To teach scientific reproducibility, start young

February 28 2014

The ability to duplicate an experiment and its results is a central tenet of the scientific method, but recent research has shown an alarming number of peer-reviewed papers are irreproducible.

A team of math and statistics professors has proposed a way to address one root of that problem by teaching reproducibility to aspiring scientists, using software that makes the concept feel logical rather than cumbersome.

Researchers from Smith College, Duke University and Amherst College looked at how introductory statistics students responded to a curriculum modified to stress reproducibility. Their work is detailed in a paper published Feb. 25 in the journal *Technological Innovations in Statistics Education*.

In 2013, on the heels of several retraction scandals and studies showing reproducibility rates as low as 10 percent for peer-reviewed articles, the prominent scientific journal *Nature* dedicated a special issue to the concerns over irreproducibility.

Nature's editors announced measures to address the problem in its own pages, and encouraged the science community and funders to direct their attention to better training of young scientists.

"Too few biologists receive adequate training in statistics and other quantitative aspects of their subject," the editors wrote. "Mentoring of

[young scientists](#) on matters of rigour and transparency is inconsistent at best."

The authors of the present study thus looked to their own classrooms for ways to incorporate the idea of reproducibility.

"Reproducing a scientific study usually has two components: reproducing the experiment, and reproducing the analysis," said Ben Baumer, visiting assistant professor of math and statistics at Smith College. "As statistics instructors, we wanted to emphasize the latter to our students."

The grade school maxim to "show your work" doesn't hold in the average introductory statistics class, said Mine Cetinkaya-Rundel, assistant professor of the practice in the Duke statistics department. In a typical workflow, a college-level statistics student will perform data analysis in one software package, but transfer the results into something better suited to presentation, like Microsoft Word or Microsoft PowerPoint.

Though standard, this workflow divorces the raw data and analysis from the final results, making it difficult for students to retrace their steps. The process can give rise to errors, and in many cases, the authors write, "the copy-and-paste paradigm enables, and even encourages, selective reporting."

"Usually, a data analysis report, even a published paper, isn't going to include the code," Cetinkaya-Rundel said. "But at the intro level, where this is the first time students are exposed to this workflow, it helps to keep intact both the final results and the code used to generate them."

Enter R Markdown, a statistical package that integrates seamlessly with the programming language R. The team chose R Markdown for its ease of use—students wouldn't have to learn a new computer syntax—and

because it combines the raw data, computing and written analysis into one HTML document. The researchers hoped a single HTML file would give students a start-to-finish understanding of assignments, as well as make studying and grading easier.

The study introduced R Markdown to 417 introductory statistics students (272 from Duke University, 145 from Smith College) during the 2012-2013 school year. Instructors emphasized the lesson of reproducibility throughout each course and surveyed 70 students about their experience using R Markdown for homework assignments.

The survey, conducted once at the beginning of the semester and once at the end, showed gradual gains in student preference for R Markdown. The percentage of respondents who indicated they found R Markdown to be frustrating at first but eventually got the hang of it jumped from 51 to 75 percent. The students vastly preferred it to the alternative, with 70 percent strongly disagreeing that they'd rather use the copy-and-paste method.

The research team also found that even when students had no prior computing experience or expressed negative attitudes toward R Markdown, their grades did not appear to suffer. Future surveys will ask more pointed questions about how much of the lesson on reproducibility students absorb from the modified curricula.

As the use and analysis of big data becomes increasingly sophisticated, the team writes, the ability of researchers to retrace steps and achieve the same statistical outcomes will only grow in significance.

  More information: *Technological Innovations in Statistics Education*, Feb. 25, 2014. escholarship.org/uc/item/90b2f5xh