

Roadmap for Smaller Statistical Programs to Incorporate 2014 Guidelines for Undergraduate Statistics Education

November 12th, 2014

Roger Hoerl – Union College

This white paper (and other related resources) can be found at <http://www.amstat.org/education/curriculumguidelines.cfm>.

Caveat: Note that this roadmap is NOT intended as the optimal approach to adopting the 2014 guidelines for undergraduate statistical education under ideal circumstances. Rather, it is intended to describe a *viable* approach to rethink a traditional undergraduate statistics program under significant practical restrictions that may exist in a smaller program. We focus on reasonable steps a school with a minimal statistics major might take.

Background: We assume that a school has a traditional undergraduate statistics program, involving courses such as introductory statistics, experimental design, regression, and perhaps probability and mathematical statistics, taught at an undergraduate level. Further, we assume that there are only two or three instructors in the department who teach statistics courses, who may not all have a Ph.D. in statistics.

Strategy: Small programs probably cannot just add more courses to meet the objectives of the 2014 guidelines. Therefore, we recommend a more feasible three-pronged approach toward meeting key aspects of the 2014 guidelines:

- Revise current courses to make them more consistent with the guidelines.
- Partner with other departments, such as computer science, to ensure availability of needed courses not offered by the statistics or mathematics department.
- Ensure that limited statistics faculty are best deployed to teach courses that only statisticians can teach. For example, mathematicians may be able to teach probability.
- Incorporate non-course activities, such as internships, seminars, and projects to help address some of the guidelines.

Recommendations: We consider each of these three areas individually.

1. Revise current courses.
Current courses could be revised to incorporate key elements of the 2014 guidelines.
 - a. For example, we would recommend enhanced emphasis in the following areas across the curriculum:

- i. Greater use of statistical software and programming languages, including more modern languages, such as R, Python, and SQL. If possible, include access to databases, and database management for large data sets.
- ii. Analysis of real rather than “textbook” data sets in methods courses. Such data would include typical problems seen with real data, such as outliers, missing values, and lack of a single “correct” answer.
- iii. Projects or extended homework problems that require an iterative approach, and cannot be solved via application of the one “correct” statistical method intended by the instructor.
- iv. Design of real experiments or surveys, which require the student to actually obtain data within the context of practical constraints.
- v. Incorporation of more modern methods, such as bootstrapping throughout the curriculum and text mining in more advanced courses.
- vi. Reconsider the traditional probability course within the context of modern technology, such as Markov chain Monte Carlo (MCMC) or other computational methods.
- vii. Incorporation of written and oral presentation assignments, to develop communication skills.
- viii. Team projects that require students to work together.
- b. A “capstone” or thesis course might be added in a manageable way, for example by working one-on-one with a faculty member, as opposed to offering another class. This would be an opportunity to develop practice in the art of statistics, incorporate data management and computing, and several other of the 2014 guidelines.
- c. The mathematical statistics course might be renamed “Statistical Theory”, as this is what it should actually focusing on. Theory is correlated with mathematics, but certainly not synonymous. This course might also incorporate:
 - i. More computing, such as maximizing likelihoods that cannot be solved in closed form.
 - ii. Both Bayesian and frequentist paradigms and methods.
 - iii. Simulation and bootstrapping

2. Partner with other departments.

Partnership with computer science departments, in particular, can be a viable approach to incorporating modern data science without having to create additional statistics offerings. In addition to potential courses in statistical computing, text mining, machine learning, or data science, computer science departments no doubt already offer courses in various programming languages and database management, which could be relevant for statistics majors. Physics, operations research, electrical engineering (if

present) and others may be potential partners. Partnerships with biology, finance, psychology, and others provide an opportunity to develop subject matter knowledge in application domains.

3. Non-course alternatives.

There are many ways to develop skills and provide experience; formal coursework is only one such avenue. Some alternative approaches include:

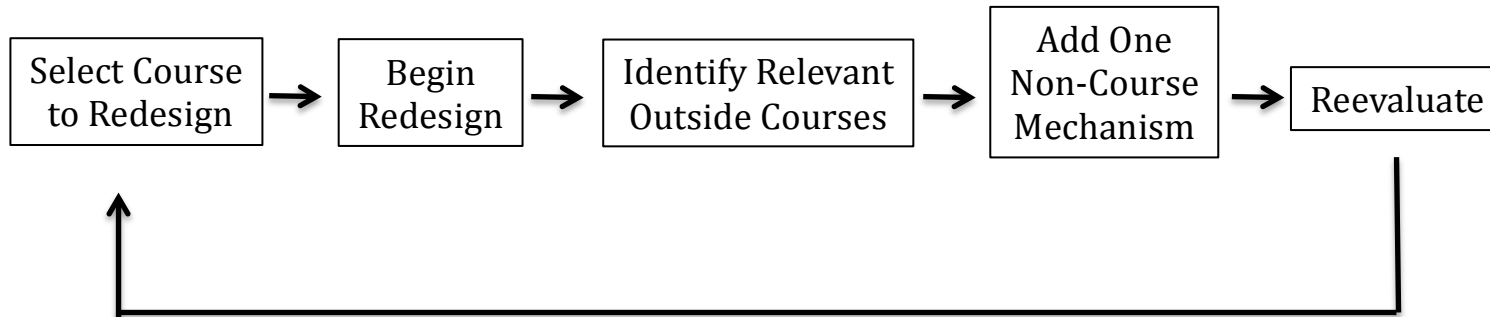
- a. Internships, possibly internal to the college (if the student does not yet have marketable skills)
- b. A seminar series highlighting non-technical skills, or specialized technical skills not covered in the formal coursework
- c. Students presentations during this seminar series, to gain experience communicating statistical results
- d. Undergraduate research projects, possibly conducted over the summer, which require extensive computations, development of subject matter knowledge, or integration of multiple statistical methods. The NSF offers the Research Experience for Undergraduates (REU) program.
- e. Involvement with statistical consulting centers to obtain experience consulting on real problems. For example, see www.math.smith.edu/~nhorton/statconsultreport.pdf. If this is not possible, consider statistical consulting workshops where students learn how to ask appropriate questions and interact with non-statisticians on an equal footing
- f. Involvement with DataFest or other student competitions, either virtual or onsite
- g. Courses at nearby institutions which may fill a gap in the curriculum

Potential approaches to implementation: We realize that implementing everything listed above in a small school may be overwhelming. Therefore, in terms of a specific roadmap, there are several alternative approaches. One potential approach would be:

- Select one course currently being offered, in which the department feels the greatest changes are possible and warranted. Focus on redesign of that course per the comments above. Move on to subsequent courses as time permits.
- Meet with the computer science, and other relevant department, to determine which courses currently offered would be most advantageous to statistics majors. Discuss the possibility of having some statistics majors join these courses. St. Olaf College, for example, has combined mathematics, statistics, and computer science into one department.
- Add one new non-course opportunity to develop the desired skills and experiences in the first year. Add additional opportunities as feasible.

- Reevaluate after the first year, and consider what additional changes are feasible in the second year. It will likely take several years to develop the curriculum and associated opportunities needed to fully implement the guidelines.

This approach is illustrated in the following flowchart.



Iterative Approach to Incorporate 2014 Guidelines in Smaller Programs