ON THE TEACHING OF STATISTICS SERVICE COURSES

BU-587-M* Walter T. Federer July, 1976

ABSTRACT

A frequently occurring phenomenon is the presentation of introductory statistics courses by subject matter specialists who are not primarily, and often only minimally, trained in statistics. The quality of such courses varies widely, as does the correctness of presentation. This must be a concern for statisticians, as it can lead to incorrect usage of statistics in published literature. The question of who should teach introductory statistics courses is addressed in this paper. A definition of a statistics service course is presented and five such courses are described in section two. The third section of the paper deals with the various kinds of subject matter specialists who have frequently taught such courses. The different kinds of courses that would ordinarily be taught by a particular specialist is presented in the fourth section of the paper. An evaluation of teaching performance of various specialists in each of the five types of service courses is given in the fifth section of the paper. A discussion of various disadvantages and advantages of having statisticians and nonstatisticians teach statistics service courses is presented in the last section.

---

* In the Mimeo Series of the Biometrics Unit, 337 Warren Hall, Cornell University, Ithaca, New York - 14853.

ON THE TEACHING OF STATISTICS SERVICE COURSES

BU-587-M* Walter T. Federer July, 1976

1. INTRODUCTION AND SUMMARY

A frequently occurring phenomenon is the presentation of introductory statistics courses by subject matter specialists who are not primarily, and often only minimally, trained in statistics. The quality of such courses varies widely, as does the correctness of presentation. This must be a concern for statisticians as it can lead to incorrect usage of statistics in published literature. The question of who should teach introductory statistics courses is addressed in this paper. A definition of a statistics service course is presented and five such courses are described in section two. The third section of the paper deals with the various kinds of subject matter specialists who have frequently taught such courses. The different kinds of courses that would ordinarily be taught by a particular specialist is presented in the fourth section of the paper. An evaluation of teaching performance of various specialists in each of the five types of service courses is given in the fifth section of the paper. A discussion of various disadvantages and advantages of having statisticians and nonstatisticians teach statistics service courses is presented in the last section.

2. DEFINITIONS AND KINDS OF STATISTICS SERVICE COURSES

A statistics service course is herein defined to be a college of university statistics course given for credit and with less than 50 percent of the enrollees

---

* In the Mimeo Series of the Biometrics Unit, 337 Warren Hall, Cornell University, Ithaca, New York, 14853.

being statistics majors. It is realized that this definition is arbitrary, but it does form a base for the ensuing discussion. Statistics service courses can be, and are, of many kinds, and in discussing who should teach these courses, it is essential to indicate the kind of course under discussion. Although there are many kinds, only five of the more prevalent are discussed below.

Course I. Introductory Statistical Methods

Perhaps the most frequently presented type of statistics service course is one on statistical methodology for performing statistical analyses on results from an investigation, whether it be an experiment, a survey, or an observational study. The general contents of these courses would be a part of the material presented in the G. W. Snedecor and W. G. Cochran text entitled Statistical Methods. The particular course content and emphasis depends upon students' subject matter fields and interest. Students majoring in agriculture would have measures of location and dispersion, regression, correlation, one-, two-, and perhaps three-way analyses of variance, analyses for completely randomized, randomized complete block, and latin square designs, multiple comparisons procedures, hypothesis testing, and perhaps other topics such as nonparametric and covariance. There would be no calculus prerequisite.

Courses for biologists would probably substitute more material on chi-square, probability, and bioassay for some of the above topics. Courses in business statistics would sometimes place more emphasis on sampling, hypothesis testing, Bayesian procedures, and operations research aspects at the expense of topics for agricultural students. City and regional planning students would have some combination of the above, while engineering students would have a statistical methods course utilizing their calculus background; this course is usually a

combination of statistical methods and the probability and statistics course to be described subsequently.

In education and social science, more emphasis is sometimes placed on such topics as measurement, factor and latent structure analyses, path coefficients, and computational aspects using the SSPES or some similar computer package.

Thus, a statistical methods course will tend to describe and illustrate the statistical procedures employed in a subject matter discipline. The person developing such a course must be aware of the procedures in common use in the major fields of his students. Otherwise, there will be a gap between the statistics courses and the procedures utilized by a particular discipline. Services courses taught by statisticians would tend to be less variable than those taught by specialists whose major field was not statistics.

Course II. Introductory Probability and Statistics

Probably the second most popular type of statistics service course is one which has a calculus prerequisite and generally follows the material in one of the texts by, for example, Mood and Graybill, Hogg and Craig, Hoel, Freund, Larson, Lindgren, and several others. This type course is heavily weighted toward probability and the mathematics of statistics with little emphasis being placed upon applications of statistical procedures to results from real life investigations. The emphasis is on mathematical manipulations and this is where a student finds out if he understands calculus or not. This is in contrast to Course I, which places heavy emphasis on number manipulation. Both courses can contain considerable theory, but sometimes the mathematical courses are designated as theory courses when, in fact, they contain less statistical theory than the so-called applied statistical methods.

Course III. Introductory Mathematical Statistics

A third type of service course is one taught primarily for students majoring in mathematics. This course will be given in a mathematics department by a member of their staff, whether he be a mathematician, a probabilist, or a statistician. Quite often additional courses beyond calculus serve as a prerequisite for the course. The goal of the course may be to illustrate how mathematics is useful in probability and statistics, with much of the material from Course II above being included.

Course IV. Data Collection and Statistics of the World We Live In

This type of service course caters to lower division undergraduates who wish to learn about the nature and concepts of data from Real World investigations. The emphasis is on the concepts of populations, measurements and measuring instruments, variation, expressions of variation, variability, sampling and data collection procedures, steps involved in an investigation, statistical design, data summarization (using graphs, tabular arrays, figures, measures of central tendency, and measures of variation), patterned variations such as the binomial, Poisson, and normal distributions, confidence interval estimation, sample size, federal and international data collection agencies, and some elementary statistical methods. These students obtain first-hand experience and information about statistics by conducting surveys and/or experiments and by summarizing the results from their investigations.

Mathematics sufficient for entrance into a university or college is the only mathematics requirement of the course. The type of student enrolled in this course is not interested in statistical methods per se and is more interested in the nature and uses of statistics in more general terms. He desires a more liberal rather than more technical exposure to statistics such as given in Courses I and II above.

Course V. Introductory Data Analysis

This is a relatively new type of statistics service course, and has a noncalculus mathematical prerequisite. Emphasis is placed on data analysis rather than on data collection methods as in Course IV. Through a study of residuals, re-expressions, and resulting interpretations of the data, using methods developed by J. W. Tukey and coworkers, the student is made aware of patterns and the nature of variation present in a particular expression of a set of data. Although the methods were developed for a general investigation rather than an investigation in a particular subject matter field, the interpretations are in line of realism in the particular subject matter area.

Five types of statistics service courses have been described above. Several other types are possible. Courses with special emphases of certain statistical procedures have been developed for students of medicine, forestry, plant breeding, animal science, plant pathology, history and government, etc. Many of the statistical procedures described in Course I will be included, but some special statistical procedures may be incorporated.

3. SPECIALTIES OF TEACHERS OF STATISTICS SERVICE COURSES

The training of individuals who teach statistics service courses is extremely varied. Their formal training in statistics ranges all the way from zero courses to a Ph.D. in statistics. Likewise, the training of statisticians may be exceedingly varied with a new Ph.D. being required to teach Course I without ever having had a course in statistical methods, at least at the Ph.D. level. Some may have had considerably more training in computer science than in statistical methods and data analysis. Some subject matter specialties of

teachers of statistics service courses are:

Statistics	- statistical methodology - mathematical statistics
Probability	
Mathematics	
Education	
Agricultural Science	- agronomy - animal science - forestry - plant science
Biological Science	- ecology - entomology - genetics - veterinary medicine
Medical Science	- clinical psychology - epidemiology - pharmacology
Physical Science	- biochemistry - chemistry - computer science - engineering - operations research - physics - soil science
Social Science	- business - economics - psychology - sociology - urban and regional planning

4. SPECIALTY VERSUS TYPE OF COURSE

Since there are several types of statistics service courses and many types of specialists, it should be apparent that not all types teach all courses. A two-way array of type of training of instructor (specialty) and type of course (I, II, III, IV, and V) is presented in Table 4.1. The entries (X) in the table indicate the type of course usually taught by a particular specialist. For example, a specialist in statistics would teach all types of service courses,

whereas a specialist in mathematics or probability (a nonexperimental scientist) would ordinarily not teach a statistical methods course. A number of experimental scientists would usually teach a statistical methods course (I) and could teach a data analysis course (V). The methods in the latter course are still too new to the empirical sciences to be in general usage.

	Course				
Specialty of Instructor	I Statistical Methods	II Probability	III Mathematical Statistics	IV Data Collection	V Data Analysis
Statistics	X	X	X	X	X
Agricultural Science	X				
Biological Science	X				
Education	X	X			
Mathematics		X	X		
Medicine	X				
Physical Science	X	X	X		
Probability		X	X		
Social Science	X				

Table 4.1. Type of statistics service course usually taught by various specialists.

5. EVALUATION OF TYPE OF INSTRUCTOR FOR EACH TYPE OF COURSE

There appears to be a general feeling that "just about anyone can comprehend and teach" a university level, introductory statistical methods course. The wide proliferation of this type of course in many subject matter areas by specialists in the area and the numerous statistical methods texts written by instructors attest to this fact. Furthermore, as with introductory mathematics, graduate students, instructors, and newly appointed assistant professors are often assigned to teach an introductory statistics course, frequently a statistical methods course.

The exception to using inexperienced individuals to teach a statistical methods course is when individuals continue to teach the course and make their reputation by their teaching of the course. They often teach the course until they retire, and the cycle is started all over again.

Since "just about anyone" can master the computational aspects (the manipulation of numbers) of statistical methodology, this is what tends to be emphasized in statistical methods courses; this is what is taught by many subject matter specialists in these courses. It should also be noted that the manipulation of numbers is the main preoccupation of many writers of statistical methods textbooks. It is appalling to see 90 to 99 percent of a textbook devoted to statistical computations and only one to ten percent devoted to statistical theory, concepts, design, models, and inferences. Students in these courses are able to parrott back definitions and computational procedures, but have only the vaguest notion of what it's all about.

The statistics profession itself must take much of the blame for the manner in which statistical methods courses are taught. As long as teachers of statistics obtain their populations and statistical sampling procedures with a wave of their hands and via definition (Note the impossibility of this for the unfortunate experimenter!), obtain their response model via definition (Linear models only, if you please, and only additive, homoscedastic error structures!), and remain preoccupied with hypothesis testing within the framework of their definitions and for nonsequential sampling procedures, it is highly questionable if model selection and verification procedures will become a part of the subject of statistics. Perhaps a separate field will evolve. Data analysis does tend to consider model expressions and re-expressions for the data within the framework of an experiment, but not necessarily within the subject matter area in which the data were obtained.

Likewise, there is a general feeling that an introductory probability and statistics course is one of the easiest to teach, and one which varies little over time. Depending upon how the course is taught, both opinions can be incorrect. If one were mainly interested in the manipulation of symbols at the expense of statistical theory and concepts, this course can be satisfactorily taught by a variety of specialists. Mathematicians and physical scientists, having the necessary mathematics fully under control, quite often feel fully competent to teach this course. The emphasis will usually be on the manipulation of symbols rather than on the underlying statistical theory.

An evaluation of the kind of performance that could be expected of various specialists teaching various types of introductory statistics service courses is given in Table 5.1. The symbol U stands for unsatisfactory teaching

Specialty of Instructor	Course I Statistical Methods	II Probability	III Mathematical Statistics	IV Data Collection	V Data Analysis
Statistics	S	S	P to S	S	S
Agricultural Science	P to S	U	U	P	S
Biological Science	P to S	P	U	P	S
Education	P to S	U	U	P	S
Mathematics	U	P to S	S	U	U
Medicine	P to S	U	U	P	S
Physical Science	P to S	P to S	P to S	P	S
Probability	U	S	S	U	U
Social Science	P to S	U	U	P	S

Table 5.1. An evaluation in general of abilities of various specialists to present a particular type of course. S = satisfactory, P = passable, and U = unsatisfactory.

of a particular type of course, P stands for a passable performance, and S stands for satisfactory performance. It should be noted that there are exceptions, but the rating is what would be expected in general of all specialists in a particular area teaching a particular type course; some nonstatistics specialists may, and occasionally do a better job teaching a particular course than do some statisticians. Empirical investigators will generally perform much better in a statistical methods course than will the analytical researchers such as mathematicians, probabilists, and theoretical physicists.

A satisfactory teaching performance in a statistics service course is defined to be one in which

(i) a correct presentation of the statistical theory and application of the statistical procedures is given, and

(ii) the student obtains a correct impression of the statistical procedures presented.

A spell-binding presentation by a lecturer is not necessarily an indicator of satisfactory presentation of the material in a statistics service course. A course may be very popular with students but it may be entirely unsatisfactory from the viewpoint of correct presentation.

In addition, if a student "understands" the statistical concepts, theory, and procedures, he should be able to apply the procedures to situations other than those encountered in the classroom. His knowledge should not be superficial, but should be deep enough to have a lasting effect (at least longer than the next examination). This means that the material presented will need to go considerably beyond the computational procedures and examples frequently presented in statistical methods courses.

6. DISCUSSION

From the preceding discussion it can be implied that the writer feels that statisticians should teach statistics courses. This is a correct impression. Although economists, psychologists, and sociologists, for example, feel fully competent to offer a statistics service course, the instructor is not primarily a statistician and the subject of statistics is not his major interest. Thus, it will be highly unlikely that the course will be kept up-to-date with new research results and theory. The instructor will not usually have the depth of training in statistics to understand and to apply the finer points of a statistical procedure. Another disadvantage is that students within a subject matter specialty are often required or pressured into taking the statistics service course given by their department. The department then justifies the course presentation (to their captive audience) by pointing to high enrollment figures. Over the period of several years a mathematics course presented by a nonmathematician and a statistics course offered by a nonstatistician will tend to degenerate and become out of date. One advantage of a subject matter specialist as a teacher of a statistics service course is that he may be able to collect interesting examples for students from their area.

In deciding whether or not statisticians should be the only ones teaching statistics courses, perhaps one should ask if statisticians should teach courses in economics, biology, business, psychology, sociology, etc. When a professor in Natural Resources asked me if his advisees should take a course taught by a nonstatistician in preference to one taught by a statistician, he was asked what his recommendation would be if a statistician taught a course in Natural Resources and a similar course was taught by his department. He had little trouble making

up his mind as to where to send his advisees. A good counter weapon might be to offer an introductory subject matter course by whatever department offers a statistics service course and then use teaching load to justify a larger statistics faculty. A better approach might be to obtain administrative support to have statisticians teach the less satisfactory courses already being given and gradually mold them into satisfactory statistics courses. Close cooperation with the subject matter specialists should be maintained. Their input can be via illustrative examples and problems, or by participation in a part of the course. The participation may be in handling a discussion section which contains students from his area or by offering the lectures for a part of the course.