Statistics Program Area Study Group Report

Co-Chairs: John McKenzie, Jr. (Babson College) and Mary Parker (University of Texas at Austin and Austin Community College)

Patti Frazer Lock (St. Lawrence University)

Christopher Malone (Winona State University)

Roger Woodard (North Carolina State University)

Introduction

The world is awash in data, and there is a growing demand for decisions informed by data. In colleges and universities, request for statistics courses and statistics degrees has increased substantially. From 2005 to 2011, there has been a 40% increase in demand for elementary statistics and probability courses increased 40%; and undergraduate statistics degrees increased 62%. These and additional statistics from the CBMS and NCES can be found in [1].

How much and what statistics should be taught in a college or university department of mathematics or mathematical sciences? Mathematics departments should first ask themselves: Where will our students learn the statistics they need? Statistics courses are often offered in several departments. In this report we suggest strategies for mathematics departments to address this under various circumstances. Most of the report addresses mathematics departments in colleges or universities without separate statistics departments. In such institutions it is important for a mathematics department to work with other departments to plan for and meet the needs of their students in statistics.

Even at a college or university with a separate statistics department, a mathematics department is usually involved in some aspects of statistics education and should both stay abreast of important trends and coordinate with the statistics department. One such area is the preparation of K-12 teachers, who teach statistics in 6-12. Mathematics departments usually teach probability, and sometimes also the subsequent mathematical statistics course. It is also important for a mathematics department to provide good advice to students with mathematics majors who may be interested in statistics – a thriving career field.

How should statistics be taught now? New technology enables major changes in how easily statistics courses can be made engaging to students, in what data are collected, and in what statistical methods can be used. While these changes began several decades ago, the impact on students and on practitioners has accelerated in the past decade and even greater changes are likely in the next decade. By 2025 (or earlier!) all institutions should have revised their curricula substantially from late 20th century norms. The American Statistical Association makes, and will continue to make, its own recommendations about statistics curricula [2].

Program recommendations

It remains true that most statistics degree holders are at the master's level, with a substantial minority of doctorates. While there have been relatively few students with statistics majors at the undergraduate level, there have always been a substantial number of jobs and careers in data analysis; many of those are held by people with bachelor's degrees in other areas. As technology has enabled statisticians and data analysts to do more varied and more sophisticated analyses, which may not rely on calculus or advanced probability, it has become clear that colleges and universities need to better prepare students for jobs in data analysis and for the life-long learning they will need to progress in their careers.

Programs of study

a. ***For mathematics majors.*** Require all students with mathematics majors to take an introductory applied statistics course focusing on data analysis; for further description, see the Course Area Report on Statistics in this Guide [3].

A deep understanding of statistics requires a solid command of probability and calculus. But that command may not assure students' fluency and ease in discussing "statistics in practice" – which many will require in their careers. An applied course in statistics, with a focus on data analysis, should be an important part of the preparation of students in mathematics. Indeed, for several decades the CUPM guidelines have recommended that students with mathematics majors take such an applied statistics course. Some colleges and universities will use the same introductory applied statistics course for all students at their school. If the numbers of students and scheduling constraints allow for a department to offer a separate introductory applied statistics course for mathematics and other math-intensive majors than for the general student population, then it is worth considering what might be gained by doing so.

b. ***For applied mathematics majors*****.** Require students with applied mathematics majors to take both the introductory applied statistics course mentioned above and a second applied statistics course. A good choice would be a course with substantial coverage of analyzing multivariate statistical data.

Students majoring in applied mathematics should have a stronger statistics foundation than students with general mathematics majors because they are likely to need to deal with data and issues involving variation in data, including multivariate data. Regression, time series, and data mining courses are among the possible courses that deal with such data.

c. ***For students who want a minor or concentration in statistics.*** Work with other departments as needed, to identify and/or offer coherent sets of courses for a minor or concentration in statistics for students. Different sets of courses will be appropriate depending upon whether the student's goal is a data analysis position after a bachelor's degree, graduate school in statistics, or graduate work in another field besides statistics.

In a college or university without a statistics department, the mathematics department should undertake to provide a minor or concentration in statistics. A mathematics department need not offer all the courses involved in such a minor, but the department should lead in organizing coherent sets of courses, in statistics and in allied areas, for at least two categories of students: those whose majors require calculus and those from other areas.

The 2000 American Statistical Association Curriculum Guidelines [2] indicate that the number of courses could vary, depending on institutional requirements, but would usually be between three and seven. The following core topics are identified:

- general statistical methodology (statistical thinking, descriptive statistics, estimation, testing, etc.)
- statistical modeling (simple and multiple regression, diagnostics, etc.)
- exposure to professional statistical software.

The guidelines offer a non-exhaustive list of possible topics for other courses. Courses from other departments might count toward a minor or concentration – with due care taken to avoid undue duplication of content among courses offered by different departments.

To support any minor or concentration in statistics, appropriate student advising is essential to identify appropriate course paths for students with varying career paths: data analysis career after bachelor's degree, graduate school in statistics, or graduate work in another data-related discipline.

One challenge to address is how to handle introductory applied statistics courses taught in several departments. It is seldom productive to ask other departments to give up their courses. A better idea is to work toward agreement that students may not earn degree credit for more than one of the various introductory courses. Even where several beginning statistics courses are offered, courses are seldom offered at the "second level". This need provides an impetus for discussions among departments about what they want their students to learn in a second course, and how to accomplish that. Such consultation might also assure that all (or most) of the introductory courses will prepare students adequately for the second course.

We do not recommend that every mathematics department offer a major in statistics. However, if a department is interested in pursuing that, the first step is to work toward locating in other departments, or creating, appropriate courses to provide enough choices for a minor/concentration in statistics and then build from there.

d. ***For students with mathematics majors who want a minor or concentration in statistics.*** Provide appropriate advising to these students, depending upon whether the goal is to prepare for a graduate degree in statistics or an initial employment in data analysis after the undergraduate degree.

In addition to an introductory applied statistics course, the student should have a two-semester probability and mathematical statistics sequence. A course such as the "second" applied course mentioned above would be also useful.

Essential mathematics courses include multivariate calculus and applied linear algebra. Some work in computer programming and database management would be helpful. Real analysis will give students a solid preparation for graduate work in theoretical statistics. Encourage students interested in applied statistics to pursue substantial study in a field where statistics is used extensively.

One of the main paths to a good career with a substantial component of data analysis comes from having a major in a subject area, such as biology, and a minor or even a double major in statistics. Thus we encourage mathematics departments to work with colleagues in other departments to identify coherent interdisciplinary programs of study. Doing so may require a mathematics department to offer some courses in statistical analysis that rely on an introductory applied statistics course, but go on to address more advanced techniques in more general forms. Such courses should include both enough theory that students can assess when a technique is applicable, and exposure to fairly complex, realistic, applications.

Courses needed; some new

Statistics is the science of data—focused on obtaining useful information from data. Students in all statistics courses should be encouraged to fit the ideas and methods they are learning into that overview. The variability in the data includes evidence about the patterns that form the useful information and the "noise" that must be "seen though" to see these patterns in order to make appropriate choices among multiple possible patterns that could be used.

Most mathematics departments have been offering at least an elementary statistics course and a two-semester sequence of probability and mathematical statistics. It is important to continue offering such courses that are meeting the needs of the students.

a. **An introductory applied statistics course focusing on data analysis and following the various guidelines as described in the report from the Course Area Report on Statistics [3].**

We recommend that a mathematics department use the Course Area Report on Statistics [3] to consider whether its current introductory course would be appropriate as an introduction to data analysis for mathematics majors. Perhaps it would be more appropriate to develop a course similar to this for the particular audience of students from mathematics and math-intensive disciplines, which might be able to get to some more exciting and deeper data analysis in one semester than the general population of students in the introductory applied statistics course.

b. **A "second course" in statistics, with only a prerequisite of an introductory applied statistics course, and including substantial work with modeling and multivariate data.**

This course should not have a calculus prerequisite so that it can be taken by students in the social sciences and other statistics-intensive fields where calculus is not required. The Common Core State Standards (CCSS) for K-12 have recently been adopted by a large number of states. These include most of the concepts and procedures in the introductory college-level statistics course in the middle-school and high-school mathematics curriculum. As these are implemented, colleges should change their introductory statistics course to include more work with modeling and multivariate data and then revise their second course to allow students to work on projects with a wider variety of applied techniques, including exploring the techniques needed with Big Data.

Most introductory courses in applied statistics do not have time to go far enough into the subject for students to have much insight into how statistics is used for multidimensional problems. And yet most real-life problems have multidimensional aspects. Various texts and syllabi are available for a second course which have a substantial amount of multiple regression and then some other topics. Ideally this course will include having students carry out a complete statistical process (identify a problem and choose appropriate methods, collect the data, edit the data, process the data with appropriate technology, analyze the output, and communicate the results in writing, visually, and orally). This experience helps students move to a clearer understanding of statistics as the science of obtaining useful information from data rather than thinking of statistics as simply a collection of procedures.

c. **Other courses as needed to support a minor/concentration in statistics.**

If the mathematics department organizes a discussion with other departments about what needs their students have for more statistics education, that could lead to recognition of what courses should be developed. There is always a delicate balance between making the new course deal adequately with the particular types of application problems that are of interest to the other department(s), and yet be general enough to provide students from several different departments with a good educational experience. Some possible topics for parts of courses or a course to be included are resampling methods, design of experiments, time series, categorical data analysis, non-parametric statistics, survey sampling, Bayesian methods, database management, data mining, and a capstone experience.

d. **Continue to offer the two-semester probability and mathematical statistics sequence of**

courses.

Students heading for graduate school in statistics, and all students who want to understand the theory of statistics, need a solid foundation in probability at the level usually taught in the typical junior-level post-calculus probability course. This allows students to understand how to build and interpret probability models and to understand how the standard statistical inference techniques are based on sampling distributions of statistics. As we broaden our repertoire of statistical techniques to include Bayesian methods, resampling methods, and the various techniques used with Big Data, it is even more important for students to understand the foundations well enough to compare and contrast the meaning of the results of analyses from these techniques.

e. **Continue to offer elementary statistics courses appropriate for the audience of students who have traditionally taken such courses in the mathematics department.**

Elementary statistics courses can be designed at a variety of levels, including a course focused mostly on statistical literacy, with little work on procedural fluency, up through courses which combine statistical literacy, data analysis, and procedural fluency for a variety of types of data and a variety of methods (including resampling or Bayesian methods, as well as the traditional normal-distribution methods). Our recommendation to offer an introductory applied statistics course emphasizing data analysis appropriate for students with mathematics majors should not be taken as a reason to stop offering other elementary statistics courses which are working well. We do encourage departments to review their course periodically in light of the GAISE recommendations [4] and other guidelines about introductory courses.

Leadership and advising

a. **Designate a department faculty member to provide coordination and leadership of the statistics program and to coordinate with other departments. Give that faculty member appropriate course release time and other support.**

b. **Expect those faculty members who guide and teach statistics offerings in a department to monitor and communicate with colleagues about recommendations for content, pedagogy, and assessment from the foremost statistics professional associations.**

None of these recommendations can be effectively implemented without a substantial amount of faculty time and attention. And they cannot be efficiently and effectively implemented in a department without some level of coordination among faculty members inside the mathematics department and with faculty members in other departments, which, to a large extent, requires an identified leader or at least a committee chair. Because of the rapid pace of changes in statistical practice, it is important that the faculty members providing this leadership participate in the conversations about these topics available through the professional associations listed in this report's references. Thus, the mathematics department must identify one or more faculty members to work on this, give them adequate credit for their work as service to the department, and make clear the expectation that they will become/stay involved in the conversations in professional associations about undergraduate statistical education.

Needed skills and recommendations for delivery

In 2000 the American Statistical Association approved curriculum guidelines for undergraduate programs in statistical science [2], which covers majors in statistical science departments and minors or concentrations in other departments. Five necessary skills were identified for effective statisticians. Here is the ASA list:

1. *Statistical*: Graduates should have training and experience in statistical reasoning, in designing studies (including practical aspects), in exploratory analysis of data by graphical and other means, and in a variety of formal inference procedures.

2. *Mathematical:* Undergraduate statistics majors should study probability and statistical theory, along with the prerequisite mathematics, especially calculus and applied linear algebra. Programs for non-majors may require less study of mathematics. Programs preparing for graduate work may require additional mathematics.

3. *Computational*: Working with data requires more than basic computing skills. Programs should require familiarity with a standard statistical software package and encourage study of data management and algorithmic problem-solving.

4. *Nonmathematical*: Graduates should be expected to write clearly, speak fluently, and have developed skills in collaboration and teamwork and organizing and managing projects. Academic programs often fail to offer adequate preparation in these areas.

5. *Substantive area*: Because statistics is a methodological discipline, statistics programs should include some depth in an area of application.

The least taught of these skills are the non-mathematical skills. Often referred to as the soft skills, collaboration, teamwork, and organizing and managing projects are among the hardest skills to introduce into the classroom, even though they are necessary for success in numerous jobs. See [5] for an excellent paper on these important skills.

These guidelines refer to seven position papers and a website with some resource material for statistics undergraduate minors or concentrations, which are still relevant. They have been placed in the website which accompanies this report [6].

The “statistics reform” movement over the past two decades has advocated various changes in pedagogy and assessment to the typical statistics courses of earlier times. References at the end of this report include several main papers contributing to and documenting these changes. The following non-exhaustive list of recommendations is paraphrased mainly from the ASA Curriculum Guidelines [2] and the GAISE Report [4].

1. Stress conceptual understanding more than knowledge of particular procedures.
2. Emphasize statistical literacy and develop statistical thinking.
3. Use real data and authentic applications and require students to communicate results in context. Students should be expected to communicate orally and visually, as well as in written form.
4. Encourage synthesis of theory, methods, and applications.
5. Include experience with statistical computing, both to explore concepts and to analyze data.
6. Use technology of various types to support conceptual understanding, active learning, and communication among students and between students and the instructor.
7. Provide opportunities for active learning, in contrast to traditional lecture format.
8. Use a variety of assessments designed to improve and measure student learning.

Why students should choose a statistics minor or concentration

In a college with no statistics department, the mathematics advisers should be available to offer information to all students who want a statistics minor or concentration. Students majoring in mathematics, as well as those in many other areas, have good reasons to choose a minor or concentration in statistics. The mathematics advisers should be ready to recommend which of the courses available in their institution are most useful for each of these paths.

- A statistics minor or concentration is invaluable preparation for numerous entry-level jobs and internships in business, industry, and government, where data analysis is essential.

- A statistics minor or concentration will assist students who plan further education in biology, social sciences, business, nutrition, government, and many other disciplines in which they will be expected to perform complete statistical processes (or parts thereof), from generating a hypothesis, planning a study, and choosing a procedure to communicating the results of an analysis.

- A statistics minor or concentration prepares students for graduate study in statistics, biostatistics, or data science.

- Graduates of a statistics minor or concentration will be able to confidently understand, evaluate, and discuss the results of surveys, experiments, and observational studies generated by others.

- A student who completes a statistics minor or concentration, as opposed to just one course, will be able to recognize situations where statistical thinking is applicable, to correctly use statistical thinking, and to communicate results clearly in many real-life situations.

Sample courses and programs

Statistics courses can be found in research universities, comprehensive universities and colleges, liberal arts colleges, and other institutions. Most research universities have a separate department of statistics (or biostatistics, or both); this is seldom the case for other types of institutions. Our recommendations focus mainly on institutions without such a separate department of statistics and therefore on minors and concentrations. Many of these institutions offer some version of a statistics minor or concentration, an applied mathematics minor, or an interdisciplinary joint statistics minor. In what follows we refer to all of these as minors or concentrations.

Good examples of statistics courses and programs can be found among nationally ranked statistics departments. Readily available on the Internet are lists of departmental courses; in some cases, full syllabi are available for all or most of their courses. This report's website [6] contains information about selected statistics programs and courses with current syllabi. It also contains information about representative minor or concentrations from comprehensive universities and liberal arts colleges.

The introductory course

More articles have been written about this important course than any other statistics course. For an excellent overview example, see [7]. An example of an innovative way to present this course from a Bayesian perspective is present in [8]. Most of these introductory courses use only algebra, but some use calculus. The majority of papers in the *Journal of Statistics Education* deal with such courses. The accompanying Course Area Report on Statistics [3] includes syllabi for two possible introductory applied statistics course.

Commonly offered courses in a minor or concentration

Courses in multivariate statistics, regression, and experimental design are often included in statistics minors or concentrations. Mathematical statistics may also be included as an option. Other courses now often mentioned include time series, categorical data analysis, non-parametric statistics, and survey sampling, along with a capstone course. Most of these courses use only algebra; the exceptions are multivariate statistics, which may use linear algebra, and mathematical statistics, which uses calculus and probability. This report's accompanying website [6] contains representative syllabi from such courses.

For the first several of the commonly-offered courses in a minor or concentration, we reviewed several syllabi and picked one that seemed representative. Also included are a detailed syllabus of a capstone course at a nationally ranked statistics department and some articles that describe innovative approaches to some of these courses.

Future changes

All that endures is change. Heraclitus, pre-Socratic Greek philosopher.

Much has changed with respect to teaching undergraduate statistics since the 2004 CUPM report on statistics programs within mathematics departments [9]. Many of these changes will continue, including increasing numbers of students entering college who have taken Advanced Placement (AP) courses in Statistics, whether or not they receive advanced placement or credit. Technology for statistics education will also continue to evolve: hardware (from computer labs and personal computers to tablets and smartphones); computational software (from TI-84s and Excel to statistical software such as SAS and open-source R); educational software (from specialized packages to widely available applets); course management systems (from expensive commercial packages such as Blackboard to open-source Moodle), delivery methods (from blackboards to PowerPoint, videos, and on-line courses), and data (from small sets of artificial data to readily available large sets of real data present on the Internet).

Here are some possible changes to expect by 2025.

The Common Core State Standards for Mathematics, if successfully implemented, will change the first introductory course in applied statistics. Because the statistics content in CCSS for Mathematics Practice is essentially the same as the content of the AP Statistics course, which is similar to the current introductory applied statistics courses, college credit cannot be given for such courses. Hence there will a need for a revised introductory applied statistics course at the college level. This is especially from those states that require the CCSS for graduation. One possibility is a modeling course that assumes builds upon the current introductory applied statistics course. Possible topics include simple and multiple regression, analysis of variance (ANOVA), and logistic regression.

Massive, Open, On-Line Courses (MOOCs) may change the instruction of introductory statistics somewhat, except for elite universities and colleges who are less concerned with the cost of the current delivery model. Even if MOOCs do not drastically change education, they will lead to more flipped classrooms.

The arrival of Big Data has just begun to affect undergraduate statistics education. At a minimum there will be a need to acknowledge its presence in the introductory applied statistics course. Because traditional inferential methods cannot be used to analyze such data, there will be a need to develop additional electives, in addition to the Data Mining course already found at some institutions. One such Data Science course at Smith College [10] was recently taught as a Five-College course. Here is its description:

> "Computational data analysis is an essential part of modern statistics. This course provides a practical foundation for students to compute with data, by participating in the entire data analysis cycle (from forming a statistical question, data acquisition, cleaning, transforming, modeling and interpretation). This course will introduce students to tools for data management, storage and manipulation that are common in data science and will apply those tools to real scenarios. Students will undertake practical analyses using real, large, messy datasets using modern computing tools (e.g. R, SQL) and learn to think statistically in approaching all of these aspects of data analysis.

Other possible special-topic elective course topics – all computer-intensive – include Bayesian statistics, simulation and optimization, and advanced graphics for visualization. Courses related to the analysis of Big Data are likely to become part an interdisciplinary joint minor with other areas, such as computer science and operations research.

An additional area of potential change is in addressing the "soft skills" of collaboration, teamwork, and organizing and managing projects. Some colleges are increasing the use of teams in classes and it is expected that more will provide that experience for students. Participation in competitions is often useful in engaging students. The professional associations and some companies sponsor some competitions open to individuals and teams, which makes it relatively easy for an institution or an individual instructor to provide that experience for their students.

Between now and 2025, statistics educators will continue to work on bringing the best of these opportunities to undergraduates. The ASA is currently working on a revision of its curriculum guidelines to be available in 2015 [2]. Departments are urged to review these and any later updates as statistics educators continue to develop curricula and pedagogies to prepare students to extract highly useful information from the data of the 21st century.

Concluding remarks

This report has benefited greatly from previous work by three groups: [2], [9], and [3]. We feel that the statistics-related discussion of the "2000 MAA Guidelines for Programs and

Departments in Undergraduate Mathematical Sciences" is another valuable resource, which we have placed in the website which accompanies this report [6]. The same website contains references to two statistics-related reports for the "CUPM Curriculum Guide 2004" [9].

Finally, we acknowledge that portions of this report will be dated as soon as it is published. Hence we encourage individuals interested in how much and what statistics should be taught in a college or university departments of mathematics or mathematical sciences to be aware of these four important associations promoting collaboration and innovation in undergraduate statistical education:

- American Statistical Association (ASA) http://www.amstat.org/education/index.cfm and http://www.amstat.org/sections/educ/
- Consortium for the Advancement of Undergraduate Statistics Education (CAUSE) https://www.causeweb.org/
- International Association for Statistical Education (IASE), http://iase-web.org/
- Mathematical Association of America Special Interest Group on Statistics Education (Stat-Ed SIGMAA), http://sigmaa.maa.org/stat-ed/

References

1. Pierson, Steve, "Growing Numbers of Stats Degrees" *Amstat News*, May 1 (2013). http://magazine.amstat.org/blog/2013/05/01/stats-degrees/

2. American Statistical Association, "Curriculum Guidelines for Undergraduate Programs in Statistical Science" (2015). http://www.amstat.org/education/curriculumguidelines.cfm. First published in 2000 and due to be updated in 2015. The 2000 guidelines will continue to be archived on this website.

3. Committee on the Undergraduate Program in Mathematics, "Course Area Report on Statistics" (A separate report in this current guide.) (2014).

4. Aliaga, Martha, et al. "Guidelines for Assessment and Instruction in Statistics Education (College Report)" (Endorsed by the American Statistical Association.) (2005). http://www.amstat.org/education/gaise/

5. Higgins, James J., "Nonmathematical Statistics: A New Direction for the Undergraduate Discipline, *The American Statistician* 53 1 (1999), 1-6.

6. Parker, Mary, and McKenzie, John, "CUPM Program Area Report on Statistics Resource Website" (2014). http://www.ma.utexas.edu/users/parker/cupm-stat-resources/index.php

7. Garfield, Joan et al. "First Courses in Statistical Science: The Status of Educational Reform Efforts", *Journal of Statistics Education* 10 2 (2002). http://www.amstat.org/publications/jse/v10n2/garfield.html

8. Albert, James, "Teaching Introductory Statistics from a Bayesian Perspective", *The Proceedings of the Sixth International Conference of Teaching Statistics* (2002). http://www.stat.auckland.ac.nz/~iase/publications/1/3f1_albe.pdf

9. Committee on the Undergraduate Program in Mathematics, "CUPM Curriculum Guide 2004" (2004). http://www.maa.org/programs/faculty-and-departments/curriculum-department-guidelines-recommendations/cupm/cupm-guide-2004

10. Smith College, "MTH 292: Data Science" (2013). http://sophia.smith.edu/blog/4cbc/2013/04/01/mth-292-data-science/