Faculty Q&A with Stephen Stigler

Stephen Stigler is the Ernest DeWitt Burton distinguished service professor of statistics and a member of the Committee on Conceptual and Historical Studies of Science. He earned his bachelor’s degree in mathematics from Carleton College in 1963 and his PhD in statistics from the University of California, Berkeley, in 1967. He has served as president of both the Institute of Mathematical Statistics and the International Statistical Institute.

How did you become interested in the history of statistics?
Well, basically, I got tenure. I had all along been interested in the history, but in my graduate study, I’d never heard of any work that had been done earlier than that of the major professors in my department.

But when I discovered accidentally that there was quite a bit of interesting work at an earlier period, I looked into it and I got hooked. For the next 20 years or so, I was dabbling in history as well as working on more standard research topics, but in recent years, it’s been much more concentrated in history than on other things.

What are you working on right now?
I’m partway through writing a book tentatively called “The Five Most Consequential Ideas in the History of Statistics”; I hope to have a first draft by the end of the summer. That’s one answer.

Another answer is that I’ve been working on a project that explores how modern statistics, and in particular modern multivariate analysis, evolved from questions that were inspired by Darwin’s On the Origin of Species. It attempts to deal with problems in the theory of evolution that Darwin didn’t fully realize, but whose solutions turned out to be to be immensely fruitful for all sciences up to the present day.

It was absolutely amazing how simple questions that could be asked based on On the Origin of Species led to an overthrow of a mindset that had existed in science since Euclid. I think it’s one of the great and inspiring stories in the history of science, and it hasn’t been appreciated.

What kind of statistical problems are we talking about?
Well, there is a doctrine in Euclid that used to be a standard tool in elementary math classes called the rule of three. Basically, it says that if A is proportional to B, and C is proportional to D, and these proportions are equal, then all you need to know are three of the numbers and you can calculate the fourth. It’s a simple exercise in linear extrapolation. The need for it comes up all the time in commercial arithmetic and in scientific questions.

What was discovered in the late 1800s was that this was exactly the wrong thing to be doing whenever there’s variability in your data. There are much more effective methods—we call them now by names like regression analysis—that will avoid the biases that come from a straightforward application of Euclid.

Much of your work has to do with not just the history of statistics, but also the history of the way people understand and use statistics.
History can be a dry recitation of facts or can be an exploration of intellectual history, by which I mean the way ideas have developed and changed. Ideas change as they migrate between different fields of application. Methods that were based on a framework of Newtonian gravitational theory in astronomy, like the method of least squares and the use of probability to deal with error, became the means of creating the objects of modern social science.

What concept in statistics, or perhaps in the scientific method in general, do you think is most misunderstood?
It is often accepted in the naive history of science that one conceives of a theory and then one takes measurements. The process of taking measurements is not that simple. It can involve deep statistical ideas; it can involve serious conceptual problems; and if handled poorly, it can involve major errors. Even modern scientists are often unaware of some of these nuances and problems.

One of the things I look for when I’m studying old materials is not just what people did but how they did it and whether they did it well. In particular, I keep my eyes open for blunders. And I find them, in old works and in modern works. It’s not unusual to pick up a current issue of Nature or Science and be able to locate some significant blunder that has been propagated by a noted and excellent scientist who was ignorant of the sovereign science of statistics.

I will say that we don’t make quite as many elementary errors as we used to, but that’s not to say that you can’t find some really elementary errors in modern scientific literature as well. I’m not studying it just because of perversity, but because it sheds some light on how understanding spreads and develops and how precarious that understanding can be.