Curriculum Guidelines for Bachelor of Arts Degrees in Statistical Science

Thaddeus Tarpey, Wright State University
Carmen Acuna, Bucknell University
George Cobb, Mount Holyoke College
Richard De Veaux, Williams College

November 6, 2000

Origins of these Guidelines. The proposed guidelines described here originated in a workshop, partially supported by the National Science Foundation, which was held in Alexandria, Virginia, on April 28-29, 2000. There were 17 representatives from larger universities with separate statistics departments, 18 representatives from smaller post-secondary institutions without statistics departments, and 5 representatives from non-academic organizations that employ statisticians. Scheaffer (2000) provides more details and a list of the participants. A major objective of the workshop was to develop preliminary curriculum guidelines for modern undergraduate programs in statistics. The first day of the workshop focused on guidelines for undergraduate degrees, with four independent teams working simultaneously. On the second day five new teams were organized. Three worked on curriculum issues of concern to smaller institutions without statistics departments; the remaining two worked on curriculum for students preparing either for graduate work in statistics or a job as an applied statistician. Two writing teams were chosen from the participants in the workshop. One was to focus on curriculum for a Bachelor of Science degree in statistics, and on minors and concentrations with a similar emphasis. Their report forms a companion to this one and is recommended as a source of more detail in three areas: (1) some history related to curriculum guidelines in statistics; (2) an account of the discussions at the Alexandria workshop; and, especially, (3) a description of core skills and concepts needed by statistical scientists. The second writing team, consisting of the authors of this report, was to focus on curriculum guidelines suitable for liberal arts colleges, and other institutions where statistics is taught in departments of mathematics or mathematical sciences. Preliminary drafts of these position papers were presented at the ASA Symposium on the Undergraduate Statistics Education Initiative (USEI) during the 2000 Joint Statistics Meetings in Indianapolis, Indiana.

Statistics and the Liberal Arts. In the spirit of the liberal arts, the suggestions for curriculum proposed here are not intended to guarantee that a graduate will have a specific and comprehensive set of job-related skills. (Guidelines for degrees with that goal are set out in the parallel report for programs leading to a Bachelor of Science degree (Bryce et. al., 2000).) Rather, the Guidelines described here are intended to parallel the sort of curricular recommendations one might see for programs in history, or psychology, or biology. Statistics can be defined as the theory and practice of constructing knowledge from empirical data. As such it is central to the Liberal Arts. A liberal arts program in statistics should pay explicit attention to what kinds of inferences from data are, and are not justified. Courses (at all levels) should illustrate and evaluate the role of data, and of statistical thinking, in the various sciences, social sciences, and professions (medicine, law, public policy, business management). Finally, courses should pay explicit attention to the different kinds of thinking involved in the practice of statistics: logical/deductive, computational/algorithmic, graphical/dynamic, and verbal/interpretive. In particular, "although statistics requires mathematics for the development of its underlying theory, statistics is distinct from mathematics and uses many non-mathematical skills. Thus the curriculum *must* be more than a sequence of mathematics courses. Faculty trained in statistics should be involved in developing an undergraduate curriculum in statistical science at any particular institution. For the undergraduate the focus should be on data analysis skills rather than on statistical theory, although some theoretical development should be required." (Bryce et. al., 2000)

Overview of the proposed Guidelines. Liberal arts colleges and other institutions with an emphasis on teaching are often the places where new courses are developed. Indeed curricular experimentation at such places is one of the main ways that new knowledge from research-oriented universities is made accessible to undergraduates. Accordingly, the Guidelines that follow have deliberately kept proposed requirements to a minimum. They are intended to illustrate by example the elements that distinguish a modern statistics program with a liberal arts orientation from other programs, while leaving plenty of flexibility. The body of our proposal is in three parts. In mathematics, we recommend that a BA degree in statistics should include five particular courses in mathematics: three semesters of calculus, and a semester each of linear algebra and probability. In statistics, rather than recommend particular courses, we recommend that the courses leading to a BA in statistics cover a small set of core topics dealing with theory and methods of data production, applied statistical modeling, and inference. Although we consider these topics essential to any major in statistics, we recognize that there are a variety of course structures that present them effectively. The final recommendation is coursework in a substantive area as a minor or concentration. In what follows, we give a telegraphic summary of each part, then provide more detail in a series of notes. A good way to start building toward a statistics major is to offer a minor in statistics if that is not already available. Guidelines for a minor in statistics are presented following the guidelines for the major in statistics.

A. Mathematics

- **Calculus I, II, III**
- **Linear algebra**
- **Probability**

Notes:

1. The **linear algebra** course should include basic introductions to matrix algebra, abstract vector spaces, linear transformations, projections in Euclidean space, and eigenvalue/eigenvector decomposition.

2. The **probability** course need not necessarily approximate the first semester of the standard two-semester sequence in probability and mathematical statistics. Good alternatives include courses with emphasis on stochastic processes or other applied probability models. However, a probability course that serves as prerequisite for mathematical statistics should emphasize connections between probability concepts and their applications in statistics, such as the way various distributions are related to random sampling from a population and data analysis based on random samples. Experience suggests that for many students, such a course is more meaningful if they have already taken at least one course in applied statistics.

3. Students considering graduate work in statistics should take **real analysis**, and are encouraged to take other mathematics courses as well. Alternatively, they could consider a minor or the equivalent in mathematics, or even a major in mathematics with a concentration in statistics.

B. Core statistics

Topics:

- **Production**
- **Modeling**
- **Inference**

Approach:

- **Real data**
- **Computing**
- **Synthesis**
- **Communication**

Notes:

1. **Core topics**
   a. **Data production** should cover
      - experimental design,
      - sampling and surveys, and
      - observational studies.

   b. **Statistical modeling** should have an applied emphasis, and should include
      - applied regression,
      - analysis of variance, and
      - models for categorical data.

   c. **Statistical theory** should include
      - estimation by least squares and maximum likelihood,
      - the logic of hypothesis tests and interval estimates.

2. The **approach** to teaching these topics should
   - emphasize **real** (not merely realistic) **data** and authentic applications;
   - include experience with statistical **computing**, both for data analysis and for simulation or modern computer intensive methods like the bootstrap, using one or more software packages of the sort used by professional statisticians;
   - encourage **synthesis** of theory, methods, and applications;
   - and offer frequent opportunities to develop **communication** skills through group work, oral presentations, and writing assignments.

Although a student graduating with a BA in Statistical Science should not be expected to have a comprehensive set of job related skills, the BA program should provide a solid base for statistical reasoning and equip its graduates with the necessary skills for becoming lifelong learners (of statistics). The core topics should provide a solid foundation of basic statistical competencies in the areas of data production, statistical modeling and statistical theory from which graduates of such programs can build upon in the future. Experience interpreting the results of statistical analyses in context and understanding the limitations of the methodology is important. Overall, the BA program should stress statistical reasoning and problem solving.

We emphasize that the program should provide ample opportunities for students to develop communication skills (verbal and written). In addition, teamwork skills and computing skills should be encouraged. Communication skills and experience with teamwork can be gained within statistics courses by requiring students to work in teams, to do oral presentations and to write reports of their statistical analyses. In addition to working with statistical software, gaining experience with computer programming by taking a computer science course would be useful, if institutional resources allow it.

3**. Course structure.** In the spirit of the liberal arts, these Guidelines are intended to be flexible so as to encourage curricular experimentation. Ordinarily, a BA in statistics should include five or more courses in statistics. The balance between required core courses and electives will depend on the institution. There are many different courses and course sequences that cover the core topics mentioned above. Here is an example of one sequence:

Introduction to Statistics
Applied regression
Statistical theory
Capstone course

An **introductory course** (see the associated position paper `First Courses for Statistical Science' by Garfield, Hogg, Schau, and Whittinghill (2000)) is often based loosely on the Advanced Placement syllabus, although we encourage structures that offer multiple entries into the statistics major. Whereas an introductory statistics course is recommended, there are successful majors that do not include the usual sort of introductory course. An **applied regression** course is recommended as an excellent vehicle for teaching core material, and such a course is described in detail in the associated position paper `An Example of a Second Course in Applied Statistics: Regression Analysis' by Halvorsen, McKenzie and Sullivan (2000). **Statistical theory** has most commonly been taught in the second semester of a yearlong sequence in probability and mathematical statistics. While that course is not unacceptable, the usual version is neither representative of modern statistical practice nor a good introduction to statistical thinking, and we encourage alternatives. For example, the applied regression and theory courses might be replaced by a two-semester sequence combining theory and applications. Finally, by a **capstone course** we mean a senior-level course involving one or more data-production-and-analysis projects intended to lead students to synthesize and apply what they have learned in their other statistics courses. A capstone course also provides an opportunity to integrate statistical knowledge with other substantive areas (see part C below). Throughout the statistics curriculum and particularly in the capstone course, students should become familiar with the ASA ethical guidelines. Of special importance are issues of confidentiality and impartiality when dealing with data collection and the reporting of results. In smaller departments where it may be difficult to offer a capstone course each year, a senior thesis project could be required instead. Hands on experience can also be gained through the use of internships and consulting practice.

.

4. There are many **statistics electives** that might be included in a statistics major. In some instances a single one of the topics listed below might be the focus of an entire undergraduate course. In other instances topics from this list might be interwoven with topics from the statistics core.

- Applied multivariate analysis
- Time series
- Generalized linear models
- Statistical genetics
- Categorical data analysis
- Statistical process control
- Mathematical statistics
- Data Mining
- Computer Intensive Methods (bootstrap, permutation methods, et cetera)

This list is not necessarily exhaustive of all possible elective course topics. It is anticipated that new topics may become appropriate as statistical methodology evolves. Depending on content, upper division courses like econometrics, which have a statistics emphasis but are taught outside a department of mathematical sciences, might well be acceptable as statistics electives counting toward the major.

C. Substantive Area.

Statistics, being a methodological discipline, must be applied in some other area. Students whose interest is primarily in applied statistics are encouraged to plan a coordinated program of courses (e.g., a minor) in an area of application such as economics or biology.

Statistics Minor or Concentration

A minor in statistics is designed to encourage students studying another discipline to gain a deeper understanding of the field of statistics than might be gained from just one or two introductory courses. The number of courses required for a minor may vary, typically from five to seven, depending on individual institutions. Since students will be undertaking the coursework for this minor in addition to requirements for a major in another area, the recommended requirements focus on courses in statistics, without including other courses in mathematics or computer science. Although some of the possible statistics electives might have mathematical prerequisites, *a student in the social sciences should be able to complete a statistics minor without taking a calculus course*. Note that courses from several departments might be allowed to count toward a statistics minor, although care must be taken that the content of each of the courses differs substantially from the others.

By a concentration, track, or emphasis in statistics we mean a collection of courses within a major (such as mathematics) that demonstrates a special focus in the field of statistics. We assume that additional courses from the discipline would be used to complete the major.

Curriculum for a Minor/Concentration

1. Core Statistics Topics (2 courses)
A sequence of two courses (for example, an introductory statistics course and an applied regression course) that cover the core topics (data production, inference, and applied modeling) as described in section B. Courses in a student's major discipline that overlap substantially with either course in the core sequence may be substituted.

2. Electives (3 to 5 courses - depending on individual institutions)
Possible electives include any courses that would qualify for the major in statistics. In particular, some electives might be courses like econometrics that are taught in other departments and have a substantial statistical component distinct from the introductory material of the core sequence. Although a probability course would not ordinarily be part of a minor with an applied emphasis, it might be appropriate for some students.

SUMMARY

A. Mathematics: Calculus I, II, III; linear algebra; probability

B. Core statistics

- **Data production (design, sampling, observational studies)**
- **Applied modeling (regression, ANOVA, categorical data)**
- **Inference (least squares, maximum likelihood, tests and intervals)**

Emphasize real data, statistical computing, synthesis, and communication.

C. Substantive Area: a minor or concentration

Acknowledgements. The authors acknowledge the contributions of the participants of the Undergraduate Statistics Education Initiative (USEI) Workshop in Alexandria, Virginia. In particular, we thank Ann R. Cannon, Bradley A. Hartlaub, Robin H. Lock, and Mary R. Parker who drafted the guidelines for the BA minor in statistics.

REFERENCES

Bryce, G. Rex, Robert Gould, William I. Notz, and Roxy L. Peck (2000). "Curriculum Guidelines for Bachelor of Science Degrees in Statistical Science: A Preliminary Proposal." http://www.amstat.org/meetings/jsm/2000/usei/curriculum.PDF

Garfield, Joan B., Robert V. Hogg, Schau, and Dexter C. Whittinghill, III (2000). "Best Practices in Introductory Statistics." http://www.amstat.org/meetings/jsm/2000/usei/usei_1st.PDF

Halvorsen, Katherine T., John D. McKenzie, and Mary M.Sullivan (2000)."An Example of a Second Course in Applied Statistics: Regression Analysis," in preparation.

Scheaffer, Richard L. (2000). "NSF-Funded Workshop Points the Way to Undergraduate Symposium," Amstat News, June 2000, no. 276, pp. 18, 19.