STATISTICAL APPROACHES TO STUDIES INVOLVING ANNUAL CROPS

Elmer E. Remmenga[1]
Department of Statistics, Colorado State University, Fort Collins, CO 80523

The development and use of statistical techniques in conjunction with experimentation on annual crops has been widely accepted and recognized for many years. Simple, straightfoward field research with annual crops enhanced the validation of the statistical procedures. Simultaneously, progress in field research benefitted from application of statistical methodology.

Much of the early development of statistical ideas in association with field experimentation is a result of this mutually beneficial system. Therefore, current discussion of the use of statistical methods with annual horticultural crops should be a natural. There are, however, many unsolved problems and misunderstood techniques.

As modern technology has become more complex and research equipment more sophisticated, the tendency has been to make statistical procedures more complicated; after all, with computers it is possible to process any data set and perform an analysis. Overzealous people have dashed madly down this path; and, as a result, many research projects have been poorly conducted at best, and much honest effort has been totally wasted.

The research world, even in the area of annual crops, has become a vast maze of technical complications. It is no longer possible for the researcher to work alongside or with a small group of technicians. It is necessary for the horticulturalist to utilize the services of, or to work in close conjunction with, many specialists in electronics, computers, genetics, pesticides, fertility, nutrition, pollution, and many others, including the statistician.

Field aspects of experimentation

Research technology has developed in such a manner that statistics must play a very important role. The statistical aspects of experimental design must be considered early in the planning stages. The design must be statistically valid and analyzable. When a problem is posed and defined, research objectives should be stated. These objectives lead to statistically testable hypotheses which are guides to the choice of treatments and selection of experimental material. In the final stages, data analysis and statistical evaluation become critical to valid interpretations.

However, statistics in the absence of subject matter knowledge is not a solution to anything and tends to interfere with good research operation. Biologists, along with many others, tend to get carried away with sophisticated technology and elegant procedures. Statistics is too often used as a veneer to cover up the lack of a good field operation. A poor set of results is too often masked under a mass of statistical jargon, or a fancy data analysis is substituted for adequate knowledge of the subject matter area. Statistics cannot replace good basic research methodology; it cannot salvage a poorly done experiment and should not be used to cover up research inadequacies. Many researchers use statistics in this manner, however, and have the impression that statistical manipulation is the panacea to all research problems. This overreliance on statistics is a serious mistake.

Field problems. Good research starts with a basic knowledge of the subject matter at hand and with close study and understanding of the field problems that may be encountered. These ideas have been presented many times in bits and pieces in most experimental design and/or statistical technology texts, but they are so important that they are worth repeating.

Many different types of studies with annual crops are conducted, and each specific types has its own conditions and special problems. Fertility studies have different characteristics than variety studies; cultural investigations have unique problems, and pesticide studies have a different range of peculiarities. Some of these will be mentioned, but all field studies have a set of common problems. Effect on soil and climate are often of greater magnitude than any of the anticipated treatment responses. Localized drainage or disease problems may affect or mask the responses to treatments being evaluated. Unplanned changes in techniques or equipment may become completely confounded with treatments. Most experiments must be repeated several times at different locations in different years in order to establish valid inferences.

Soil. Soil heterogenity, in terms of fertility, texture, drainage, and other aspects, is of major concern. Great care should be taken in selecting the location of research plots. Soil samples ought to be evaluated and soil pits dug to aid in selecting the site. Most people have few options in the choice of site, so this information also can be used to critically evaluate results. Natural soil problems can be illustrated by small sand or gravel bars or by hidden drainage channels resulting in local differences in fertility, organic matter, and moisture. Some of these are obvious or at least apparent; others are not and appear only after the experiment is underway or completed. Additional soil irregularities are created by previous soil treatments, fence rows, plowing, and other aspects of cropping history. Of particular concern are previous research treatments, such as fertility studies, which have been imposed on the area. Such fields should be subjected to uniform, nonexperimental cropping for one or more years to deplete such effects.

Weather. Differences in weather between years and within years create large effects on field research results. The effect of precipitation patterns during a season may be as great as the effect of total amount of precipitation. Local irregularities caused by screening due to trees or buildings should be avoided. The effects of irregular drifting of snow and resulting moisture differences can affect growth and yields the following summer. There is no way that a researcher can control or eliminate the effects of weather; but it is necessary that annual and within-year effects be recognized and considered in the conduct and interpretation of a field study. Any study whose results are to be used as a basis for inference to the future must be conducted over a period of years.

Plot size. Plots should be sufficiently large to provide an adequate amount of material so that the responses can be judged without the precision in measuring or weighing unduly affecting the results. There is always the possibility of loss of some material in harvest, and measurement errors always exist; these effects are minimized with reasonable-sized plots. Typically, precision increases as plot size increases, reaching near constant at a reasonable size. Beyond the optimum size, precision then may decrease due to increasing soil heterogenity or to labor problems. The need to keep treatments close together so they can be compared directly is at odds with the need to have a large enough plot to evaluate each treatment adequately. Only experience with a given crop and the local situation, including soil problems, can determine the near optimum plot size. Unfortunately, most plot sizes are determined by efficient work conditions based primarily on equipment. Efficient conduct of a study cannot be ignored but should not be the primary control of statistical precision.

Plot shape. Plot shape has a major effect on precision and accuracy

[1]Professor.

of results. This is partially due to machine and equipment limitations and partially due to border or edge effects. Rectangular plots, often long narrow strips, are the usually recommended shape for efficient use of equipment but also so that every plot has equal exposure to other field conditions. The long axis of the plots should be oriented across; that is, at right angles to, the major direction of soil heterogenity. Small square plots often defeat the goal of obtaining homogeneous plots. Each small plot may be unique. It is necessary to have plots sufficiently large and properly oriented so that they contain a composite of soil heterogenity and thus become similar.

Plant problems. These plot size and shape problems cannot be determined unless the characteristics of the crop are known. Plant size and shape, growth habits, seed size, planting and harvesting procedures, and purpose of the study must be considered. Genetic uniformity of the crops must also be considered. Response to fertility may be masked by extreme variability in plants, in germination, or vigor. Sometimes the treatment itself has an effect on the plants, to injure them or to enhance growth, in some nonrandom fashion, such as a nitrogen application making plants vigorous and thus more resistant to disease or insects and thereby interfering with the disease study.

Competition. Competition between plants (or plots) treated differently must be controlled, eliminated, or standardized. Plant characteristics, such as height leading to shading or growth rate leading to competition for light or moisture, must be considered. Studies involving planting or harvesting dates must be organized so that early growth is not affected by the large plants in the next plot or late growth changes due to removal of competition by early harvest of an adjacent plot.

Competition within plots is another source of variability. It is often difficult to get uniform application of experimental treatments ranging from uniform planting rates, uniform application of fertilizers, irrigation water, pesticides, diseases, weeds, and any other effect. Whether these effects are experimental or incidental is immaterial; lack of uniformity can make the results meaningless. These effects are at least partially eliminated by randomization.

Field access. Access to plots is required for numerous field operations as well as for research purposes. Plants at sides or ends of plots adjacent to these access alleyways will grow differently than plants in the interior due to differences in competition for light and nutrients. A row or rows of plants of the same species must surround these plots to provide uniform competition everywhere.

Border rows. Experiments involving any treatment which may move in the soil after application, such as irrigation, fertilizer, some tillage effects, or drift during application of herbicides or other pesticides, must be organized with border rows between treatments to provide uniform plant competition and unique treatment effects. It may be necessary to plant 3 to 6 rows and to harvest only the center 1 to 3 rows for study. As a result, the efficiency of use of space decreases. Sometimes only ⅓ to ½ of the plot area is actually used for experimental plots.

General. The list of such ideas is endless and most are species and location specific. The purpose here is to emphasize that a good study demands that an exhaustive critique by the researcher of all nonstatistical techniques should be made before statistical procedures are incorporated. Some years ago, I read a comment about design which I have been unable to find at the present time; however, I believe it was in one of R. A. Fisher's books and will give him the credit. While the exact source and the precise quote escape me, the gist of the comment was: "The results of properly designed and well-conducted experiments do not need elaborate analysis." This statement is as true today as it was 40 or 50 years ago.

Good experimental design requires that the researcher take a good deal of care and do his homework well in such areas as selecting experimental material; selecting experimental units compatible with the crop grown; selecting treatments appropriate to the objective; refining research techniques to apply treatments uniformly, measuring results suitably and unbiasedly, and controlling outside effects; obtaining additional related measurements (but within reason); and *then* looking to statistical techniques.

Statistical aspects of experimentation

Recognition of the many sources of variability that exist in field experimentation and controlling or eliminating as much of this variability as is practicable is the first step in conducting a study on annual crops. Proper interpretation of results from such trials depends on the estimates of experimental error. A device or procedure must be developed to permit correct and appropriate computation of the statistical variability inherent in the study. Only with this estimate of variability can differences between treatments be evaluated. Experimental error can be estimated only after all sources of variability have been eliminated except for variability due to chance.

The validity of this error depends upon the degree of correspondence between the conduct and results of the experiment and the mathematical basis for the statistical analysis. Reasonable assurance of a high degree of correspondence between the real world and the abstract mathematics can be obtained from repeated experiments that follow sound statistical practices.

Statistical requirements. The basic statistical requirements have not changed over the years and are discussed in numerous publications. There is no substitute for randomization, replication, and local control. Randomization is essential for valid estimation of experimental error and to obtain estimates of means within a minimum bias. It is also the basis for tests of significance and confidence intervals; without randomization the probability statements have little credibility. Replication is needed in order to compute an arithmetic estimate of variability and to provide a more reliable estimate of means. A large sample is generally more representative than a small one simply due to the effect of numbers but also because the treatment has been exposed to a wider range and variety of the natural conditions. A large sample can be obtained from a few large plots or many small ones and, in a given instance, say for a fixed area, some compromise based on the specific crop, available equipment, and other factors must be made. The exact number of replicates needed is impossible to determine. It is a function of variability, difference of interest, and probability, but is most often determined by economic factors such as cost, space, and labor. However, the experimental error ought to be based on 15 to 20 degrees of freedom. The concept of local control is affected by restrictions on randomization to eliminate recognized sources of variability that are not pertinent to the required comparison. Arranging treatments in simple blocks to remove the between-block variability from the experimental error is a simple illustration.

Experimental design. The specific experimental design used depends on the experimental conditions and limitations of each case. The available experimental material or area, the treatments chosen to answer the objectives, equipment, labor, and costs all enter into the actual design. However, the simplest possible design that fills the needs is the best choice. Not that complex designs are to be disregarded, but they should be selected only as dictated by experimental conditions, not simply for the sake of sophistication. Too often, the complexity of conducting a high-powered design to overcome an experimental shortcoming creates problems and gross mistakes which distort the error to a greater extent than would happen if the original limitations were ignored. This is particularly the case when computational difficulties occur. If the experiment is apt to involve data loss, the effect of imbalanced data on a complex design is much more of a problem than a similar amount of missing data for a simpler design. Every text on experimental design advocates factorial experiments as preferable to single factor studies, and there is no disagreement provided everything works; but a factorial experiment is no better than the researcher's ability to carry it through. If time and space do not permit conducting, analyzing, and interpreting such a study, the researcher had better reduce the size of the experiment to a study he can handle. This problem is exaggerated by complex designs and analyses. A solid conclusive result from a simple *t* test is preferable to a vague result from a correct but complicated model.

Given reasonably homogeneous experimental material, a completely random experimental design will give maximum degrees of freedom for error and allow the researcher to concentrate on his research material uncomplicated by design and analysis problems. A random sequence must be followed in all manipulation of plots and the resulting material throughout the experiment so that each treatment and plot has an equal opportunity for whatever good or bad events may occur. However, even under the best of conditions, the researcher should consider a randomized block design where each block or set of treatments constitutes a unit of effort which can be performed in a short period of time. Any interruption, such as a rainstorm, in the conduct of the various field operations ought to occur equally for a block and can be removed from the error as a block effect. Further partitions, into incomplete blocks, follow naturally from the same line of thought either based on real or potential stratifications. There is a limit on the complexity of field design that can be successfully undertaken, particularly in a university atmosphere. The size of the experiment leading to the necessity of a complex design is often more than can be handled with the area, time, and funding allocated. Even if this can be overcome, the need for additional technicians introduces another factor of such major concern as to force the project into a smaller sphere for any valid results.

The set of treatments involved may have an effect on the design (assuming adequate replication). A straight-forward cultivar trial presents no problem, except to keep the number of entries to a reasonable number. A simple fertilizer study to evaluate the response curve needs an adequate number of increments in the correct range to determine the form of the response (say 8 to 10). A factorial involving cultivars and fertility, cultivars and cultural practices, or other arrangements can be handled in a randomized block; or the researcher, because of application problems or drift problems, may wish to form a split-plot design based on the randomized block. Competition problems, border effects, drift problems, etc., can be controlled and partially eliminated by placing these types of treatments on large plots and splitting these plots into subplots for the other factors. There are statistical problems and advantages associated with split-plot designs, but most often split plots are used for convenience, in that some treatments are easier to handle and control on large areas. If the concepts associated with these designs are understood, the move to more advanced designs as required by experimental conditions is straightfoward and natural. These designs should be utilized as needed, but only if the researcher can adequately conduct the program.

Analysis. Once the experiment is carried out and the data collected, the proper analysis should be completed dependent on the design. The analysis should be pertinent to, and directed by, the objectives of the experiment. The analysis of variance provides the appropriate estimate of the experimental error and can be used for an F test on treatment means. However, such a test will seldom be sufficient; it is almost necessary to conduct a specific test to isolate and identify specific treatment effects and differences of interest. These tests should be part of the overall experimental design, in that the most valid tests are based on information independent of the experiment. It is preferable to preplan meaningful interpretable comparisons and thus have the techniques laid out before the data are available, as everyone tends to be guided by the results. Linear contrast procedures, including orthogonal polynomials where appropriate, are recommended, but limitations to orthogonality should be ignored if the contrasts are meaningful. Orthogonality, per se, does not ensure meaningful nor interpretable contrasts.

If meaningful interpretable contrasts cannot be preplanned, such as in a cultivar trial, this does not connote a poor experiment. Many studies are of this type and are totally valid. For these studies, a recommended mean separation test is Tukey's HSD, as the alpha level is stable; however, several equally valid tests are available. Homogeneous subgroups are identified, and this is usually a satisfactory procedure.

Presentation. Simple, understandable tables and/or graphs should be constructed to present the means obtained from the experiment. In fact, these presentations should be prepared to correspond with the objectives of the study; the experiment should be designed to gather data which can be compiled to complete the tables or graphs, and the analysis should be performed to verify and respond to this set of information, which supports conclusions drawn to answer the objectives. A recommended procedure is to write a tentative conclusion for every experimental objective at the time the research is being planned. Each conclusion will require a set of data, in a table or graph, to support it. The experiment can be designed specifically to obtain the data for the tables or graphs. The tentative conclusion may need to be altered when the data are obtained, but this procedure gives direction to the entire operation. Sometimes, unexpected, interesting results pop up during an experiment, but such occurrence cannot be counted on to justify a study; the best results are obtained from a well-organized, simple experiment directed toward specific objectives.

Conclusion

Nothing new has been presented here — only a compilation of many ideas from numerous sources, some of which can no longer be found. Many years of experience have lead to the conclusion that the best experiments require little statistical support while poor ones require a lot from statistics. The best experiments are probably conducted in the presence of good statistical procedures which do not, and need not, show in the end product. The poor experiment is often supported by statistical flourishes because the right concepts were absent in the beginning.

The best results come from simple, well-planned experiments with a few straightforward objectives carried out thoroughly and analyzed complete by knowledgable conscientious people.

Selected References

Cochran, W. G. and G. M. Cox. 1957. Experimental designs, 2nd. ed. Wiley, New York.

Fisher, R. A. 1937. The design of experiments, 2nd ed. Oliver & Boyd, London.

LeClerg, E. L., W. H. Leonard, and A. G. Clark. 1962. Field plot technique, 2nd ed. Burgess, Minneapolis, Minn.

Little, T. M. and F. J. Hills. 1975. Statistical methods in agricultural research. Univ. of California, Davis.

Snedecor, G. W., and W. G. Cochran. 1967. Statistical methods, 6th ed. Iowa State Univ. Press, Ames.

Steel, R. D. G. and J. H. Torrie. 1980. Principles and procedures of statistics, 2nd ed. McGraw-Hill, New York.

Wishart, J. and H. G. Sanders. 1935. Principles and practice of field experimentation. Empire Cotton Growing Assoc., London.