
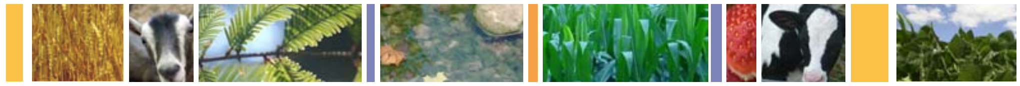

FACT SHEET
Agriculture and Natural Resources

Statistics and Agricultural Research

Robert Mullen, Soil Fertility Specialist, Ohio State University Extension
Ed Lentz, Extension Educator, Seneca County, Ohio State University Extension
Greg Labarge, Extension Educator, Fulton County, Ohio State University Extension
Keith Diedrick, Research Assistant, The Ohio State University

When a state Extension specialist or Extension educator makes a presentation, the individual will occasionally make reference to "statistical significance" or some variant that alludes to statistical analysis and its use in determining treatment differences. So what is meant by statistical significance? Why should a producer, consultant, or retailer care about statistics? Can the average of treatment effects be used alone to evaluate differences? These are often-asked questions that need clarification. Within this fact sheet we will attempt to explain why researchers use statistics as a tool, why statistics are useful and necessary, and why it is difficult to draw meaningful conclusions from split-field, unreplicated data.

Useful Terms

Observation — a measurement that is made for some output(s) of interest (yield, plant stand, nutrient status, disease incidence, insect infestation, etc.).

Treatment — the controlled application of a process or product (seeding rate, fertilization rate, insecticide application, fungicide application, etc.) to an experimental plot that will hypothetically have an impact on an output(s) of interest (growth, yield, etc.). Usually more than two treatments are utilized in a well-designed experiment.

Experimental error — differences in observations from treatments due to environmental conditions that cannot be controlled by the experimenter (differences in soil texture, topography, soil compaction, rainfall, nutrient status, disease infestation, etc.).

Statistics 101

Any observation made within an experiment has a certain amount of error associated with it. In order to determine whether or not numerical differences in observations are due to treatments, we need to know how much error was encountered within the experiment. Statistics allow us to quantify and assess this error (experimental error). If only a single observation is made you cannot estimate experimental error. Multiple observations associated with each treatment, or replications, are needed. In a field experiment, the observations can be confounded with a multitude of uncontrolled soil and environmental factors; therefore, we must replicate the treatments across the landscape. To ensure the estimates of experimental error for each treatment are unbiased (not systematically influenced by underlying environmental conditions like soil type, topography, etc.), the replications should be randomly placed within the field. We have just discovered the two most important concepts of modern statistics: (1) to estimate the experimental error of treatments requires replication, and (2) to ensure an unbiased estimate of experimental error requires randomization of the treatments.

Statistical Significance

Statistical significance is often mentioned but seldom explained. When an experiment is conducted (properly replicated and randomized), the experimental error is computed and used to assess whether or not treatments differ "significantly" from one another. Statistics are based on probability, and researchers select what level of probability constitutes significance. The probability level (often referred to as "*p*" in scientific studies) selected is solely at the discretion of the researcher. The scientific community in general prefers a probability level of 90% or 95%, meaning that a researcher can state with 90% or 95% probability that the difference between treatments did not occur by sheer chance. If the 95% probability criterion is met, then the treatments are "significantly" different. This is where some gray area enters into research; what is the appropriate probability level? Each researcher has his or her own set of criteria. The next time you attend an Extension event and the speaker is discussing some research data, think about what level of probability is being used to evaluate treatment differences.

Statistics allow researchers to assess the error associated with conducting an experiment and to separate real treatment differences from differences caused by uncontrollable environmental factors. Researchers can separate the grain from the chaff as it were. Like any tool, it must be used properly to be effective (replication and randomization).

Importance of Replication

Assume you want to evaluate a fungicide treatment on your farm, so you split a field in two and apply the treatment to one half and leave the other half untreated. At the end of the year you harvest each of the two halves and observe a 3 bushel per acre increase in yield on the treated side. This 3 bushel per acre difference seems like a good deal, so you decide that next year all of your acres will be treated with this new fungicide. Are you sure that the additional 3 bushels per acre was due to the application of the fungicide? Closer inspection of the field reveals that the half of the field that showed the yield response was dominated by a lighter texture soil that drained better than the other half of the field. Due to excessive moisture the half of the field with better drainage might be expected to perform better. With the field split in two, it is impossible to determine what factor contributed to the yield increase. There are a multitude of other possible explanations for the yield increase: historical management differences, fertility level differences, insect pressure, disease pressure, natural variation in soil productivity, etc. Since we have no replication it is very difficult to reach a definite conclusion as to the cause of the yield increase. This is not to say that the 3 bushel per acre increase was not real; you just do not know that the yield difference was due to the treatment you applied or to some other factor.

Replication allows us to estimate the error associated with carrying out the experiment itself. Let's revisit the fungicide experiment. Assume you split the field into strips and established three strips that were treated with the fungicide and three that were not. We will look at two different scenarios based on the harvest information.

Scenario 1

At harvest the yield levels of the three treated strips are 50, 59, and 50. The three untreated strips yielded 44, 57, and 49. The average yield levels for the treated and untreated strips are 53 and 50 bushels per acre, respectively. Statistical analysis reveals that the probability of the fungicide treatment resulting in greater yield by sheer chance is 57% ($p = 0.57$). Thus, as experimenters, we are concerned that the differences in yield between the two treatments may have occurred by chance alone.

Scenario 2

At harvest the yield levels of the three treated strips are 54, 53, and 52. The three untreated strips yielded 50, 52, and 48. The average yield levels for the treated and untreated strips are 53 and 50 bushels per acre, respectively (same as scenario 1). Statistical analysis reveals that the probability of the fungicide treatment resulting in a 3 bushel per acre yield increase by sheer chance is 8% ($p = 0.08$). Thus, we are comfortable stating that the treated plots yielded significantly higher than the untreated plots.

The only difference between the two scenarios is the variability (and resulting experimental error) in the data collected. The averages for each treatment have not changed, but notice the spread in the data in scenario 1. Large error in the data makes it much more difficult to identify treatment differences. In other words, some underlying source of error exists that we cannot control or possibly even measure.

Importance of Randomization

While not stated explicitly in the "Importance of Replication" section, randomization is just as important as replication. Think about our initial experiment where the field was split in two. There was an underlying difference in soil productivity due to soil texture and drainage that could affect the experimental outcome by biasing (confounding) the data. To properly conduct the experiment, this variation should be accounted for in the experimental design. Even if you replicated both treatments (with and without fungicide) three times as you did in the Replication section, the conclusions you reach may not be correct if the fungicide treatment was always applied to the same half of the field. The data would be biased (confounded) based on its location in the field.

Least Significant Difference

The next item in our discussion of agricultural statistics is the term "least significance difference" or LSD. This number is often mentioned at Extension meetings and in university publications that provide information and summaries of research. The question is, "What does this number mean?"

Least significant difference is used to compare means of different treatments that have an equal number of replications. What does that mean? Let's take our example above. We had two different scenarios which can be seen below:

Scenario 1	
Treated plots	Yield, bu/acre
Rep 1	50
Rep 2	59
Rep 3	50
Average	**53**
Untreated plots	Yield, bu/acre
Rep 1	44
Rep 2	57
Rep 3	49
Average	**50**
LSD (0.10)	7.4

Scenario 2	
Treated plots	Yield, bu/acre
Rep 1	54
Rep 2	53
Rep 3	52
Average	**53**
Untreated plots	Yield, bu/acre
Rep 1	50
Rep 2	52
Rep 3	48
Average	**50**
LSD (0.10)	2.0

Recall back to the previous discussion that for the first scenario the probability of the treated plots being different from the untreated plots by sheer chance was 57%. For scenario 2, the probability that the treated plots would be different than the untreated plots by sheer chance was 8%. This was primarily influenced by the amount of error associated with the experiment for each scenario, and we are much more comfortable attributing yield differences to scenario 2 because the probability of the yield differences existing due to sheer chance is low.

Now let's look at this another way using LSD.

For scenario 1, at a significance level of 90% (or stated as 0.10) the LSD value is 7.4. For the treated plots to be different than the untreated plots they must differ by at least 7.4 (which they do not).

For scenario 2, at a significance level of 0.10 the LSD value is 2.0. Since the differences between the treatments are greater than 2.0, we feel comfortable stating that the treatments were significantly different.

Hopefully this will help you understand whether or not two treatments are different the next time you are sitting in an Extension meeting or reading a research summary. Remember, as has been mentioned before, research studies should be conducted over multiple locations and under different environmental conditions to prove their robustness.

Summary

Statistics allow us to evaluate treatment differences and determine whether or not the differences that exist are due to the treatments applied. It allows us to make meaningful comparisons to help us decide what production practices are beneficial and those that are not. A general understanding of statistics will help you as end-users to understand how university Extension personnel arrive at their recommendations and prompt you to question information that is being sold to you. Remember that in order for an experiment to be properly carried out it must contain replication and randomization. Split field information, while useful, has severe limitations that should be viewed with caution.