EFFICIENT STATISTICAL CONSULTING IN THE SMALL UNIVERSITY ENVIRONMENT

David W. Smith, North Dakota State University

1. Introduction and Background

The purpose of this paper is to share some information concerning the development and organization of statistical consulting at a medium-sized, land-grant university where there is no Ph.D. offered in Statistics. North Dakota State University has a student population of 8,500 with about 1,000 graduate students. The state has a population of approximately 600,000. Note that this is a site-specific organization and can, in no wise, be advocated for other sites. However, some facets of this consulting organization may prove useful at other locations.

The first task is to detail the statistical environment in which our program must function. As of three years ago there were no statistics programs in the state and relatively few courses. There were documented cases of state government agencies unable to attract even one applicant for a statistics position. The 1978 Directory of Statisticians lists 12 members of statistical societies in the entire state of North Dakota. From a statistics viewpoint, the situation was primitive.

In the past two years, several positive changes in the situation have occurred. First, a Master of Science in Applied Statistics was approved for the Department of Mathematical Sciences at NDSU in the spring of 1981. Second, our consulting center in the Department of Mathematical Sciences has grown to a fair size and shows every sign of a continued and sustained growth. Thus, we are now in a position to begin producing master's level statistics students with a heavy emphasis on Applied Statistics.

Often a successful new enterprise is the result of an almost random juxtaposition of circumstances, personnel, and technology. It seems appropriate to point out that one of the most helpful bits of technology is SAS. The single characteristic which makes SAS so useful to us is the wonderful property of allowing researchers across the entire spectrum of computer abilities to obtain meaningful analyses of data.

2. Components of the Present Statistics Program

Our first premise is that the practice of consulting alone, instruction alone, or research alone is not very useful. The statistics program is viewed as an integrated program in which these three parts are interdependent.

In general, teaching and consulting are essentially like those at other institutions with differences lying in the emphasis and methods.

Instruction falls into two areas - service and major. Certain courses are common to both areas. Figure 1 shows the courses available. Note that most are listed as service courses. Majors also take a substantial amount of course work in this service category. The majors category shows that the core of the masters program consists of three quarters of theory followed by two quarters of linear models. In addition, each student must take consulting practicums which presently provide about 60 hours of contact consulting. There is some thought being given to increasing this to approximately 100 hours. The goal of the practicum is to produce a student who can carry on consulting in a business or government environment. Typically, the student begins by simply sitting in consulting sessions with faculty and researchers in a passive role. Then, the students help with the organization of computing. Gradually, the student becomes more active in consulting until the point is reached where the students conduct the consulting alone with faculty as backup and support.

Undergraduate or Graduate

<table>
<thead>
<tr>
<th>Undergraduate</th>
<th>Graduate</th>
</tr>
</thead>
<tbody>
<tr>
<td>Simple Linear Regression</td>
<td>Probability &amp; Discrete Data</td>
</tr>
<tr>
<td>Intro. to ANOVA</td>
<td></td>
</tr>
<tr>
<td>Categorical Data</td>
<td></td>
</tr>
<tr>
<td>Intro. Nonparametrics</td>
<td></td>
</tr>
<tr>
<td>Sampling</td>
<td></td>
</tr>
<tr>
<td>Tables</td>
<td></td>
</tr>
<tr>
<td>Panels</td>
<td></td>
</tr>
<tr>
<td>S</td>
<td>E intro.</td>
</tr>
<tr>
<td>R</td>
<td>Statistics</td>
</tr>
<tr>
<td>V</td>
<td>Regression Theory</td>
</tr>
<tr>
<td>I</td>
<td>Robust Estimation</td>
</tr>
<tr>
<td>C</td>
<td>Time Series</td>
</tr>
<tr>
<td>E</td>
<td>Design</td>
</tr>
<tr>
<td>D</td>
<td>Stat. Computing</td>
</tr>
<tr>
<td>A</td>
<td>Multivariate</td>
</tr>
<tr>
<td>J</td>
<td>Repeated Measures</td>
</tr>
<tr>
<td>O</td>
<td>Design Theory</td>
</tr>
<tr>
<td>R</td>
<td>Nonparametric</td>
</tr>
<tr>
<td>S</td>
<td>Sampling</td>
</tr>
</tbody>
</table>

M Probability/Math Stat I & II - Linear Model I & II

A Consulting I, II, & III

Figure 1: Coursework

Most graduate students from other areas will select courses from the service group. The courses listed in italics are one-hour credit modules which are taught during one-third of the quarter. Depending on the needs of a specific student, a standard course such as regression or design may be taken or a tailor-made course may be constructed from the menu of modules. This arrangement was originally conceived to serve a number of diverse departments with a limited amount of teaching resources. This is the first full-year trial of the modules and it remains to be seen whether they will be accepted by the academic community.
An important component of the service instructional program is the introduction to SAS. It is assumed that the student is native with respect to computing. When finally in place, the SAS instruction will begin during the Introductory Statistics course and progress through several levels of sophistication to the Statistical Computing course. Along the way, the student is strongly encouraged to take the SAS short course operated by the computer center.

For major, extensive use is made of the matrix procedure in SAS. In linear models problems are assigned which require a theoretical development followed by an implementation using the matrix procedure to deal with data sets which are non-trivial.

With this brief look at the instructional area, we now turn to the consulting and research components of the program. All statistics faculty engage in consulting. Roughly 20% of faculty time is devoted to consulting activities. Together with the graduate student consulting, there are approximately 1.5 person-years devoted to consulting.

The research component of the program is very closely related to the consulting component. Researchers are primarily encouraged to support the consulting effort rather than to be primarily interested in "new theory". This is done by counting joint publications, which contain statistical research (i.e. the development of new statistical techniques or novel applications, and not simply analysis of data), in non-statistics journals as the equivalent of publication in statistics journals.

One of the requirements of the master's program is the submission of an article to a refereed journal with the student as author or co-author. This is in lieu of a formal thesis. Consequently, the faculty and student are constantly on the lookout for consulting problems requiring statistical development and which could possibly produce work of publishable quality.

Statistical research not directly related to consulting problems is certainly not discouraged; but it is evaluated at the same level as papers arising from consulting. The emphasis is on the application of statistics to research problems in other areas.

3. Statistics Consumers

The consumers of statistics at NDSU tend to be the traditional ones found at other universities. These include the graduate student researchers, faculty researchers, and, to a limited extent, the administration. Assistance with statistics is provided to these university personnel on a first come, first served basis.

In addition, a limited amount of consulting is done for federal, state and local government agencies. Typically, this involves a minor amount of time for advice with sample surveys. However, there are several substantial projects in areas such as estimation of wildlife populations and parental attitudes toward various aspects of public school programs.

Finally, a small amount of consulting is carried on with local business and industry. This is usually on a contract basis and involves problems such as hospital personnel surveys, data analysis of patient records, and analysis of production data for a large sugar manufacturer.

4. Facilities

Facilities for consulting are of considerable importance. We are fortunate to have excellent facilities, both physical and computational. As an example, the College of Home Economics provides a consulting and research facility which contains office space for a faculty consultant, office space for a graduate student, and two computer terminals—a CRT and a Deewriter for hardcopy. In addition, a secretary in the College of Home Economics handles client flow through appointments. A similar, though less elaborate, facility is provided by the College of Pharmacy. We are contemplating establishing a similar facility in the building housing the Biological Sciences.

Consultants spend a scheduled amount of time in these satellite consulting facilities each week. This has the advantage of grouping consulting visits together. It is also very convenient for the clients because of proximity to their offices. When the temperature outside is -20° F, this is an important consideration.

As for computational facilities, there are four large clusters of terminals connected to an IBM 370-158 plus an IBM 4341 running in parallel. In addition, many departments have one or more terminals for use by faculty and graduate students. A graphics terminal cluster is maintained by the department of Mathematical Sciences. A minimal amount of computer work is by card input. Most is done directly from the terminals mentioned above.

Software packages available include, first and foremost, SAS as well as BMDP, Stat Basic, SPSS, and other specialized programs. SAS Graph is available and easily accessed via the graphics cluster. The impact of SAS, our primary software tool, will be discussed in more detail below.

5. Personnel

The statistics faculty presently consists of positions for five people with Ph.D.'s in Statistics in the Mathematical Sciences department. Three adjunct members of the Statistics faculty are located in the Agricultural Experiment Station. One other adjunct member of the Statistics faculty is located in the federal Metabolism and Radiation Laboratory. Finally, there is an M.S. in Statistics located in the Mathematical Sciences department on an instructor's appointment. Thus, altogether there are ten statisticians engaged in consulting to some degree.

6. Funding

Presently, consulting is a free commodity for campus researchers. However, the volume has grown to the point where some alterations will have to be made.
It is likely that supported research will be required to bear the costs of statistical consulting. Unsupported research will probably still be provided with consulting on a free basis.

As mentioned above, extra-university consulting is supported by the particular organization for which the consulting is done. Funding arrangements are made on a case by case basis.

7. Evaluation

Evaluation of consulting falls into two parts. First, student consultants must be evaluated on an academic basis. Second, faculty consultants must be evaluated for promotion, tenure, pay raises, etc.

This evaluation program is currently in a very rudimentary form. It is anticipated that the level of evaluation will be raised and formalized in the near future.

Sources of information for evaluation of students will primarily be faculty observation. It is likely that clients, deans, and consulting records will be utilized in the evaluation of faculty.

8. SAS

The consulting structure described above would not be possible without SAS. Graduate students in all areas do their own computing. Faculty researchers also perform their own computing. In the majority of these cases, SAS is the software package of choice both for statistical reliability and for ease of use.

With respect to data analysis, consultants generally provide advice with the computing and help with interpretation of the results. Very little computing is done by consultants.

In addition, accounting for the consulting program is done using SAS. A program has been developed which prints tables needed for reports and which provides summaries of resources expended in all areas. These reports go to the appropriate deans and other administrators.

9. Comparison to Other Types of Programs

Since NDSU is the land-grant university in North Dakota, a considerable amount of research is mandated by both state and federal governments. The relatively small size of this university in comparison to other land-grant universities essentially limits the size of the statistics program. Thus, the concentric ring organization described by Professor Bancroft (1971) is not really applicable. Instead, we have found that we have the resources for only one ring.

10. Summary

In the past two and one-half years the statistics program has expanded dramatically. We feel a great deal has been accomplished. However, a great deal remains to be done before we can say that an adequate, integrated statistics program exists at North Dakota State University.

REFERENCE