Introductory Statistics Teaching at UCLA
Overview and Proposal

Jan de Leeuw
Donald Ylvisaker
Richard Berk

June 12, 1997

Contents

Chapter 1

Introduction

We begin by arguing that the problems UCLA is having in organizing its statistics teaching and research are unique neither in time nor in space. Statistics is merely a prominent example showing how difficult it is for interdisciplinary programs to exist and prosper in universities structured along traditional departmental lines. There are many universities that espouse interdisciplinary and multidisciplinary teaching and research, but too often the administrative structures and the flow of the funding reflect conventional priorities.

In order to illustrate how conventional arrangements have affected statistics teaching and research on this campus, we review below a number of articles and proceedings from conferences. We discuss them briefly, and enclose some of them with this report.

To anticipate, our analysis clearly reveals one key question:

- How does statistics relate, both academically and administratively, to the other departments on campus ?

This question has three parts, which arise from both historical and organizational circumstances.

- How is introductory statistics teaching organized and administered ;
- What is the relation between statistics and mathematics ; and
- What is the role of an inter-disciplinary and cross-disciplinary emphasis in statistics graduate and research programs ?

In this report, we concentrate on the first part, introductory statistics teaching. In fact, we concentrate almost exclusively on lower division statistics, now taught in courses such as Statistics 50, Economics 40, Sociology 18, Psychology 41, and Political Science Sci 6. It is impossible, however, to discuss introductory statistics in isolation, and to ignore the larger problem of the place of statistics on campus.

1.1 Fifty Years Ago

In the February 1988 issue of *Statistical Science,* there is a reissue of two classic articles by Harold Hotelling from, respectively, 1940 and 1949. The first one is "The Teaching of Statistics" and the second one "The Place of Statistics in the University" [5], [6], [7]. Both topics were already seen as problematic at the time. It is remarkable, in fact somewhat eerie, to review the list of questions posed by Hotelling in 1940 (taken literally from the reprinted version, page 65, only the itemizing is ours).

1. Should statistics be taught in the department of agriculture, anthropology, astronomy, biology, business, economics, education, engineering, medicine, physics, political science, psychology, or sociology, or in all these departments ?

2. Should its teaching be entrusted to the department of mathematics, or to a separate department of statistics, and in either of these cases should other departments be prohibited from offering duplicating courses in statistics, as they are often inclined to do ?

3. To what students, and at what stage of their advancement, should a course in statistics be administered ?

4. Should there be mathematical or other prerequisites ?

5. How much of an investment in a statistical laboratory is warranted ?

6. Should courses be primarily theoretical and mathematical, or should they be made as practical as possible, equipping the student in the shortest possible time for a job as statistician, or for statistical work in the field with which a particular department is concerned ?

7. What about degrees in statistics ?

8. Eclipsing all these in importance, though it seems to have received too little of the attention of college and university administrative officers is the question, What sort of persons should be appointed to teach statistics ?

In the discussion of Hotelling's paper, David Moore remarks that his first reaction to rereading the papers was distress that the situation has changed so little over the intermediate 50 years. His second reaction was that we have mostly ourselves to blame. It is also possible to take a somewhat less fatalistic attitude. It seems that these basic questions must be answered again and again, at the very least, when another group of statisticians is planning to build a new department of statistics in a large university. There are no correct answers to these questions, but they provide the important dimensions along which any organized group of statisticians must define itself.

1.2 Statistics and Mathematics

David Moore gives a more extensive discussion of one aspect of the teaching of statistics in a paper with the provocative title "Should Mathematicians Teach Statistics ?" [11]. Not surprisingly, the answer is a resounding **NO !** Again, we list the main points from the paper by Moore. He makes a very strong case for moving statistics out of the mathematics departments, and for making statistics teaching problem-driven and data-oriented.

1. Statistics does not originate within mathematics.

2. The aims and foundational controversies of statistics are unrelated to those of mathematics.

3. The standards of excellence in statistics differ from those of mathematics.

4. Statistics does not participate in the interrelationships among subfields that characterize contemporary mathematics.

5. The perception of statistics as mathematics is in part an illusion due to the isolation and narrow training of many theoretical statisticians.

6. The mathematical theory of statistics is of secondary importance in teaching.

7. Graduate training in mathematics is no more sufficient for teaching statistics than for teaching economics.

8. Satisfactory teaching of the science of data requires experience with data.

9. Introductory courses that contain mathematically false statements but require students to work with data are less damaging than courses consisting solely of correct proofs of true theorems.

In the last ten years, statistics has indeed moved away from mathematics, partly because of advances in computer technology, and partly because of an ever-increasing avalanche of data that have to be analyzed. Telling indicators are recent changes in the two most prominent statistics journals. The *Journal of the American Statistical Association* was split, internally, about ten years ago, into a theoretical and applied section, with two separate editors. The *Annals of Mathematical Statistics* developed, around the same time, into two different journals, the *Annals of Probability* and the *Annals of Statistics*. Recently the *Annals of Applied Probability* was created, and very recently it was decided that the *Annals of Statistics* needed a second editor to handle (and encourage) more applied papers. The two major statistical societies also started journals directed at a more general audience, dealing often with "Statistics in Society". The Royal Statistical Society of London already produced its flagship journal in three separate series (Theoretical, Applied, General) for a long time.

Parallel with this development, modern introductory statistics books, such as the one by Moore and McCabe [13] and the one by Freedman et al. [4], have virtually no mathematics, almost no formulae, and many examples and actual data structures. Two influential papers in this context are the review by Cobb of 16 introductory statistics texts [2], and the related review paper by Singer and Willett on statistics teaching with real and representative data sets [16].

1.3 Hotelling's Problems, Revisited

That these problems continue to occupy the scientific community is clear from two recently published reports.

The first report, which concentrates on the situation of statistics in the university, is "Cross-Disciplinary Research in the Statistical Sciences", produced by a panel from the Institute of Mathematical Statistics [14], funded by NSF. We quote from the abstract.

> The panel endorses the principles that advances in substantive knowledge and in statistical theory and methods are virtually inseparable and that the continued health of statistics depends strongly on continuing cross-disciplinary research in many fields. Its report reviews past successes in collaborative research among statisticians and scientists in other fields and also important current problems that would greatly benefit from close collaboration among statisticians and others scientists to push forward the frontiers of theory, methods and knowledge. Yet the panel find that constrained resources and the existing infrastructure within government, academia and industry thwart the growth and development of needed cross-disciplinary work. The panel's report presents recommendations for concrete steps to help rectify this situation and to promote and encourage cross-disciplinary research in the statistical sciences.

The second one [15] are the proceedings of a symposium organized by the Committee of Applied and Theoretical Statistics; Board of Mathematical Sciences; Commission on Physical Sciences, Mathematics, and Applications; National Research Council. It concentrates mostly on graduate education, but for our purposes the strong emphasis on interdisciplinary education and training is especially relevant.

1.4 The Berkeley Model

1.4.1 History

The University of California at Berkeley has one of the strongest statistics departments in the world. It is of some interest, perhaps, to see how that department was established, and how it relates to the other parts of the university. The following quotations are taken from E. L. Lehmann's "reminiscences" [10]. Key sections for us include his comments on Jerzy Neyman.

Neyman was 44 when in 1938 he accepted the offer to start a statistics program in the Berkeley Mathematics Department.

A more ambitious development of courses and faculty to teach them had to wait until the end of the war but eventually led to the creation during the decade 1946-1956 of the Berkeley Statistics Department, which became one of Neyman's principal American achievements. For many years, it was the leading department of theoretical statistics in the country. Its curriculum sets the standards that were followed by many others; it trained hundreds of Ph.D. students from all over the world.

The year 1947 brought Neyman two great victories. From the beginning, he had envisaged an independent statistics department, separate from the mathematics department. ... Evans (Griffith C. Evans, chair of Math) adamantly opposed a separate department since he believed in a "greater mathematics department" that would include all mathematical sciences. A compromise was now reached that left statistics within the mathematics department but with a separate budget that no longer required Evans' approval, and with the right to make its own research appointments although Evans would still have a say on teaching appointments.

Neyman's wish for a completely independent Department of Statistics had to wait a few more years. When Evans retired in 1954, his successor Charles Morrey had no desire to keep a substantial group of statisticians in the Mathematics Department against their wishes and recommended the creation of a separate statistics department. Thus, in 1955 Neyman finally obtained his own department which consisted of seven tenured faculty members (), three tenure-track assistant professors () and several lecturers and visitors

The particular charge of inordinate expense in terms of number of students taught, although true at the time it was made, lost its validity as gradually the new department took over the teaching of all lower division statistics courses and as a result soon regularly taught statistics to almost 5000 students a year.

There are clearly some interesting parallels between the situation at UCB then, and the situation at UCLA now. Of course the differences are at least as large as the similarities. Almost all major research universities now have substantial departments of statistics, there are many more undergraduates taking statistics courses, and the move of statistics away from mathematics has progressed considerably.

1.4.2 Implementation

At UCB the statistics department teaches almost all introductory statistics courses (psychology has its own statistics-like course). In both semesters the department teaches three sections of Statistics 20 and three sessions of Statistics 121. Two of the six sections are large (around 200 students), the other four are much smaller (30-60 students). Statistics 20 corresponds to our Statistics 50, while Statistics 121 is the introductory statistics course for business school students. It requires some calculus.

The Statistics 20/121 courses have three hours of lecturing by the instructor, 1-2 hours of TA sessions, and four times per semester a computer session of 1-2 hours. TA office hours are organized in such a way that a "technical TA" and a "computer TA" are always available. There is one TA per session, the larger sessions usually have two TA's.

There are two computers labs, each with about 30 X-terminals. Throughout, the BLSS statistics package is used, with an X11 menu-front-end.

Instructors in the courses are tenured faculty and visitors. Roger Purves is a tenured lecturer, who teaches at least one of the large sessions each year. Courses use the book by Freedman et al. [4], which is actually a direct outgrowth of the Berkeley introductory courses

Chapter 2

Situation at UCLA

In the last ten years there have been many developments in UCLA Statistics. They have been supported by the National Science Foundation, the Office of Instructional Development, and the College of Letters and Sciences, in particular by the Deans of Social and Physical Sciences. We briefly review the recent history, in particular as it pertains to lower-division statistics teaching. The changes have been driven by two different developments. Statisticians in the Department of Mathematics were trying to become more and more independent from the rest of the department, and statisticians in the social and behavioral sciences were trying to reorganize their graduate teaching.

2.1 History

In 1986 the six statisticians in the Department of Mathematics obtained a certain degree of autonomy, and a Division of Statistics was formed in the Department. Around this time they decided to write a letter to provost Orbach, proposing the establishment of a Department of Statistics at UCLA (appendix C in [3]). For more information on the relation between Statistics and Mathematics at UCLA we refer to the 1990 presentation of Ylvisaker [18].

In 1987, the Advisory Committee for the Program in Social Statistics, was given a charge to

- construct a new interdisciplinary graduate program within the Division of Social Sciences, and

- evaluate and discuss the status and organization of statistics on the UCLA campus,

The report presented to the Dean in 1988 has a detailed proposal for a graduate program in Social Statistics [3]. However, the situation in undergraduate statistics teaching became another primary focus. Support was obtained through the Office of Instructional Development, and various inventories and evaluations were started (cf. below).

In 1990 the Dean of Social Sciences appointed a committee with the charge to survey the teaching of introductory statistics on the north campus and to make recommendations for improving its organization and instruction. The existence of the committee has been particularly beneficial in bringing together people who share the same fundamental problem. In once-a-quarter meetings, the committee discussed the similarities and dissimilarities of philosophies and the ways in which courses have been evolving. Our proposals is partly based on the committee's 1993 report to the Dean [17].

2.2 The Problem and its Magnitude

An inventory of prerequisites, emphases and enrollments in the five large introductory courses on campus - Economics 40, Political Science 6, Psychology 41, Sociology 18 and Statistics 50 - was made in Kreft [8]. That report also touched on which majors took which courses and how they performed, who taught the courses, and how the instruction was perceived by students. Inventories of the syllabi, preferences of the students, and goals of the instructors, were investigated by Kreft and Braverman [9] and Braverman [1].

Smaller introductory courses, such as Anthropology 80 and Geography 40, have not been studied in the same way. They will also not be discussed here, although it is clear from the point of view of resource allocation, that they should be included in a campus-wide plan. The same is true for the upper-division courses taught in the Schools of Medicine and Public Health, and perhaps even for some of the beginning graduate courses in some of the professional schools.

As background, it is instructive to consider enrollments in the large statistics courses for the three-year life of the introductory statistics committee with the three years preceding this period. We give enrollment and number of sections.

	1987-1990	1990-1993
Economics 40	1864/18	958/15
Political Science 6	649/8	385/7
Psychology 41	1530/17	801/9
Sociology 18	1582/13	1596/14
Statistics 50	3515/22	4408/25

One sees a decided shift away from Economics 40, Political Science 6 and Psychology 41, with close to half the difference showing up in Statistics 50 while Sociology 18 holds constant. The meaning of these changes may be addressed in part as we discuss students below. For completeness we also give the 1993-1994 figures, per quarter.

Subject	Fall 1993	Winter 1994	Spring 1994	Year
Economics 40	179/2	107/1	95/1	381/4
Political Science 6	89/1	77/1	103/1	269/3
Psychology 41	126/1	93/1	53/1	272/3
Sociology 18	194/1	158/1	92/1	444/3
Statistics 50	0/0	531/3	539/3	1070/6
Total	588/5	966/7	882/7	2436/19

We note that the Political Science department now requires a statistics course of their majors, therefore overall enrollment of Political Science 6 has increased. Also, Statistics 50 was not taught in the fall quarter, which increased enrollment in the other courses (mostly in Economics 40 and Psychology 41).

2.3 A Basic Difference

From the discussions of the committee it became clear there are two very different types of entry-level students:

- those who come to Statistics because they have specific subject material requirements in their majors, and
- and those who do not.

The first sort is illustrated well by students in Economics and Psychology. In Economics, it is useful to spend a good deal of time on regression analysis. In Psychology, the analysis of variance is the critical material. Anthropology, Geography and Political Science students need material tailored for them as well, but less directly. On the other hand, the general quantitative requirement students have no special agenda and none need be imparted to them. Rather, it is sound to provide a general introduction to statistical reasoning with many practical applications.

Statistics 50 has been paying increasing attention to the second type of student. The course brings out the ubiquity of statistics in everyday life. Graphics and tables are exploited to suggest ways in which one may make sense of complex data. Case studies are regularly drawn from newspaper articles for discussion. The nature of anecdotal evidence, observational studies and the design of experiments is discussed. Sampling is given attention, leading to the language and elementary properties of probability. Inferential approaches to problems are begun in basic settings. Of the more complex material, bivariate regression and correlation are discussed at some length. The purpose of the course is to broadly reduce statistical illiteracy.

One of the problems any statistics program has to deal with is that no single course can satisfy the needs of both groups of students. This is true, because the two groups of students will generally differ in terms of mathematical background, the departmental requirements for both groups are different, and the amount of course material is much to large to fit into a single one-quarter course.

2.4 The Future

Since 1993 UCLA has a Interdivisional Program in Statistics, supported by the Deans of Physical, Social and Life Sciences, and by the Provost of the College. The program runs UCLA Statistical Consulting, and it provides a working environment for the statistics graduate students in mathematics and in some of the social sciences. It also maintains the statistical infrastructure on campus (computing, courseware for graduate courses, preprints, announcements, mailing lists, ftp and gopher servers).

Actively participating in the program are the faculty from the Division of Statistics in mathematics (Ylvisaker, Jennrich, Ferguson, Li), from the Social Statistics Program (De Leeuw, Berk, Mason, Emigh, Reiff), as well as from other schools and departments (Muthén, Bentler). Temporary and visiting positions in statistics (Brown, Gould, Lin) are also considered to be part of the program. If the program develops, and more office space becomes available, these faculty will move all or part of their teaching and research time into the program, while maintaining strong ties with their home departments.

The current reorganizations on campus, and the budget crunch, will make it necessary to reconsider the relation between the statistics program in the College of Letters and Sciences, and the statistics programs in the professional schools, of which Biostat/Biomath is of course the most important one. We have good working relations with all of these programs, and we share various resources, but there are no plans for further integration. The same reasoning applies to the new School of Public Policy, which of necessity will be a heavy user of statistics. It seems to us that the Interdivisional Program could provide or coordinate or support introductory statistics teaching for the professional schools as well.

2.4.1 The Buck does not Stop Here

Our group, in its former incarnations as Social Statistics Program or Division of Statistics, has been for years drafting proposals to modernize and reorganize introductory and graduate statistics teaching in the College. Some of the proposals have actually been implemented, in an incidental way, in Statistics 50, Sociology 18, Social Sciences 40, and so on. But ultimately we cannot require people to enroll in our version of these courses, and we cannot forbid departments to teach their own versions. The only thing we can ask for is a level playing field. Students must be free to choose which statistics offerings best meet their needs, both as part of their general education requirement and as part of their major.

In particular, if we design alternative versions of the introductory statistics courses, then they should be supported as legitimate alternatives for each department's major. If the college prefers our coordinated approach to introductory statistics teaching, then it could support us, for instance, making Statistics 50 the general quantitative requirement, and by making sure that we are supported at a level at least commensurate with that of the departmental courses. We are quite confident (as the history of the enrollments shows) that the students will ultimately decide in our favor, if they have the possibility to choose freely.

2.4.2 A Worst Case Scenario

If something close to our proposals is not adopted, nothing much will change, at least initially. We basically go back to the situation in 1987, except there are now far fewer resources, the courses have grown, and fewer tenure-track statistics faculty teach statistics. We could not expect the large amounts of extra effort faculty have invested in trying to improve introductory statistics teaching to continue. Thus, introductory statistics teaching at UCLA, which is currently disorganized and of an uneven quality, will deteriorate further. At the same time, the importance of statistics, in the sense of working with quantitative information, is increasing everywhere. All major universities have Department of Statistics, in fact **all** universities in the UC system have either statistics departments or independent statistics programs. The disappearance of a coherent statistics research and teaching program will be a major loss to UCLA.

Chapter 3

Recommendations

This chapter consists of two parts. The first part proposes a reorganization of the lower-division introductory statistics teaching, the second part discusses the resources which are necessary to implement this reorganization.

3.1 Courses

3.1.1 Core Courses

We propose to split up introductory statistics teaching into two parts or streams. The first part, which satisfies the general quantitative requirement imposed by the university, and which teaches mostly "quantitative or statistical literacy", will be called Statistics 50. It is similar to the current Statistics 50 as taught by Ylvisaker or Draper, and to Sociology 18 as taught by Berk.

The second track aims to satisfy the requirements of the various departments that most heavily use statistics in their undergraduate majors. This is Statistics 51. It comes in three different flavors.

Statistics 51a Introductory Statistics, with regression and time-series analysis. This could be cross-listed as Economics 41 and/or Political Science 6.

Statistics 51b Introductory Statistics, with regression and categorical data analysis. This is mostly for sociology, anthropology, and history students. It could be cross-listed as Sociology 18. It could also be used in the School of Public Policy.

Statistics 51c Introductory Statistics, with regression and the linear model. This is for students from psychology. It can be cross-listed as Psychology 41. If suitably developed, it could also be used in the School of Medicine.

It is up to the individual departments to decide which form of the course is required for their major. Thus, we do not make Statistics 50 a prerequisite for Statistics 51. It could very well be, for instance, that economics students only have to do Statistics 51 for the major. Other departments may require both courses, and there is always the possibility to use Statistics 50 as a remedial course, even in cases where only Statistics 51 is required for the major.

We feel (but do not propose) that the College should consider making statistics a general requirement for all undergraduates. This would reflect the importance of quantitative information and quantitative reasoning throughout society. There is such a statistics requirement in at least some other large universities.

3.1.2 Time and Credits

Each of the courses Statistics 50 and Statistics 51 is a five-credit course. Each week, it has three hours of lecture time, one hour of computer time (run by the teaching assistants), and one hour of homework/exercise time (also run by the teaching assistants).

3.1.3 Honors Courses

We propose that there be a sufficient number of honors courses, running parallel with the Statistics 50 and Statistics 51 courses. These are additional 1-2 hour sessions with small groups (25 or less) students, for additional credit. The number of honors sessions can be smaller than the number of regular course sessions; in fact, one possible saving of the current proposal is that honors students can be pooled.

In these courses, students attracted by the material are brought further toward statistical literacy, in smaller classes and with extensive computer content. A Statistics 50B experiment was run for three years during the late 70's. The reasoning was much the same at that time and, while the course did not survive, we are far better equipped in the present environment.

With the support of Deans Sears and Waugh, David Draper also taught a course of this type under the heading of Social Science 40. Draper is a truly exceptional teacher and following him is a daunting task. Nonetheless, teaching Statistics to a small number of willing students, backed by ample computing facilities, is an especially inviting prospect to many faculty members, and participating in it an interesting avenue for students to pursue.

3.1.4 Sections

We propose to have 30 introductory statistics sections each year, 20 sections of Statistics 50 and 10 sections of Statistics 51. FTE should be allocated to insure that class size be held to no more than 120 students. This means we can handle a maximum of 3600 students each year. Except for Mathematics and Sociology, existing introductory statistics courses currently have no more than 100 students per quarter. At the other extreme, five of seven sections of Statistics 50 taught during the 1992-1993 school year contained more than 200 students. We do not see that real quality of instruction can surface in the face of audiences of this size.

3.1.5 Curriculum

This section is still quite tentative. The details have been reviewed to some extent by the Advisory Committee, and in the reports [1], [8], [9]. Final decisions should by taken by the introductory statistics team, after due consultation with the client departments.

Statistics 50

All Statistics 50 sections use the same textbook, or textbook-like material. We suggest that a book similar to the one by Moore [12] be adopted. This is supplemented by material from the statistics data bank (cf. below), and by material which is particularly interesting because it is currently in the news. Both the Berk and Ylvisaker versions of the current Statistics 50 depend rather heavily on scanning the daily newspaper for "statistical" items that are worth discussing in class.

We propose the course covers, at least, the following topics (with keywords).

Descriptive Univariate Statistics: Graphs Histograms, Stem-and-Leaf Plots, Box-Plots, Cumulative Distribution, Density.

Descriptive Univariate Statistics: Numbers Center, Spread, Skewness, Kurtosis, Mean and Median, Variance, Quantiles and Quartiles.

Descriptive Multivariate Statistics: Graphs Scatter plot, Spinning Plot, Draftsman's Plot, multiple Stem-and-Leaf and Box plots.

Descriptive Multivariate Statistics: Numbers Regression, Correlation, Cross Tables, Chi-square.

Aspects of Design Experimental design, Sampling design, Surveys, Causality, Randomization, Clinical Trial, Control Group, Effect, Observational Study, Quality Control.

Probability Models Normal distribution and friends, probability plots, urn models for sampling, random variables as models.

Inference: Sampling Distribution Standard error of the mean, of a proportion, of a general statistic (illustrate with Bootstrap).

Inference: Hypothesis Testing One-sample test, two-sample test, 2 by 2 table, m by n table, correlation (illustrate by Permutation Test).

Statistics 51

Statistics 51 material is taken preferably from the textbooks by Moore and McCabe [13] and by Freedman et al. [4] In each of the flavors of Statistics 51 instructors will of course add their own material.

Possible Alternatives

David Draper is finishing his problem-based textbook. How heavily we shall lean on his work depends, obviously, on whether he will return to UCLA. We are developing (with others) courseware for interactive computer instruction (cf. below).

3.2 Resources

3.2.1 Coordinator

We propose that the university hire a coordinator for Introductory Statistics teaching. This is a tenure-track faculty position, on at least the associate professor level, in which the emphasis is on coordinating issues connected with the teaching of the statistics classes. The coordinator teaches two sections per year, makes a plan for course and teacher evaluation, carries out course and teacher evaluations, overviews computing, researches audio-visual developments, coordinates textbook selection, keeps track of departmental wishes, and does research (which may or may not be related to introductory teaching).

3.2.2 Committee

We propose to establish a campus-wide committee, based in the Interdivisional Program, to oversee and assist Statistics instruction at the introductory level. In principle this committee could be made up of the introductory statistics coordinator (chair), and the faculty teaching the courses, whether they have their FTE in the program or not. There should also be representation from the professional schools.

3.2.3 Instructors

We propose that the university set aside a sufficient number of additional FTE to teach 30 sections. These positions need not be tenure-track, but a certain continuity is required. The positions are in the Interdisciplinary Program, and can be converted to tenure-track positions (in an appropriate department) if necessary and proper.

There are already two FTE of this type in the Mathematics department for 1994-1995. There is also faculty in other departments that currently teach introductory statistics courses. They should have at least a courtesy appointment in the Interdivisional Program. They should comply with the common decisions about curriculum and teaching resources, and they should participate in meetings of the committee mentioned above.

Thus, we do not exclude the possibility that faculty, not currently involved in the statistics program, will continue to teach the Statistics 51 type courses appropriate for their department. We do propose to involve them actively in the coordination and further development of the courses.

We also propose that tenured senior faculty involved with the statistics program teach sections of these courses, using some sort of rotation schedule.

3.2.4 TA's

We propose to rotate TA's, as is done in Berkeley, in such a way that there are permanent office hours for a "Statistics TA" and a "Package TA". This requires (at the rate of one TA per section) a total of 10 TA positions permanently allocated to introductory statistics teaching.

3.2.5 Classroom

In large lectures especially, there is currently no possibility of teaching Statistics in such a way that students can rely on a computer and its capabilities as they come to grips with real problems. In this respect, specifically in Statistics, the campus is behind the times and suffers accordingly.

We propose that a dedicated classroom, containing at least 50 computers, should be made available to facilitate computer instruction. This makes it possible that each of the 1,000 or so students in an introductory statistics class on the north campus in a given quarter can use a computer and be instructed on its use for at least one hour a week. The computers should be either Macintosh workstations or X-terminals. They should have a dedicated server, running (free) software such as BMDP, BLSS, Lisp-Stat, SchoolStat.

3.2.6 Courseware

Data Bank

We propose that funds be allocated for the maintenance of a data bank facility that can serve instructors in any department at any level.

Over the past two years, David Draper has been setting up a file of data sets for instructional use campus-wide. This project was undertaken with the assistance of some graduate students and facilitated by grants from NSF and OID. In its wake, one is able to see the background of a problem drawn from a wide variety of disciplines, the data, possible analyses and discussion. The project is an ongoing one but the material developed is already available on the Laplace server. The Interdivisional Program is committed to banking data sets from all directions and they are acquired regularly through the Consulting Service and other sources.

Software

We have had considerable experience, in our graduate courses, with Xlisp-Stat. This is a public domain statistics package, written in a mixture of C and Lisp, which is extremely graphical and interactive. It runs on Macs, DOS, MS-Windows, X11, and Sunview. It comes with source code, and we have a close working relationship with the author, Luke Tierney of the statistics department at the University of Minnesota. We have are helping him to port the program to various non-standard UNIX and Macintosh environments. We are also working on Common Lisp and XVT versions of the package.

For the graduate courses we have written extensions of Xlisp-Stat to perform various specialized jobs, such as bootstrapping, generalized linear models, Bayesian elicitation, hierarchical linear models, and so on. De Leeuw teaches Math 285 (Statistical Computing with Xlisp-Stat) and runs the seminar Math 296L (Xlisp-Stat Users and Programmers Group). The ftp server of the statistics program is a repository of Xlisp-Stat software from all over the world, and many of our extensions are used by others.

We have started a user-friendly front end to Xlisp-Stat, which will make it possible to use the package more easily in introductory statistics classes. The front-end is completely point-and-click, and the user does not have to program in Lisp, or use the command-line at all. This project is done jointly with statistics instructors at the University of Florida and at Penn State. It will result in a hypertext-based statistics course, in which the computing is done transparently in Xlisp-Stat.

We propose that the University allocates funds, through OID or otherwise, which supports this development project. We think two graduate student positions are needed to keep it up to speed.

Bibliography

[1] A. Braverman. Goals of Introductory Statistics. UCLA Statistics Series 92, University of California, Los Angeles, 1992.

[2] G. W. Cobb. Introductory Textbooks: a Framework for Evaluation. *Journal of the American Statistical Association*, 82:321–339, 1987.

[3] J. de Leeuw. Report of the Advisory Committee for the Program in Social Statistics. July 12, 1988.

[4] D. Freedman, R. Pisani, R. Purves, and A. Adhikari. *Statistics*. W.W. Norton and Company, New York, NY, 1991. Second Edition.

[5] H. Hotelling. The Teaching of Statistics. *The Annals of Mathematical Statistics*, 11:457–470, 1940.

[6] H. Hotelling. The Place of Statistics in the University. In J. Neyman, editor, *Proceedings of the Berkeley Symposium on Mathematical Statistics and Probability*, pages 21–40, Berkeley, CA, 1949. University of California Press.

[7] H. Hotelling. "The Teaching of Statistics" and "The Place of Statistics in the University". *Statistical Science*, 3:63–108, 1988. Reprinted, with discussion.

[8] I. G. G. Kreft. An Inventory of Undergraduate Statistics Courses at UCLA. UCLA Statistics Series 82, University of California, Los Angeles, 1992.

[9] I. G. G. Kreft and A. Braverman. Evaluation of Course Content and Patterns of Student Preference and Performance. UCLA Statistics Series 91, University of California, Los Angeles, 1992.

[10] E.L. Lehmann. Mentors and Early Collaborators: Reminiscences from the Years 1940-1956 with an Epilogue. *Statistical Science*, 8(3):331–341, 1993.

[11] D. S. Moore. Should Mathematicians Teach Statistics ? *The College Mathematics Journal*, 19:3–35, 1988. With discussion.

[12] D. S. Moore. *Statistics. Concepts and Controversies.* W.H. Freeman and Company, New York, NY, 1991. Third Edition.

[13] D. S. Moore and G. P. McCabe. *Introduction to the Practice of Statistics.* W.H. Freeman and Company, New York, NY, 1993. Second Edition.

[14] Institute of Mathematical Statistics Panel. Cross-Disciplinary Research in the Statistical Sciences. *Statistical Science*, 5:121–146, 1990.

[15] Committee on Applied and Theoretical Statistics of the National Research Council. *Modern Interdisciplinary University Statistics Education.* National Academy Press, Washington, DC, 1994.

[16] J. D. Singer and J. B. Willett. Improving the Teaching of Statistics: Putting the Data Back Into Data Analysis. *American Statistician*, 44:223–230, 1990.

[17] D. Ylvisaker. Report of the Advisory Committee for Introductory Statistics Teaching. October, 1993.

[18] D. Ylvisaker. Separating Statistics from Mathematics: How and Why. Presented at a Panel on Statistics at the Mathematical Sciences Department Chairs Colloquium, sponsored by the Board of Mathematical Sciences, Washington, October 19-20, 1990.