

Accuracy, Precision and Statistics

Have you ever heard the old saying, “measure twice, cut once?” Carpenters know that they might make a measuring error, and it’s better to double-check a measurement *before* cutting a board to the wrong length!

You will be making many different kinds of measurements, and you should realize that any kind of measurement is subject to error. The error might be due to the limitations of the instrument, or it may be due to your skill in measuring—or it may simply result from random, chance fluctuations that we expect in all data. Since your conclusions may depend on the measurements you made, it’s important to know how good those measurements are.

Accuracy refers to how close a measurement is to the “true” value. For example, you will make a more accurate measurement of liquid volume with a graduated cylinder than with a beaker: the calibration marks on a beaker are only approximate. We usually determine how accurate a measurement is by comparing it with a known standard. If you’re measuring a volume of water, for example, you might check the accuracy of your measurement by *weighing* the water, knowing that 1 ml of water weighs 1 g.

Precision is how much variation there is when you make repeated measurements. Suppose you use a micropipettor to measure 1 ml of water. You do this three times, weighing the water each time, and get 0.751, 0.753 and 0.750 grams. In this case, your measurements were very *precise* (not much variation among the three measurements), but not very *accurate*, because 1 ml of water should weigh 1.0 g. Perhaps your micropipettor isn’t calibrated properly, or there’s something wrong with your technique. On the other hand, if you got 1.3, 0.75 and 0.96 grams, then your measurements were not very precise. In this case, it’s hard to tell what the “true” value is because there’s so much variation in the individual measurements.

Analyzing precision: average and standard deviation

In order to evaluate the precision of data, we commonly use statistical tools such as the **average** (or **mean**) and the **standard deviation**.

The mean (\bar{x}) is simply the sum of all the measurements divided by the number of measurements. For the three measurements 0.25, 0.19 and 0.12 mg/ml, we could report a mean measured concentration of 0.19 mg/ml. However, this doesn’t tell us anything about the precision of the measurements, so we can also report the standard deviation (s). A large standard deviation tells you your measurements are far apart: not very precise. The smaller the standard deviation, the less variation there is in your data.

The standard deviation (s) is calculated by the formula at right. For each measurement, the average of all the measurements (x_{avg}) is subtracted from the measurement itself (x_i), and the difference is squared. These results are then added together, and the sum is divided by one less than the total number of measurements. The standard deviation is the square root of this number. Excel and most calculators can calculate the standard deviation automatically, so you will probably never need this formula.

$$s = \sqrt{\frac{\sum (x_i - x_{avg})^2}{n - 1}}$$

Why is the standard deviation important? Let’s imagine that you’re studying the influence of some condition on the size of tree leaves, so you measure the length of 40 different leaves from each of two trees. Figure 5 shows the results in the form of a **histogram**: it shows how many leaves from each tree were between 20 and 25 mm in length, how many between 25 and 30, and so on. For tree #1 (black bars), the data are “tight:” there is variation, but all of the measurements fall pretty close to the average. For this data set, the standard deviation is fairly small. Tree #2 actually has *exactly the same average leaf size*, but the data look very different! The much larger standard deviation tells you that there is much more variation in leaf size for this tree than for the first tree. If you only reported the average, you wouldn’t get the true picture. The standard deviation would help your reader understand the data better.

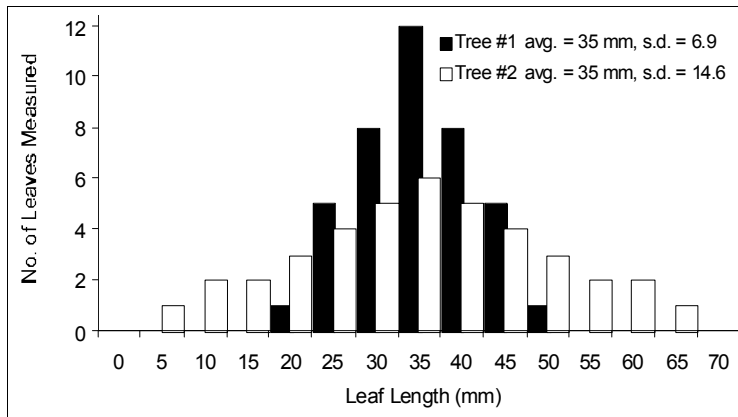


Figure 1. Histogram of leaf measurements for two different trees. Notice that the average value is the same, but the standard deviation shows how different the two data sets are.

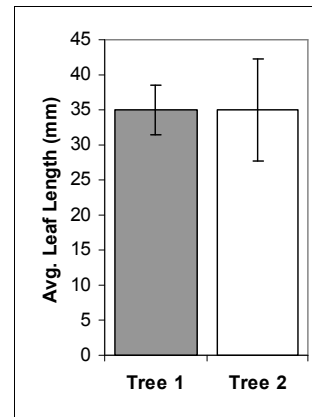


Figure 2. Average leaf length for the two trees. Error bars represent one standard deviation.

There are many ways to use the standard deviation. If you will be graphing your average data, a good way to put standard deviation to use would be to include **error bars** that show the size of the standard deviation. The section on graphing with Excel (page 49) has more detail on how to do this. Figure 6 shows how the average data would look if we used error bars whose height is the standard deviation. Notice that this figure gives you much more information than it would if you could only see the heights of the bars.

Significance

Take a look at the graph in Figure 7. It shows the results of testing a new drug which hopefully will be useful in treating the common cold. Hey, this drug looks great! Clearly, people who took the drug had colds that didn't last as long as people who weren't treated.

...Or not. How many people were tested? If only one or two people were in each group, then it is very likely that the apparent "difference" is not **significant**—it could easily be due to chance, or to individual variation, and not due to the drug. On the other hand, if you knew that 1000 people were in each group, then you might start to believe that the drug was really doing something.

Results are only meaningful if we can measure their *significance*. Averages and standard deviations can help here—if we knew that the groups were large and the standard deviations small, we'd feel more comfortable concluding that there really is a difference. In many cases, however, you will need additional statistical tests to determine whether an observed difference is significant.

In this case, for example, a T-test would be helpful: the T-test gives you the probability that there is a significant difference between two sets of data. It takes into account how much variation there was and also the sample size (higher sample sizes are much more likely to yield significant data). In other situations, you might use a χ^2 test to compare experimental results to hypothesized results, or an ANOVA to look for differences among multiple data sets. The *Handbook of Biological Investigations* that you purchased for your introductory courses can help you with this, and so can faculty members familiar with the kind of analysis you're doing. A statistics course (e.g., PSY 250) is required for a BS degree and strongly recommended for BA students; taking this course as a sophomore would make it much easier for you to deal with the statistical analysis of data in all your lab courses and research!

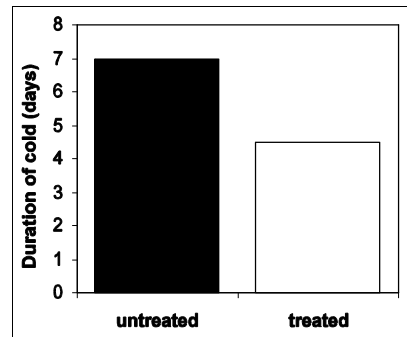
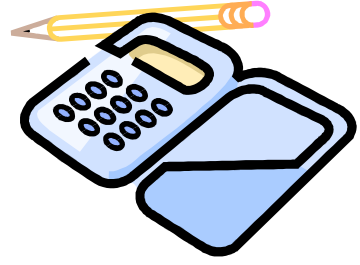


Figure 3. Duration of colds for individuals treated with a new drug vs. control group

Numerical accuracy

Remember what you learned in chemistry about **significant figures**? When your calculator spits out an average of 0.364125, think about whether your measurements are really that accurate! If you reported this number, it would mean that you could expect to accurately distinguish between a measurement of 0.364125 and a measurement of 0.364126! That's pretty unlikely, unless you're using some really fancy measuring equipment.



Typically, the kinds of measurements we make in biology support no more than 2-3 significant figures, so you should round numbers appropriately. The average above might be reported as 0.364, 0.36 or even just 0.4, depending on how the measurements were done. When you just write down whatever's on your calculator without considering significant figures, your instructor knows you haven't really thought carefully about the measurements you've made.

When you're graphing with Excel, you can change the number of decimal places displayed for the cells of a spreadsheet, the axes of a graph, or even the equation of a line. Don't settle for what the program spits out; learn to make the program work the way you want it to.

