, N = 61) (t-test, P = 0.01, Figure 4). This result supported our hypothesis that Chestnut Oak acorns from the cove would be larger than those from the plateau.

11. Reporting results from a continuous hypothesis: indicate the variables being compared, whether or not they are significantly related (as determined by the ‘significance F’ value from a regression), the significance F value, and the $R^2$ value. Also valuable are your sample size, a reference to your Figure, whether or not you supported your hypothesis, and, if you found a significant relationship, an indication of what that relationship is (e.g. as X increases, Y decreases).

   - Example: We did not find a significant relationship between White Oak acorn width and weight ($N = 173$, regression, $F = 0.36$, $R^2 = 0.12$). This did not support our hypothesis that acorn weight would increase with width (Figure 5).

12. Remember that you should never claim to have ‘proven’ your hypothesis, only to have ‘supported’ or ‘failed to support’ it.

13. It makes no difference whether you support or fail to support your hypothesis as long as you analyze your data and report your results correctly.
HYPOTHESES, DATA, STATISTICS, AND SIGNIFICANCE TESTS

Adding-in the Data Analysis package to Excel 2007 for Windows

• If ‘Data Analysis’ is not an option under the ‘Data’ menu of Excel 2007 for Windows, you need to Add-in this capability to the program. To do so, click on the round Office button in the upper left corner of the Excel window, and choose ‘Excel Options’. An ‘Excel Options’ window will open (example below).

![Excel Options window](image1)

• From this window, click on ‘Add-Ins’ on the left and then ‘Go…’ next to ‘Manage Excel Add-ins’. The ‘Add-Ins’ window will appear (example at right). Click the box next to ‘Analysis ToolPak’ and then ‘OK’.

![Add-Ins window](image2)

• From the main Excel menu, click on ‘Data’ to confirm the ‘Data Analysis’ option is now present (example below). Occasionally you may need to close and re-open the Excel program for the Data Analysis package to appear.

![Data Analysis menu](image3)