

Vital statistics

As the data deluge swells, statisticians are evolving from contributors to collaborators. **Sallie Ann Keller** urges funders, universities and associations to encourage this shift.

This summer, statistics PhD graduate Geneva Allen took on a new position — a joint assistant professorship between the statistics department at Rice University and the department of paediatrics and neurology at the neighbouring Baylor College of Medicine in Houston, Texas. She took the post for the access it will give her to important problems and great volumes of data, despite the risks for an untenured professor of splitting attention between multiple fields and institutions.

Allen's appointment is emblematic of growing links between statisticians and research scientists. Statisticians are increasingly not just contributors, asked to help out with data analysis in the end game, but collaborators, needed to help in defining research questions and designing the study from the beginning. This is a positive trend — for statisticians and scientists alike — that needs to be encouraged.

Although no thorough analyses have been done to confirm the rise of statisticians in collaborative projects, anecdotally the trend is clear. At North Carolina State University's statistics department in Raleigh — one of the top statistics groups in the United States — nearly half of all faculty publications are now published in non-statistical scientific and engineering journals. In line with this trend, the need for statisticians is rising. In the United States, master's degrees in statistics have rocketed from about 900 in 2000 to more than 1,800 in 2009 (see 'Increasing significance?'). And still, jobs for statisticians go unfilled across industry, academia and government.

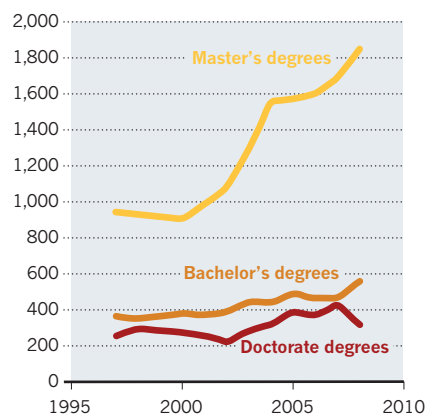
MODEL BEHAVIOUR

Statistics has had a rich history of collaboration; indeed it is sometimes called the quintessential interdisciplinary science. I was reminded of this in a recent conversation with oceanographer Walter Munk, whose collaboration with statistician John Tukey in the 1950s and '60s, for example, helped to forge the field of wave forecasting. The earliest statistical pioneers, including Thomas Bayes of Bayesian analysis fame, worked in close collaboration with scientists.

Many factors are helping to increase such links. In health science, researchers have become wary of pharmaceutical company involvement in study design and statistical analysis. In response, the *Journal of the American Medical Association (JAMA)* has since 2005 insisted that industry-affiliated

INCREASING SIGNIFICANCE?

The number of completed US statistics degrees has grown with demand for the specialists.



papers undergo independent statistical review. Fields from neuroscience to genetics and particle physics are now deluged with data. Statisticians are increasingly needed to help sort out which data are important, and what can be concluded from them. Such disciplines are calling upon statisticians such as Allen to help to find hidden gems.

Such positive moves should help to counteract the fact that many scientists still do not properly understand or appreciate the role of statisticians.

When I was head of the Los Alamos National Laboratory Statistical Sciences Group from 1998 to 2005, for example, I often encountered physicists and engineers who were dismissive of what statisticians could do to help them. They thought we planned to 'ride in on white horses' and somehow magic away their lack of data regarding the stability of the nuclear stockpile. Of course, this was not our job. Instead, we made their models more transparent and helped to quantify the uncertainty in their results, eventually winning over the sceptics.

Even the field of climate change, which is fundamentally concerned with quantifying uncertainty, hasn't fully embraced statisticians. A panel investigating the behaviour of scientists at the University of East Anglia's Climatic Research Unit in Norwich, UK, recently concluded that "inappropriate statistical tools with the potential for producing misleading

results have been used by some other [climate research] groups, presumably by accident rather than design." The panel noted: "It is regrettable that so few professional statisticians have been involved in this work, because it is fundamentally statistical."

Three things can and should be done to encourage collaborative work between statisticians and scientists. First, universities should reward junior faculty members for exploring interdisciplinary links, not penalize them for dividing their attention.

Second, new funding models are needed to support large statistics groups in academia. In the 1990s, a faculty member might have had one or two statistics students or postdocs in their group. Now, thanks to increasing numbers of graduates and modern computing power that allows for more sophisticated work, research groups of ten or more are not uncommon. Agencies including the US National Science Foundation (NSF) need to recognize this, and increase the size of awards to fund these large groups. Perhaps it will help that the NSF's Division of Mathematical Sciences, in which statistics sits, is now being led, for the first time, by a statistician — Sastry Pantula, former department chair at North Carolina State University, and current American Statistical Association president.

Finally, the statistics community needs to actively seek collaboration. There are many scientists who still labour under the misapprehension that statisticians aren't needed or can be parachuted in at the end of a project. Because there is no shortage of jobs or collaborative opportunities, many statisticians have become content to sit back and wait for a call. Others fear that by pushing the field too much towards collaborative work, they will lose their distinct identity. This fear is misplaced. Those in positions of influence — from department chairs and deans to the president of the American Statistical Association — need to help to shift the culture and encourage outreach.

Wednesday 20 October is the first World Statistics Day, sponsored by the United Nations. In its honour, we should celebrate the links between scientists and statisticians. Everyone stands to benefit. ■

Sallie Ann Keller was president of the American Statistical Association in 2006, and is director of the Science and Technology Policy Institute, Washington DC 20006, USA. e-mail: skeller@ida.org