Statistics Summary

When researchers conduct trials they not only want to know how “things” compare in that particular experiment but also how those results extrapolate to other situations. They also want to know how consistent the results will be. To address these needs experiments include replication and randomization. Biological systems are inherently variable. The researcher has to account for the variation and both replication and randomization help control or identify sources of variation. When data are collected and analyzed statistically, in many cases the first step is to run an analysis of variance test (often referred to as ANOVA). This allows the researcher to gain an appreciation for the source of variation in the experiment. Determining the ratio of known sources of variation (treatments) to unknown sources of variation (considered error variance) is important (F statistic is the term used for this ratio.) The larger the ratio of known sources to unknown sources the greater the likelihood that the researcher will have quantifiable differences between treatments. When this ratio is low, the researcher most likely will not have quantifiable differences between treatments. However, a low ratio of known to unknown variance can be caused by several things. First, it may simply be that treatment A versus treatment B impact the crop or pest, for example, the same way. If someone uses either of these treatments they will get the same result. Alternatively, it may seem like there is no difference in the treatments because the ratio is low, but this may be because there is a tremendous amount of variation in the experiment that is unknown. To draw a conclusion that the treatments perform the same when there is a great deal of random or unknown variance can be a mistake. There are many reasons for the unknown variance to be high (which results in a lower ratio) and can include poor experimental procedures and improper blocking. The researcher needs to control the unknown error as well as possible and this involves using solid experimental procedures and blocking. Larger differences in treatments (the magnitude of the differences) can overcome larger amounts of unknown variance. If response to treatments is large (my example was no nitrogen on corn versus 150 pounds actual nitrogen) the ratio of known variance to unknown variance will still be large even if the unknown variance itself is large. But if you are trying to determine if 120 lbs of nitrogen is just as good as 150 pounds of nitrogen, a larger amount of unknown variance may cause you to say there is no difference in 120 vs. 150. The challenge is whether the researcher thinks this is what is really going on with the treatments or is the conclusion an artifact of a great deal of unknown variance in the experiment (often called experimental error.)

Once the researcher has made a decision on whether or not the experiment was conducted in a reasonable manner, he or she can draw conclusions about the treatments. In doing so the research also wants to protect from making a wrong conclusion (and recommendation) based on random chance. This is done using hypothesis testing. See the previous handout for more details on that. In short, the researcher wants to be as sure as possible if the difference in treatments will occur in the future when these treatments are compared or when one is adopted at the expense of another. Generally, researchers will accept that 5, 10, or 20% of the time they make a wrong conclusion simply based on chance. You will often see a P value (probability) in data tables. Many times it will be written $P > F$ (or probability of a greater F statistics, which refers to the ratio of known to unknown variance.) The lower the p value the more confident the researcher will be that the results he or she sees in the experiment will also be the case in other situations. The p value is decided by the researcher (the community of researchers in that field of study.) For example, in agriculture, researchers are more comfortable with higher P values because the "cost" of a mistake in drawing a conclusion based on random chance is not catastrophic. In contrast, if an airplane manufacturer is comparing components of a hydraulic system or new rivets made from a lighter weight material, the concern over making a mistake due to random chance is much more important (than the agricultural example) because the result can be catastrophic.