American Statistical Association
Draft Guidelines for Undergraduate Programs in Statistical Science

Undergraduate Guidelines Workgroup[1]

July 23, 2014

Contents

Draft Statement: This is a preliminary draft: please do not quote or cite. Please send comments and suggestions to nhorton@amherst.edu by August 15, 2014.

1 Introduction

Statistics is an increasingly important discipline, spurred by the proliferation of complex and rich data as well as the growing recognition of the role that statistical analysis plays in making evidence-based decisions. Enrollments in statistics classes have been increasing dramatically. More students are entering college having completed a statistics class, and more students are studying statistics at the college level[2]. There is growing demand for a variety of strong undergraduate programs in statistics to help prepare the next generation of students to make sense of the information around them[3].

The American Statistical Association (ASA) endorses the value of undergraduate programs in statistical science, both for statistics majors and for students in other majors seeking a minor or concentration[4]. The ASA Board of Directors endorsed a set of guidelines in 2000[5], but much has changed since then. This document describes updated and expanded guidelines for curricula for undergraduate programs (majors, minors, and concentrations) in statistical science that account for important changes in the field. This document lays out general goals and some specific recommendations identified during our deliberations during 2013-2014. We begin by discussing principles that informed our thinking. Then we consider skills that students should develop in their courses along with a summary of key curriculum topics.

2 Background and Guiding Principles

Scientific method and its relation to the statistical problem-solving cycle All too often undergraduate statistics majors are handed a "canned" data set and told to analyze it. Students may glance at the title of the course and simply use the expected method. This often leaves them unable to solve more complex problems out of context, especially those involving large, unstructured data. Undergraduates need practice in utilizing all steps of the scientific method to tackle real research questions. This involves formulating good questions, considering whether the data are appropriate for addressing the problem, choosing from a set of different tools, undertaking the analyses in a reproducible manner, assessing the analytic methods, drawing appropriate conclusions, and communicating results[6]. Students need practice developing an overall sequential approach to statistical analysis, which requires integration of multiple methods in an iterative manner.

This scientific approach to statistical problem-solving is important for all data analysts: not just undergraduate statistics major or minors. It needs to start in the first course and be a consistent theme in all subsequent courses. Students need to see that the discipline of statistics is more than a collection of unrelated tools. Often there is more than one appropriate way to address a research question and undergraduates need to develop judgment to assess approaches and verify assumptions, including non-statistical justifications (subject matter knowledge) for evaluating a research conclusion. Students need to be aware of limitations of a variety of approaches, when a more complex analysis is not warranted, and when it is necessary to reformulate the question. Instructors need appropriate background in applied statistics and the statistical problem-solving cycle to be able to effectively teach these courses[7].

Real applications Undergraduate statistics programs should emphasize concepts and approaches for working with complex data and provide experience in designing studies and analyzing real data (defined as data that have been collected to solve an authentic and relevant problem) that go well beyond the content of a second course in statistical methods[8]. The detailed statistical content of these problem-solving skills may vary but should be tightly integrated with study in statistics, computing, data skills, mathematics, and, ideally, a field of application[9].

Focus on problem-solving Undergraduate programs in statistics should equip students with problem-solving skills that they can effectively apply and build on and extend over time. Some flexibility is needed, as a number of students plan graduate work in statistics or other fields, whereas many will seek employment immediately after their degree[10].

The increasing importance of data science Working with data requires extensive computing skills far beyond what was needed in the past. Statistics students need the capacity to make sense of the staggering amount of information collected in our increasingly data-centered world and to accurately and effectively communicate these findings[11]. Students need to learn to ask questions and understand the right questions to ask, to use a variety of computational approaches to manipulate data and find answers, and to communicate results in a comprehensible and correct fashion. This requires the development of data, computing, and visualization capacities to complement more traditional mathematically-oriented statistical skills. Therefore, it is essential that faculty developing statistics curricula and teaching courses be trained in statistics and experienced in working with data[12].

Previously, it was often sufficient for undergraduate students who had knowledge of statistical software to successfully navigate analytic tasks that were assigned to them. In addition to facility with professional statistical analysis software, students need to be able to access and manipulate data in various ways and to utilize algorithmic problem-solving. With data now taking all shapes and formats, statistics majors need to be able to program in higher-level languages[13] as well as fluently interact with relational databases. This additional need to compute with data—in the context of answering a statistical question—represents the most salient change since the prior guidelines were endorsed in 2000.

Creative approaches to new curricular needs Many programs will require considerable creativity to fully integrate additional data-related skills into the curriculum. Relationships with computer science and allied disciplines that teach applied statistics will become increasingly important. A number of data science topics need to be considered for inclusion into introductory, second, and advanced courses in statistics to ensure that students develop the ability to frame and answer statistical questions with rich supporting data early in their programs, and move towards dexterous ability to compute with data in later courses. To make room, some traditional topics will need to be dropped from the core curriculum[14].

One of the main goals of our recommendations is to ensure that undergraduate statistics students remain useful in a world with increasingly more complex data. If we don't prepare them to learn new techniques and work with various forms of data, it will be difficult for them to compete for jobs. We need to pay attention to the core foundations of statistical thinking and practice without shying away from increasingly important data science skills. Some recommendations may be difficult to implement for many programs. We recognize the hurdles and challenge at all ends of the spectrum to provide students with modern statistical experiences. For programs that are already implementing computational courses, faculty should be encouraged to share resources, make course content available, and help train the next generation of teachers and scholars. For programs who are unable to implement an entire minor or major program, we suggest that key topics or skills be added to each of their classes in the current curriculum. Additional co- and extra-curricular experiences which enhance the formal statistics curriculum should be embraced and encouraged.

Relationship with mathematics Though the practice of statistics requires mathematics for the development of its underlying theory, statistics is distinct from mathematics and uses many non-mathematical skills. Few statistics students need the mathematics used to derive classical statistical formulas, many of which are often superseded by computational approaches that are more accurate and better facilitate understanding. Students planning doctoral study in statistics, however, need a strong background in mathematics and theoretical statistics in addition to strong computing skills.

Flexibility Institutions vary greatly in the type and breadth of programs they are able to offer, but the ASA believes almost all institutions can provide a level of statistical education that is useful to both students and employers. Programs should be sufficiently flexible to accommodate varying student goals. Institutions should adapt these guidelines to meet the needs of their own students, potentially with tracks within a single program[15]. Each institution will also need to regularly review their programs over time to reflect new developments in this fast-moving field.

3 Skills Needed

Effective statisticians at any level need to master an integrated combination of skills that are built upon statistical theory, mathematics, statistical application, computation, data manipulation, and communication. Beginning students cannot be assumed to fully comprehend these myriad connections, and an appropriate developmental progression is required to develop mastery. These skills need to be introduced, developed, and reinforced throughout a student's academic program, beginning with introductory courses and augmented in later classes (ideally culminating in a capstone experience). Such scaffolded exposure helps students connect statistical concepts and theory to practice.

We have not specified a minimum number of classes (or equivalent) expected in each area, though programs need to provide preparatory, introductory, intermediate, and advanced skill development with an integrated approach. Ideally, there should be many opportunities for topics and concepts that cut across numerous classes to be referenced and integrated in multiple places within the curriculum. Programs should provide sufficient background in the following areas:

Statistical Methods and Theory: Graduates should be able to efficiently design studies, use graphical and other means to explore data, build and assess statistical models, employ a variety of formal inference procedures, and draw appropriate scope of conclusions from the analysis[16]. They need knowledge and experience applying a variety of statistical methods, assessing the appropriateness of these models, and communicating results. They need a foundation in theoretical statistics for sound analyses[17].

Data-related and Computational: Graduates should be able to program in a higher level language[18], to think algorithmically, and to design and carry out exploratory simulation studies[19]. Graduates should be able to manage and manipulate data, including joining data from different sources and formats and restructuring data into a form suitable for analysis. They should be facile with professional statistical software and other appropriate tools for data exploration, cleaning, validation, and analysis. Their statistical analyses should be undertaken in a well-documented and reproducible way.

Mathematical foundations: Graduates should be able to apply mathematical ideas from linear algebra and calculus to statistics and to set up and apply probability models. Minor programs will generally require less study of mathematics. Students preparing for doctoral work in theoretical statistics should generally complete additional mathematics courses[20].

Statistical Practice: Graduates should be expected to write clearly, speak fluently, and have developed skills in collaboration and teamwork as well as the ability to organize and manage projects. They should be able to communicate complex statistical methods in basic terms to managers and other audiences, as well as to visualize results in an accessible manner[21]. Undergraduate majors in statistics will often be hired into positions as analysts, where they need to be able to understand and communicate statistical findings[22].

Students should be able to apply statistical reasoning to discipline-specific questions. This includes translating research questions into statistical questions and communicating results appropriate to different disciplinary audiences. Because statistics is a methodological discipline, statistics programs should include some depth in an substantive area of application. Some programs might include a required second major, co-major[23], minor, or sequence of related courses to accompany the completion of a statistics degree.

4 Curriculum Topics for Undergraduate Degrees

The approach to teaching should model the correct application of statistics[24]:

- Emphasize authentic real-world data and substantive applications related to the statistical analysis cycle[25];
- Develop flexible problem-solving skills;
- Present problems with a substantive context that is both meaningful to students and true to the motivating research question;

- Include experience with statistical computing and data-related skills early and often;
- Encourage synthesis of theory, methods, computation, and applications;
- Incorporate training in professional conduct and ethics[26]; and
- Offer frequent opportunities to refine communication skills, tied directly to instruction in technical statistical skills.

Statistical Methods and Theory

Statistical thinking begins with a problem and explores data to answer key questions. Undergraduate statistics students need a deep understanding of fundamental concepts as well as exposure to a variety of topics and methods[27].

- Statistical theory (e.g., distributions of random variables, likelihood theory, point and interval estimation, hypothesis testing, decision theory, Bayesian methods, resampling methods)
- Exploratory data analysis approaches as well as graphical data analysis methods[28]
- Design of studies (e.g., random assignment, random selection, bias, and efficiency) and issues of bias and confounding in analysis of observational data[29]
- Statistical models (e.g., variety of linear and non-linear parametric, semi-parametric, and non-parametric regression models[30], model building and assessment, multivariate methods, and predictive analytics[31])

Data-related and Computational topics

Undergraduate statistics majors need facility with computation to be able to handle increasingly complex data and sophisticated approaches to analyze it. Graduates need the ability to flexibly and correctly manage and restructure data. Such skills underpin strategies for assessing and ensuring data quality as part of data preparation and are a necessary precursor to many analyses[32].

- Use of one or more professional statistical software environments[33]
- Basic programming concepts (e.g., breaking a problem down into modular pieces, algorithmic thinking, structured programming[34], debugging, and efficiency)
- Computationally intensive statistical methods (e.g., iterative methods, optimization, resampling[35], Monte Carlo methods, and simulation[36])
- Data manipulation using software in a well-documented and reproducible way[37], data processing in different formats[38], and methods for addressing missing data.

- Use of multiple data technologies, so graduates are not wedded to one, with facility to learn new technologies[39]

Mathematical foundation topics

The study of mathematics lays the foundation for statistical theory and undergraduate statistics majors should have a firm understanding of why and when statistical methods work. Undergraduate statistics majors should be able to communicate in the language of mathematics and explain the interplay between mathematical derivations and statistical applications.

- Calculus (integration and differentiation) through multivariable calculus
- Linear algebra (e.g., matrix manipulations, linear transformations, projections in Euclidean space, eigenvalues/eigenvectors and matrix decompositions)
- Probability, properties of random variables, multivariable continuous as well as discrete distributions
- Emphasis on connections between concepts in their mathematical foundation courses (primarily probability and linear algebra) and their applications in statistics[40]

Statistical Practice

Strong communication skills are particularly necessary for statistics and complement technical knowledge to ensure that results are made available in an accessible and accurate manner. Graduates need the ability to use technical skills to undertake analyses and communicate results and conclusions concisely and effectively.

- Effective technical writing, presentations, and visualizations
- Teamwork and collaboration
- Ability to interact with and communicate with a variety of clients and collaborators

Undergraduate curricula must provide ample opportunities to practice the work of being a statistician. The completion of such practice requirements in statistics can help ensure that graduates have the skills to work as practicing statisticians. Ethical issues are critical to incorporate throughout a program[41]. Whenever possible, the undergraduate experience should include one or more opportunities for internships[42], senior-level capstone courses[43], consulting experiences, research experiences, or a combination[44]. These and other ways to practice statistics in context should be included in a variety of venues in an undergraduate program.

5 Curriculum Topics for Minors or Concentrations

It is challenging to develop the capacity to be able to analyze data in the manner that we describe within the constraints of an undergraduate program that might include 10–12 courses. These issues are even more difficult to address for minor programs or concentrations, which typically feature a much smaller number of courses as part of their requirements[45].

In some cases, however, statistics minors or concentrations for quantitatively oriented students in fields such as biology, mathematics, business, and behavioral and social science or those planning to teach at the K–12 level may be more feasible than a full statistics major[46], and institutions need to design such programs to ensure that graduates posses a core set of useful skills. These programs will necessarily be more varied than major programs. The core of a minor or concentration in applied statistics[47] should consist of the following:

- General statistical methodology (statistical thinking, descriptive, estimation, testing, etc.)
- Statistical modeling (simple and multiple regression, diagnostics, etc.)
- Facility with professional statistical software, along with data management skills
- Multiple experiences analyzing data and communicating results

This recommendation focuses on statistical fundamentals, data technologies, and communication and is intended to ensure that students develop significant data-related skills, understanding of key statistical concepts and perspective on the field of statistics.

The number of credit hours for minors or concentrations will depend upon the institution. Additional topics to consider include applied regression, design of experiments, statistical computing, data science, theoretical mathematical statistics, categorical data analysis, time series, Bayesian methods, probability, database management, high performance computing, and a capstone (or similar integrative) experience[48]. Ethics is another key topic to integrate within these courses. For many students, a methods course in an application area might be an appropriate option. Courses from other departments with significant statistical content might be allowed to count toward a statistics minor or concentration. A capstone or other integrative experience is particularly useful for minors or concentrations.

6 Additional Points

Relationship with high-school and community college courses in statistics: The dramatic growth of the number of students completing the Advanced Placement course in Statistics combined with the augmented role for statistics as part of the Common Core State Standards has increased the exposure of the discipline at the high-school level[49]. College and universities may need to re-evaluate their introductory courses[50].

The number of students studying introductory statistics courses at two-year (community) colleges has increased to more than 134,000 per year[51] (larger than the total enrollment in calculus classes at this level, up from a previous ratio of ten calculus sections per statistics section in the 1960's). This shift reflects the belief that statistics is a universal discipline, not just needed for a handful of students, but required of a number of disciplines and recommended for many others.

Anecdotal evidence suggests that many statistics majors are transferring to universities from community colleges[52]. A key question is how to facilitate this transfer and ensure that students can successfully undertake preliminary coursework and general education requirements prior to completing a statistics degree at another institution. Further efforts are needed to streamline articulation agreements with community colleges as well as to support faculty development and curricular development at two-year colleges[53].

Relationship with master's programs in statistics Graduates from undergraduate programs in statistics are generally employable as analysts or in similar positions that utilize a number of statistical skills. In addition, a bachelor degree can and should be considered an attractive option as a liberal arts degree. Both bachelors and masters graduates are needed to help address the shortage of workers with the skills to make evidence-based decisions informed by data.

There are a some differences between the learning outcomes of master's programs and bachelor's programs, primarily related to level, breadth and depth[54]. There is a presumption that master's graduates are statisticians and as such are able to move directly towards accreditation or other professional recognition[55]. There may be opportunities for institutions with both graduate and undergraduate programs in statistics to facilitate access by undergraduate statistics students to curricular innovations that address these recommendations[56]. Further efforts to assist with professional development and continuing education are needed to help ensure that bachelor's graduates can stay engaged with new developments in the field.

Learning outcomes and assessment There is a growing awareness of the importance of learning outcomes (a detailed list of what a student is expected to know, understand, and demonstrate after completing a program) and assessment of these learning outcomes[57]. Many internal and external groups (such as accreditors, legislators, parents, and students) are calling upon institutions to demonstrate accountability by defining learning goals and objectives at the program level (in addition to the course level) and devising strategies for assessing whether these goals and objectives are being met.

Assessments can be structured in a number of ways. They can be direct (e.g., tests, projects) or indirect (e.g., surveys, focus groups). For higher-order thinking skills, which encompass much of a statistics program, assessments should be relevant, open-ended, and complex[58]. A sound assessment plan will include indication of where (e.g., which courses, experiences) students are expected to develop the skills, and when they are expected to be introduced to, practice, and master the skills. Further work is needed to identify appropriate learning outcomes and assessment strategies for statistics programs[59].

Closing These guidelines are intended to provide an overview of a principled approach to ensure that undergraduate statistics majors have the appropriate skills and ability to tackle complex and important data focused problems. Additional resource materials and an annotated bibliography are available at `http://www.amherst.edu/~nhorton/undergrad`.

Notes

[1] The American Statistical Association undergraduate guidelines working group was convened by ASA President Nathaniel Schenker in the spring of 2013. Members included Beth Chance, Steve Cohen, Scott Grimshaw, Johanna Hardin, Tim Hesterberg, Roger Hoerl, Nicholas Horton (chair), Chris Malone, Rebecca Nichols, and Deborah Nolan.

[2] See for example `http://magazine.amstat.org/blog/2013/05/01/stats-degrees`. While the number of bachelor level statistics graduates is still relatively small in absolute terms (1,522 according to the IPEDS (Integrated Post-secondary Education Data System Completions Survey, `https://ncsesdata.nsf.gov/webcaspar`) data from 2012 that counted first and second majors in Biostatistics, Statistics, Mathematical Statistics and Probability, Statistics [Other] and Mathematics and Statistics [Other], this number has increased markedly from 2003, when only 673 statistics undergraduate degrees were conferred. Purdue University (n=100) was the largest producer in 2012, with University of California-Berkeley a close second (n=99). Other institutions with 40 or more graduates in 2012 included University of California-Davis, University of Illinois at Urbana-Champaign, University of Minnesota, and Carnegie Mellon University. There may be substantial undercounting, since students completing a Mathematics major with a concentration in Statistics are not included in these numbers.

[3] The widely cited McKinsey & Company report stated that "by 2018, the United States alone could face a shortage of 140,000 to 190,000 people with deep analytical skills as well as 1.5 million managers and analysts with the know-how to use the analysis of big data to make effective decisions" (`http://www.mckinsey.com/insights/business_technology/big_data_the_next_frontier_for_innovation`). While some of these will need graduate training, much of the demand is expected to be at the bachelor's level (see `http://www.ingeniousmathstat.org`).

[4] We focus primarily on majors, since the development of a deep understanding of statistical science and associated computational and data-related skills require extensive study. We also describe key points related to minor programs and similar types of concentrations or tracks through other majors.

[5] See `http://www.amstat.org/education/curriculumguidelines.cfm` for the 2000 guidelines and related resources.

[6] The K–12 GAISE guidelines (`http://www.amstat.org/education/gaise`) define statistical problem-solving as an investigative process that involves four components: (1) Formulate questions (clarify the problem at hand, then formulate one (or more) questions that can be answered with data); (2) Collect data (design a plan to collect appropriate data, then employ the plan to collect the data); (3) Analyze data (select appropriate graphical and numerical methods, then use these methods to analyze the data); and (4) Interpret results (interpret the analysis, then relate the interpretation to the original question). It should be emphasized that this process is rarely sequential.

[7] There is a need for additional continuing professional development for instructors as well as revisions to the graduate curricula that will prepare future instructors. Since many faculty teaching statistics do not have a graduate degree in statistics, there is a need for creative approaches to ensure that they have appropriate background (see for example the 2014 MAA/ASA guidelines for teaching statistics, `http://magazine.amstat.org/blog/2014/04/01/asamaaguidelines`).

[8] There is not a single definition of what is appropriate as a second course, and a number of different options can be found many institutions. But we believe that it is not possible to develop a comprehensive understanding of the range of key statistical concepts after only two courses.

[9] There are many electives that might be included in a statistics major. As resources will vary among institutions, the identification of what will be offered is left to the discretion of individual units.

[10] Data from a survey of graduates from California Polytechnic State University, San Luis Obispo (Melissa Bowler, unpublished senior project) found that 60% of bachelor graduates eventually completed a graduate degree, but for many of them, this was after years in the workforce. Additional studies of graduates and their early career profiles would be valuable for the community.

[11] This has been elegantly described by Diane Lambert of Google as the ability to "think with data". See also the ASA report on "Discovery with Data: Leveraging Statistics with Computer Science to Transform Science and Society", `http://www.amstat.org/policy/pdfs/BigDataStatisticsJune2014.pdf`.

[12] This is fully consistent with the 2014 MAA/ASA guidelines for teaching introductory statistics, `http://magazine.amstat.org/blog/2014/04/01/asamaaguidelines`. We note that mathematical expertise is not a substitute.

[13] We define this as a programming environment which supports abstraction from the specification to the computer (e.g. which hides many aspects of the underlying computational environment). This could include Python, R, SAS, and other environments.

[14] We do not try to specify which topics are central, and which could be covered in electives or dropped entirely. Given that most undergraduate statistics majors enter the workforce as analysts where data-skills are primary, we suggest that helping them to master a smaller set of methods, rather than a comprehensive laundry list, is likely to be more useful to them in the long-term. If students develop a clear understanding of basic statistical theory that allows them to select, utilize, and assess a model, it is more likely that they can learn to effectively use other approaches that they were not exposed to in college (or did not exist before they graduated).

[15] See the draft white paper by Hoerl "Roadmap for Smaller Schools".

[16] Our enumeration of key statistical skills is intentionally short, since these are likely most familiar to the statistics community. More detail is provided regarding computation and data-related skills, because they have not played as large part in the undergraduate statistics curriculum in the past. We reiterate, however, that statistical fundamentals are at the core, with the data-related skills supporting the ability to analyze complex data.

[17] Understanding the theoretical underpinnings of statistical methods is a vital component of modern statistical practice. While we do not presume to specify how the ideal statistical theory (previously called mathematical statistics) course should be structured, we do believe that aspects of the traditional probability/inference sequence, with its emphasis on large sample size approximations and lists of distributions, does not capture current statistical practice. A lively panel discussion from JSM 2002 captured many of these issues (see `www.amstat.org/sections/educ/MathStatObsolete.pdf` for details). A modern statistical theory course might, for example, include work on computer intensive methods and non-parametric modeling. This course should provide students with an overview of statistics and statistical thinking that builds on what is provided in their introductory statistics courses. It may be useful to incorporate computing, data-related, and communication components in this class. If included early on in a student's program, it will help to provide a solid foundation for future courses and experiential opportunities. Because the traditional mathematical statistics course requires probability, which in turn often necessitates three semester of calculus as a pre-requisite, students often take this course late in their programs. This precludes other upper level applied statistics courses building on this important theoretical foundation. Some models have been successful in allowing students to learn the theoretical underpinnings of the discipline (e.g., the model at Brigham Young which splits the traditional probability into a course with a focus on discrete random variables (taught at the sophomore level) and a more advanced course (with a focus on continuous random variables) which can be taken later.

[18] This might consist of the ability to write functions and utilize control flow in a variety of languages and environments such as Python, R, SAS, or Stata. Facility with spreadsheet tools such as Excel is useful for a variety of other purposes, but is not ideal as a programming environment.

[19] We anticipate that the capacity to undertake and interpret simulation studies as a way to complement students analytic understanding and/or check their results will be increasingly useful in the workplace.

[20] Many graduate programs strongly recommend at least a year of mathematical analysis and/or advanced calculus while other upper-level mathematics courses such as “Graph Theory”, “Differential Equations”, “Optimization”, “Combinatorics”, “Algebraic Statistics” and/or “Abstract Algebra” may also be helpful.

[21] These skills need to dovetail with their technical and statistical knowledge: excellent communication of inappropriate or incorrect analyses is counterproductive.

[22] Data from a survey of graduates from California Polytechnic State University, San Luis Obispo (Melissa Bowler, unpublished senior project) was used to generate a listing of current jobs for n=62 graduates from Cal Poly San Luis Obispo’s undergraduate statistics program found that 12 had “statistic” in the title (e.g., Statistician, Senior Statistical Analyst, Statistical Programmer I) while 20 had “‘analy” in the title (e.g., Data Analyst, Marketing Data Analyst, Research Analyst, Business Systems Analyst). We suspect that many of those with “Statistician” in their job title completed a higher degree. Better data on the outcomes of graduates would benefit the profession as a whole.

[23] See for example the Analytics co-major in the Department of Statistics at Miami University, `http://miamioh.edu/cas/academics/departments/statistics/academics/majors/analytics-comajor`

[24] It goes without saying that just modeling the process is insufficient: students need repeated experiences undertaking analysis of real-world data. It is also important that instructors have a history of such experiences (see the 2014 MAA/ASA guidelines for teaching statistics, `http://magazine.amstat.org/blog/2014/04/01/asamaaguidelines`).

[25] While the GAISE college report (`http://www.amstat.org/education/gaise`) focuses on the introductory statistics course, many of its tenets are broadly applicable for the principled teaching of statistics.

[26] The American Statistical Association issued a statement on continuing professional development (`http://www.amstat.org/education/cpd.cfm`). Statisticians are encouraged to undertake continuing professional development to: (1) In methodology and practice, by keeping abreast of new techniques and theory, staying connected with best practice, growing in areas not previously studied (or refreshing forgotten material), and gathering ideas and direction for future research; (2) In technology, by learning about new computational techniques and software tools, and by staying on top of trends in technology and new sources of data that are creating major new opportunities for statisticians, (3) In subject matter needed for successful collaboration with other disciplines, to strengthen the interdisciplinary contributions and capabilities of statisticians, and (4) In career success factors such as communication, leadership, and influence skills, which are vital to the impact of individual contributions and the visibility of our profession.

[27] Given how quickly the discipline of statistics is changing, it is not feasible or appropriate to attempt a comprehensive overview of the entire field at the undergraduate level, and we do not attempt it.

[28] This might include advanced visualization techniques, smoothing/kernel estimation, spatial methods (see work by Christou draft manuscript “Enhancing the teaching of statistics using spatial data”), and mapping)

[29] Issues of confounding and causal inference are central to the discipline of statistics, and there are many settings where the ideal randomized experiment cannot be undertaken. To avoiding pitfalls of drawing conclusions from observational data, students need a clear understanding of principles of statistical design as well as tools to assess and account for the possible impact of other measured and unmeasured variables.

[30] This will generally include many of the following topics: simple and multiple linear regression, generalized linear models, mixed models, survival, model selection, diagnostics, and L_1 regularization (e.g., LASSO).

[31] We define data analytics as the use of modeling, machine learning, and data mining to make predictions about future events: see Breiman (2001), “Statistical modeling: the two cultures”, *Statistical Science*, 16(3):199-231. See also Harville (2014), “The need for more emphasis on prediction: a ‘non-denominational’ model-based approach, *The American Statistician*, 68(2):71-92 and related discussion as well as the draft white papers on a model data science course (Baumer) and the role of data science in the undergraduate statistics curriculum (Hardin et al).

[32] See Zhu et al (2013), “Data acquisition and pre-processing in studies on humans: what is not taught in statistics classes”, *The American Statistician*, 67(4):235-241 which includes detailed steps: (1) get to know the study; (2) assess the validity of variable coding; (3) assess data entry accuracy; (4) perform data cleaning;

and (5) edit identified data errors.

[33] While we acknowledge that Microsoft Excel is a common platform for data exchange, we do not recommend it as a primary analysis environment.

[34] We define structured programming as the ability to use functions and control structures (such as "for" loops).

[35] Resampling methods include bootstrapping and permutation tests: see the draft white paper on "Resampling Within the Undergraduate Statistics Curriculum" by Hesterberg.

[36] This is consistent with the efforts of Conrad Wolfram and the Computer Based Math initiative, see `https://www.computerbasedmath.org` and `https://www.ted.com/talks/conrad_wolfram_teaching_kids_real_math_with_computers`. The incorporation of more computing may be particularly valuable at the bachelor level, since students will generally have less technical knowledge (and need to be able to simulate to generate insights and/or check analytic results).

[37] This could include shell scripts, R, Python, and SAS code.

[38] This should include formats such as CSV, JSON (JavaScript Object Notation, a data-interchange format that is easy to read, parse, and generate: see Nolan and Temple Lang (2014) *XML and Web Technologies for Data Sciences with R*, Springer), XML, databases (see for example Ripley (2001) "Using databases with R", *R News*, 1(1):18-20 and Wickham (2011) "ASA 2009 Data Expo", *Journal of Computational and Graphical Statistics*, 20(2):281-283, and text data. Because many faculty were not trained in these technologies, continuing education in this area needs to be made a priority.

[39] We are not prescriptive regarding which technologies are incorporated into the curriculum, as long as they are sufficiently flexible and powerful. Many undergraduate statistics students develop expertise in environments such as R/RStudio, Python, and SAS.

[40] This includes topics such as the delta method. In addition, many students might benefit from exposure to modeling and simulation (empirical) in their mathematics courses as a way to reinforce their computational skills.

[41] See the white paper on "Ethics and the undergraduate curriculum" by Cohen as well as Utts (2005) "Seeing through statistics" chapter 26 (Ethics in Statistical Studies), which includes topics such as the ethical treatment of human and animal participants, assurance of data quality, appropriate statistical analyses, and fair reporting of results.

[42] See the draft white paper "Undergraduate Internships in Statistics" by Cohen.

[43] See the draft white paper by Malone "Capstones in the Undergraduate Statistics Curricula".

[44] A number of innovative programs have been created in recent years to address the need to provide undergraduate statistics students with authentic experiences with posing and answering statistical questions. These include DataFest (`http://chance.amstat.org/2013/09/classroom_26-3`), Explorations in Statistics Research (see the draft white paper by Nolan et al), and the Summer Institutes in Biostatistics (`http://www.nhlbi.nih.gov/funding/training/redbook/sibsweb.htm`).

[45] See Cannon et al (2001), "Guidelines for Undergraduate Minors and Concentrations in Statistical Science", *Journal of Statistics Education*, 10(2), `http://www.amstat.org/publications/jse/v10n2/cannon.html`.

[46] There is a pressing need for additional K–12 teachers with the capacity to teach the Common Core State Standards. See the forthcoming ASA SET (Statistical Education of Teachers) report for specific guidance.

[47] A minor in mathematical statistics could also be considered, but it may be challenging to ensure that students develop sufficiently strong computational and data-related skills.

[48] See the draft white paper by Malone "Capstones in the Undergraduate Statistics Curricula".

[49] See `http://www.corestandards.org/Math/Content/HSS/introduction/` for an overview of statistics and probability topics at the high-school level.

[50] A number of innovative approaches have been suggested for this problem. One approach is to consider a year-long introductory statistics course at the undergraduate level, the first half of which students who have completed Advanced Placement Statistics would place out of. Offering a sequence of courses (e.g., Applied Statistics I and Applied Statistics II) would facilitate integration of additional topics into that are not feasible within the syllabus of a single course.

[51] See the Conference Board of the Mathematical Sciences 2010 report `http://www.ams.org/profession/data/cbms-survey/cbms`.

[52] Data from University of California/Berkeley indicate that these numbers are large and increasing (Deb Nolan, personal communication).

[53] See the draft white paper "The Key Role of Community Colleges to Support the Undergraduate Teaching of Statistics" by Horton et al. and the California online student-transfer information website `http://www.assist.org/web-assist/welcome.html`. Revision of introductory lower division introductory statistics courses to introduce computing and data science skills will facilitate articulation in the future, however the challenge of faculty development for an increasingly adjunct workforce that is teaching in these environments.

[54] The ASA workgroup on guidelines for master's programs made a number of recommendations (`http://magazine.amstat.org/blog/2013/06/01/preparing-masters`): (1) Graduates should have a solid foundation in statistical theory and methods, (2) Programming skills are critical and should be infused throughout the graduate student experience, (3) Communication skills are critical and should be developed and practiced throughout graduate programs, (4) Collaboration, teamwork, and leadership development should be part of graduate education, (5) Students should encounter non-routine, real problems throughout their graduate education, and (6) Internships, co-ops or other significant immersive work experiences should be integrated into graduate education. We note that many of these recommendations also apply to undergraduate programs.

[55] Most undergraduate programs are not intended to train accredited (professional) statisticians, though some graduates may reach this level through work experience or further study. The American Statistical Association allows graduate statisticians not yet eligible for accreditation because of a lack of experience to be designated GStat (`http://stattrak.amstat.org/2014/05/01/gstat`). A similar pathway might be considered for those with training at the undergraduate level.

[56] In addition, the growing number of statistics graduates at the bachelor's level may have implications for the structure and content for master's programs.

[57] See the draft white paper by Chance and Peck on "From Curriculum Guidelines to Learning Objectives: A Study of Five Statistics Programs".

[58] They should be "authentic". See for example Gould (2010) "Statistics and the Modern Student", *International Statistical Review*, 78(2):297-315 and Brown and Kass (2010) "What is Statistics?", *The American Statistician*, 63(2):105-110 and associated discussion and rejoinder.

[59] See the draft white paper by Chance and Peck on "From Curriculum Guidelines to Learning Objectives: A Study of Five Statistics Programs".