Frederick Mosteller and Applied Statistics

As a middleman between statisticians and other scientists, Mosteller is acutely aware of current research problems

Frederick Mosteller of the Harvard School of Public Health is a quiet, soft-spoken man whose life, for more than 40 years, has revolved around statistics. Primarily an applied statistician, he is most famous for bringing statistics to other fields. Persi Diaconis, a former student of Mosteller's who is now at Stanford University, explains that Mosteller "thinks of simple, widely applicable analyses and is able to communicate with people in other sciences." As a middleman between statisticians and other scientists, Mosteller is acutely aware of current research problems in statistics. He spoke to *Science* about his interest in statistics and about the current turmoil among statisticians over how best to analyze data.

This is the second of a series of occasional articles about mathematics as seen through the eyes of its most prominent scholars.

Mosteller's love of statistics began when he was an undergraduate at the Carnegie Institute of Technology. He was asked in a course on physical measurements to decide the probability that the numbers on the faces of three dice add up to ten. "After working on the problem, I found a way to solve it, but I realized that my method would not work if I were given six dice instead of three," he relates. "I asked my professor if he knew a way to handle the problem and my professor referred me to [the statistician] E. G. Olds."

Olds showed Mosteller an ingenious way to solve the problem—a way that demonstrated the usefulness of what are known as generating functions. This was tremendously exciting to Mosteller who, like many students, had learned about generating functions as part of his course work but had never realized the clever uses to which they can be put. Olds then took Mosteller under his wing, giving him material to read and problems to solve. Finally, he steered Mosteller to Princeton for graduate work in statistics and probability.

In the years since, Mosteller has been fantastically productive. "No one actually knows all the things Fred is involved in," Diaconis says. "When friends of his, like [William] Kruskal, [Paul] Meier, and myself get together at meetings and talk about work we're doing with Fred, none of us knows about the projects he's involved in with the others." Diaconis says that despite the large numbers of projects he has worked on there is nothing slipshod about Mosteller's work. "He wants the work done right and he wants it to be clear," Diaconis explains.

One way Mosteller manages to be so prolific seems to be by carefully regimenting his time. He used to have a sign in his office asking, "What have I done for statistics in the past hour?" and he told Diaconis that he once kept a record of what he did every 15 minutes of the day.

Mosteller has seen statistical methods become increasingly prevalent in all fields of science, especially the social sciences. He says that an economist coming out of a first-class graduate school today probably knows more statistics than a person receiving a Ph.D. in statistics in 1940. But statistics as a discipline is still widely misunderstood by nonspecialists. A common view is that using statistical methods is like using a cookbook: one simply chooses the appropriate recipe to analyze a particular set of data. In fact, statisticians have for years been arguing over which recipes to choose and how to choose them. There is often no consensus on how best to analyze data or even what is meant by "best." And competing methods of data analysis often lead to quite different results.

One area of continuing controversy in which there has been a recent surge of research is exploratory and confirmatory data analysis. Confirmatory analysis, as its name suggests, is used to decide whether data confirm hypotheses they were designed to test. Exploratory analysis, on the other hand, is used to look for unexpected patterns in data and to evaluate whether those patterns are significant.

Frederick Mosteller Photo by Eric Poggenpohl

0036-8075/79/0427-0397$00.50/0

The general problem in exploratory analysis, Mosteller explains, is to find routine ways to look for the unexpected in data. "All data have funny aspects," he says. "In fact, if they didn't have anything funny about them, that itself would be funny." But there will always be the problem of what credence to give these unexpected findings and how to analyze them. "Part of exploratory data analysis is to decide the best form for variables, which variables to use, and how to combine them," Mosteller says.

Mosteller tells a story indicative of the kinds of difficulties that can arise with exploratory data analysis. Recently, the Food and Drug Administration (FDA) funded an experiment in which red dye No. 40 was fed to mice to see if the substance causes cancer. The data did not support this hypothesis, but they did lead one investigator to suggest that the dye might speed up the development of tumors. The FDA called in Mosteller and other statisticians to decide whether this exploratory analysis was convincing. After much debate, most concluded that it was not and that further experiments with the dye were needed. Mosteller explains, "The incident is one in which exploratory data analysis led to a lot of complications because there is no routine way to extend the confirmatory notion to exploratory analysis." These sorts of difficulties arise constantly in large-scale data analysis, such as those of clinical trial data.

Closely related to these problems in exploratory data analysis is research on resistant statistical measures. These are measures that change very little when a small part of the data changes, even when that small part of the data changes a great deal. The purpose of using resistant measures is to discount what may be meaningless data fluctuations. Their relation to exploratory analysis is that they may enable investigators to temper their explorations, to decide which are the meaningful fluctuations in data.

A simple example of a resistant measure, Mosteller says, is the median. This is essentially the midpoint of the sample —half the data are greater than and half less than it. If a sample consists of about 100 measurements and one of these is moved to infinity, the median hardly moves at all. In contrast, the mean, which is the arithmetic average of a sample, deteriorates badly in such circumstances.

"Resistance is a mixed blessing," Mosteller says. "Although desirable, it is not desirable in the extreme. You need some sensitivity to the data you are looking at." He explains that one way of deciding how good resistant measures are is to use classical methods of analysis that measure the efficiency of certain kinds of estimates.

Recently, investigators such as John Tukey of Princeton University and Peter Huber of Harvard University developed a large collection of resistant measures that are highly efficient in a variety of circumstances. These methods, which were deliberately searched for, "help with data that have a lot of outliers," Mosteller says. These are data that contain a core of truth but the truth can be obscured some of the time by bad data.

Mosteller himself is more inclined to put such new methods to use than to develop them. Although he has made some really solid contributions to theoretical statistics in the areas of nonparametric methods and contingency table analysis, he thinks of himself mostly as an applied statistician. He enjoys working on practical problems and seeing how choices of statistical techniques affect the analysis of the problem. For example, about 20 years ago, he and David Wallace of the University of Chicago decided to use what are known as Bayesian statistics to study the authorship of the disputed Federalist Papers, a series of essays written by Alexander Hamilton and James Madison in 1787–1788 in support of the proposed U.S. Constitution. (They concluded that Madison wrote all of the disputed papers, with the exception of three papers that were started by Hamilton and finished by Madison.)

Bayesian statistics require that prior information be incorporated into the analysis. For example, one type of prior information is the probability that Madison, say, wrote a particular paper based on the observation that Madison used the word "while" whereas Hamilton used "whilst." The competing method of analysis, called frequentist statistics, handles prior information quite differently, such as by including the information in cautionary remarks at the end of the analysis. Statisticians have disputed which method to use for 250 years and their debate continues today.

Mosteller and Wallace made the unexpected discovery that the result of a Bayesian analysis varies greatly depending on what sort of model they use for their data. "This was awesome," Mosteller says. "We went in thinking that everything depended on what sort of prior information you used. Our finding says that statisticians should remove some attention from the choice of prior information and pay more attention to their choice of models for their data."

Mosteller is well known for the number of ways he has applied statistics to other disciplines, including medicine and learning theory. For example, he collaborated with John Gilbert and Bucknam McPeek of the Massachusetts General Hospital in a study of 107 published papers on surgery and anesthesia. Their goal was to see how likely it is that a surgical innovation will be an improvement. They found that almost half the innovations were of some benefit and a fair number of those were of significant benefit in that they reduced complication rates and increased survival rates for patients. In contrast, Mosteller and his associates had found previously that most new social programs do not work. But Mosteller, Gilbert, and McPeek also discovered that medical researchers must be more careful in their randomization procedures when they test new treatments with clinical trials. In addition, they found that more long-term follow-ups are needed to assess how the new therapies affect patients' quality of life.

Although most of Mosteller's work is serious, he also has received some publicity for more whimsical work. Several years ago he used statistics to design a new system for scoring football. In his system, which was reported by the news services, teams receive credit not only for scoring but also for defense.

Mosteller has a number of interests outside statistics. He is fascinated with abstract art and ancient history, for example. His office is decorated with abstract art prints, and he has framed a crayoned copy of a Stella print that he made himself.

Mosteller also was once interested in magic. Diaconis, who himself is a magician, says that Mosteller invented a number of famous magic tricks when he was in graduate school and then did magic seriously for about 5 years, after which he dropped it. Mosteller explained to Diaconis that he thought magic interfered with his credibility as a teacher. If he did a magic trick in the midst of a sampling demonstration with cards, for example, the class would no longer believe the sampling demonstration. (Mosteller takes quite seriously his job as a teacher, and students routinely give him a standing ovation at the end of the last class each semester.)

Although Mosteller has these outside interests, he apparently keeps them strictly compartmentalized. "A lot of people, even friends who have known him for years, say Fred is an enigma to them," Diaconis says. Perhaps this is the result of his highly efficient life-style and his shying away from self-aggrandizement.—GINA BARI KOLATA