Calvin Cochran, Quantitative Reasoning
05/19/2020

Project Title: Developing a simulation and programming focused introductory statistics course

Project Description

The Mathematics department and QR program are excited to implement two new intro level statistics courses in the 2019-2020 school year: Stat 160 and QR/Stat 150. Previously, Stat 101 served as a catch-all introduction to statistics both for students who planned to take further courses in inferential statistics and data science and for students who did not. These two groups of students often wanted different things from the course, making it difficult to teach. Starting this school year, Stat 101 will be retired and "split" into Stat 160 and QR/Stat 150. The former will serve as an onramp for students who plan to take further statistics beyond the intro level. The latter will serve the students who do not. This proposal is to support the development of Stat 160 for the upcoming school year. The first section of this course will be offered in the spring of 2020 to between 25 and 30 students. It will be offered at least once per school year (perhaps more frequently if there is demand) and will cover essential descriptive statistics, probability, hypothesis testing, confidence intervals, linear regression, and other topics.

Besides the new student audience, Stat 160 sports two additional deviations from its predecessor, Stat 101. First, Stat 160 students will learn to create data visualizations, run hypothesis tests, and generate confidence intervals using the programming language R. The ability to program in R has become an indispensable and sought after skill for statisticians and data scientists to master, and it is important that students who may major in a related field be exposed to R as early as possible in their education.

Second, Stat 160 will utilize simulation based approaches to illustrate critical statistical concepts. Too many students walk away from their intro statistics course as computation machines: they are armed with the ability to execute the hypothesis test algorithm but are unable to explain their rationale or the real significance of their findings. This outcome is often the result of an overemphasis on calculation and mechanics (i.e. the "how" of statistics) without sufficient reference to the "why". With the advent of ubiquitous and affordable technology, instructors are no longer bound to the traditional tools of statistics pedagogy that so often rely on abstract distributions and computation. Instead of just giving students the ability to calculate a p-value (coupled

with the “rule” of rejecting the null when the p-value is less than 0.05), online applications like “StatKey” allow students to create randomization distributions which illustrate how (a)typical their sample results would be if the null were true.

Research evidence supporting this teaching innovation

Recent literature has been encouraging of these pedagogical innovations in introductory statistics courses. R has been found to be an effective and practical tool for demonstrating introductory statistics concepts[1], and its use in the classroom complements the simulation-based method well[2].

In addition, a growing number of papers[3,4] encourage and give guidance on the use of simulation based approaches to teaching introductory statistics, and many empirical papers lay out the benefits of this method[5,6] to students’ initial comprehension and retention of deep statistical concepts. The simulation method is both effective and memorable!

Assessment

Students’ comprehension and retention of material (excluding R programming) will be measured using well-established ARTIST assessment tools. The ARTIST website contains a repository of challenging, concept-focused statistics problems which can be integrated into daily formative assessments (at the end of each class) or into summative assessments in the form of exams. It also sports ideas for writing assignments and projects, though we discourage the latter in this class to maximize time for learning new topics. Students’ ability to program in R will be assessed using homework problems.

Itemized budget

This grant would help support the following:

- Student graders for Stat 160. ($900)
- Travel, hotel, and registration funds to present pedagogical results at the 2021 National Numeracy Network conference based on my experience teaching Stat 160 in the spring. ($650)
- A small faculty stipend for the development of new course materials for Stat 160 ($2,000)

Total: $3,550

Citations

1. Salim, N.R., K. Gopal, A.F.M. Ayub (2018) “Experiential Statistics Learning with RStudio: Study on Students’ Engagement.” *Journal of Physics: Conference Series* 1132, 841-852.
2. Zhang, X. and Z. Maas (2019) “Using R as a Simulation Tool in Teaching Introductory Statistics.” *International Electronic Journal of Mathematics Education*, 14(3), 599-610.
3. Cobb, G. W. (2007). “The Introductory Statistics Course: A Ptolemaic Curriculum?” *Technology Innovations in Statistics Education* 1(1). Retrieved from https://escholarship.org/uc/item/6hb3k0nz
4. Rossman, Allan J., Beth L. Chance (2014) “Using simulation-based inference for learning introductory statistics.” *WIREs Computational Statistics* 6, 211-221.
5. Chance, Beth, Jimmy Wong, and Nathan Tintle (2016) “Student Performance in Curricula Centered on Simulation-Based Inference: A Preliminary Report.” *Journal of Statistics Education*, 24(3), 114-126.
6. Hildreth, Laura A., Jim Robison-Cox, and Jade Schmidt (2018) “Comparing Student Success and Understanding in Introductory Statistics Under Consensus and Simulation-Based Curricula.” *Statistics Education Research Journal* 17(1), 103-120.