

Kansas State University Libraries

**New Prairie Press**

---

Conference on Applied Statistics in Agriculture

1989 - 1st Annual Conference Proceedings

---

## TEACHING STATISTICAL METHODS TO GRADUATE STUDENTS IN COLLEGES OF AGRICULTURE

D. F. Cox

Follow this and additional works at: <https://newprairiepress.org/agstatconference>



Part of the [Agriculture Commons](#), and the [Applied Statistics Commons](#)



This work is licensed under a [Creative Commons Attribution-Noncommercial-No Derivative Works 4.0 License](#).

---

### Recommended Citation

Cox, D. F. (1989). "TEACHING STATISTICAL METHODS TO GRADUATE STUDENTS IN COLLEGES OF AGRICULTURE," *Conference on Applied Statistics in Agriculture*. <https://doi.org/10.4148/2475-7772.1459>

This is brought to you for free and open access by the Conferences at New Prairie Press. It has been accepted for inclusion in Conference on Applied Statistics in Agriculture by an authorized administrator of New Prairie Press. For more information, please contact [cads@k-state.edu](mailto:cads@k-state.edu).

## TEACHING STATISTICAL METHODS TO GRADUATE STUDENTS IN COLLEGES OF AGRICULTURE

D. F. Cox  
Iowa State University, Ames, Iowa

### Abstract

One method of judging the effectiveness of the teaching of statistical methods is to rate the quality of their use in the research journals of the agricultural sciences. A conclusion that improvement is possible is easily supported by such a review. Basic concepts such as the meaning of replication, the definition of an experimental unit and the nature of experimental error are misunderstood, and this leads to faulty analyses and incorrect conclusions. Changes in teaching that would improve the product are not specified easily because research on the topic is difficult. Perhaps more emphasis on the basic concepts of experimental science could lead to more complete understanding of the nature of statistical thinking and, consequently, more effective and correct use of the methods available.

Keywords: statistical methods, teaching statistics.

The teaching of statistical methods to research workers in agriculture is at least 50 years old. One view is that it started with Fisher's "Statistical Methods for Research Workers"(1) in 1925 and with Snedecor's "Statistical Methods"(2) in 1937. The objective here is to examine the effectiveness of the teaching in this area.

Universities have a continuing interest in evaluating teaching but always find objective, quantitative criteria difficult to establish. Those who teach statistical methods to prospective research workers have a unique source of information on which to judge the effectiveness of the instruction. The source consists of the journal articles eventually published by the research workers who learned the methods that they use in the courses taught in the graduate schools of colleges of agriculture. Unfortunately, almost any review of this published material leads to a conclusion that improvement is possible.

Many have written and spoken on this subject. Often the scientists themselves are criticized, then the journal editors and reviewers are admonished, finally, the outputs of the software packages that do the computing are faulted. What may be overlooked in this listing is that the main problem resides in the textbooks and the classrooms where the subject is taught.

First, the claim that the use of statistical methods in the publications of agricultural research could be improved needs some support. There are many levels of misuse and incorrect understanding but the most destructive are those that involve basic concepts. For example, the definition of an

experimental unit and what is meant by replication is often missed in studies that involve environmental chambers. Each chamber is set at a different temperature and each chamber may contain 100 plants. The notion that judgments about temperature effects can be based on an experimental error derived from the variance among plants within a chamber is incorrect but often held. The problem is not confined to the plant sciences but often occurs in animal studies where pen and family differences are ignored in the same way that chamber effects are not accounted for in the previous description.

Most test statistics require an assumption of homogeneous variance. The analysis of variance is a procedure that provides a pooled estimate of experimental error. However, the literature abounds in examples in which standard errors of treatment means are found by using only the units on a given treatment to estimate error variance for that group. Much of the software presently used in data analysis seems to encourage the practice.

The device of splitting the primary unit into subunits over which other treatments are assigned is common in much of agricultural experimentation. A typical plot of the results from such experiments shows the response versus levels of the subunit treatment for each level of the whole unit factor. Often, error bars are added to these graphs. The source of the error used for these bars is often chosen incorrectly. The majority used the subunit error, but the expectations of differences among levels of the whole unit treatment at a given level of the subunit treatment requires a combination of the errors in this mixed model case, and to ignore this can substantially underestimate the realities involved.

No statistical method has received more attention, seen more use and been more commonly misused and abused than the operations of multiple comparisons. An example taken from an actual journal article reporting the results from the applications of factorial treatment design in which four amounts of factor A were combined with five amounts of factor B is given in Table 1. Much simpler and more meaningful ways to display the main effects and interactions of these two factors exist than the attempt made by the clutter of letters given here.

Often trends in a response over the levels of a quantitative treatment factor are correctly summarized by using a simple straight line. However, many authors also include the additional test of all pairwise differences among the means without recognizing the redundancy and inconsistency in the logic of using these two summaries simultaneously.

Documenting the case that there is need for improvement in the practice of using statistical methods is not difficult. Suggesting ways that the teaching of these methods might change to improve the situation is a much more challenging assignment. Lecturing on what should be done in the classroom invites ideas, hypotheses, suggestions, subjective feelings and emotions together with almost no scientific experimental evidence on which to resolve any differences in opinion. There cannot be other than many different views.

I believe that courses in statistical methods need to move toward emphasizing basic concepts in experimental science. We need to explain how

science manages, controls and describes the uncertainties with which it must deal. The difficulty of assigning causes to the responses that we measure needs to be appreciated. All teachers of statistical methods try to do this, but it is often covered with unnecessary detail. Many students enter graduate school without any introduction to the concept of the inherent variability of experimental material or the idea of experimental error. Often formal definitions of things such as experimental units are not very helpful, and it is only after students begin to do experimental work that such ideas take hold. Replication is a difficult concept unless one understands the nature of the experimental units involved.

Randomization seems naturally abhorrent to many people who wish to be scientists. Most of them are careful, orderly, organized people who work in very disciplined ways. For them to give up and purposely inject chaos into their work through randomization goes against training and belief. They need some strong justification for such a procedure, and I do not think that the protection against bias is enough. Rather, I believe, one needs to demonstrate that randomization provides at least one rational basis for inference. The simplicity of the randomization test does not seem to appeal widely, and its logic is not easily accepted. It deserves time and care in presentation.

Demonstrations of how randomization reduces bias can be useful. Ask classes to choose random numbers from the digits 0 to 9 and be assured you will generate a nonuniform distribution, there will be too many 7's. A colleague of mine carries a set of 100 rocks to class and asks students to choose a representative sample of five and then estimate the total weight of the 100 rocks. People asked to do this will overestimate the total weight by an average of 15%.

The structure of data is important when one uses statistical software. However, the concepts of crossed and nested classifications are not well defined in the textbooks. The concepts are new to most students and familiarity does not come easily.

Obviously the best thing that can happen to promote the concepts of experimental science is for students to become involved in doing experimental science and that means to collect data, to summarize and display what is collected, to deal with missing, incorrect and uniform data and all the other problems data collectors face. Mostly the student's major field must supply this experience, but often it comes too long after they have taken the courses in statistical methods for the interconnections to be appreciated. Term projects in statistical methods courses that involve planning, conducting and summarizing an actual experiment are one way to provide the experience that seems vital.

Most of the textbooks on statistical methods descend from Snedecor's early efforts, and that book is now in its eighth edition. The wisdom in that textbook is excellent and the coverage broader than most such texts. However, the outline was set down 50 years ago in an era when the central chore of data analysis was the computation of sums of squares and cross products. In one sense, computation is the least of our concerns today, but we do not let go of it easily. Textbooks may now spend too much time demonstrating calculations that students will seldom make except on the exams

we give them. The argument that such calculations increase understanding and appreciation for underlying concepts needs scrutiny. Textbooks often rush to show how to calculate statistics and make tests of significance between group means without enough discussion of where the numbers in these groups came from, how the groups were formed and the nature of the units involved.

Textbooks often suggest operations that are seldom needed, and the overuse of multiple comparisons could come from this tendency. An example used to motivate range tests often starts by stating that one might be faced with testing differences among 30 varieties of a crop and then continues by providing a means of handling the 435 pairwise t-tests involved. The reader is not told that agronomists almost never have 30 varieties with no structure at all, but rather, that there are almost always early and late varieties, susceptible and resistant varieties and all sorts of other reasons for many planned contrasts among the 30 means that answer relevant questions, and no need for the noninformative range tests. However, there is much more to the overuse of multiple comparisons, and their appeal is subtle and widespread as any software package will attest.

In summary, the evidence in the scientific journals indicates something far from ideal in the use of statistical methods. The responsibility for this must be partly with the teachers of these methods. Changing teaching methods is not easy or without risk because we have no assurance that what we do will improve the situation. However, the risks seem small when viewed against the consequences of accepting current conditions. You may not agree with my prescriptions, but if you will simply think about other alternatives, I will have been justified in taking your time.

Table 1. Results of an experiment involving all combinations of four amounts of A and five amounts of B

		B					
		1	2	3	4	5	A-means
A	1	0 <sup>k</sup>	21 <sup>bcd</sup>	28 <sup>a</sup>	25 <sup>abc</sup>	14 <sup>de</sup>	18 <sup>w</sup>
	2	0 <sup>k</sup>	23 <sup>abc</sup>	27 <sup>ab</sup>	19 <sup>cd</sup>	9 <sup>fg</sup>	16 <sup>wx</sup>
	3	1 <sup>jk</sup>	22 <sup>abc</sup>	24 <sup>abc</sup>	11 <sup>ef</sup>	5 <sup>hi</sup>	12 <sup>xy</sup>
	4	1 <sup>ijk</sup>	18 <sup>cd</sup>	16 <sup>de</sup>	7 <sup>gh</sup>	4 <sup>hij</sup>	9 <sup>y</sup>
B-means		1 <sup>z</sup>	21 <sup>w</sup>	24 <sup>w</sup>	16 <sup>x</sup>	8 <sup>y</sup>	

a-k means that do not bear the same superscript differ ( $P < .05$ ) by HSD test.

w-z overall means of a row or column that do not bear the same superscript differ ( $P < .05$ ).

REFERENCES

- (1) Fisher, R. A. Statistical Methods for Research Workers. Oliver and Boyd, Edinburgh, 1st ed., 1925; 13th ed., 1958.
- (2) Snedecor, G. W. Statistical Methods. Iowa State University Press, Ames. 1st ed., 1937; 8th ed., 1989.

Journal Paper No. J-13559 of the Iowa Agriculture and Home Economics Experiment Station, Ames, Iowa. Project No. 0101.