touch cells that lack processes. Mutations in a sixth gene cause touch cells to degenerate soon after they are formed. And mutations in the other two genes affect the lineages that generate the touch cells.

Chalfie believes that he and Sulston have detected most, if not all, of the genes that, when mutated, can have major deleterious effects on this set of neurons. Interestingly, no one has found a mutation that specifically affects the direction of growth of touch cell processes without having a similar effect on other neurons. Nor is there yet a mutation that specifically affects a subset of the six touch cells in any way.

Overall, then, the differentiation and positioning of the touch-sensitive neurons is brought about by a combination of the activities of a small group of genes specific to these cells and an unknown number of genes that affect these and other cells.

Using this type of genetic approach, genes and gene products known to control the development and functioning of the neurons involved in touch sensitivity, as well as in other behaviors, can be identified and perhaps isolated. "The *C. elegans* genome is 20 times bigger than that of *Escherichia coli* and half that of *Drosophila*," says Sulston. "There is currently a big push to clone the genes, to map the whole genome, and then we will be able to focus on specific bits." Although some of the genes involved in neurogenesis are likely to specify cell surface components, for instance, which may play a direct and clear role in assembly of the nervous system, others may well prove to have functions that are less obviously related to development. "We might find an enzyme that is involved in some aspect of biosynthesis, for instance, but as a practical matter it would represent a gene that is involved in the logic of development."

This comment illustrates the recognition by Brenner and his associates that development is unlikely to be the result of a discrete, sequential developmental "program," but instead is the outcome of a more holistic logic of molecular assembly (see *Science*, 22 June, p. 1327). In which case, one will need to know everything about the molecular genetics and biochemistry in order to understand how the animal is put together. "And what you will see," says Sulston, "is just one of many possible ways in which *C. elegans* can be made. Development has evolutionary history built into it, and the result is not always the most obvious or logical way of doing things."

Brenner and his colleagues are occasionally asked whether knowing everything about *C. elegans* reveals anything about the rest of the biological world, much of which appears at first sight to be a good deal more complicated. "This is the same question that was asked when phage genetics was started," comments Horvitz. "Although certain detailed aspects of molecular genetics have turned out to differ, many basic principles have proved to be universal. I expect the same thing will happen with *C. elegans*. We should learn a lot both about how it differs and how it is similar to other organisms. From these studies of *C. elegans* fundamental principles concerning both development and behavior may well emerge."

Caenorhabditis elegans may be a small organism, but it is by no means simple.—ROGER LEWIN

The Art of Learning from Experience

Statistician Bradley Efron tells what his field is about and how a new method, the bootstrap, exploits the power of large-scale computing

"Statistics," says Bradley Efron of Stanford University, "is quite underappreciated." Most people who think of statistics at all consider it as simply a tool—a way to tell if data are significant or to estimate confidence intervals. But statistics is a deeply philosophical subject that tries to get at how we learn from experience. It is a dynamic field, full of arguments and beginning to change its very nature as its practitioners exploit the power of large-scale computing.

Efron, 45, is one of the leaders in the new statistics. He has invented an extremely promising new statistical tool, called "the bootstrap" and which, he says, "substitutes computing for thinking."

Although statistics is often thought of as a branch of mathematics, it actually lies on the border between mathematics and philosophy. "Obviously," says Efron, "statistics has mathematical structure—that's the only way anyone has found to say things in statistics." But the subject matter of statistics does not concern itself with typical mathematical reasoning in which results are deduced from axioms. Its logic goes in the opposite direction. Statisticians start with examples of things that are and try to determine what axioms could have given rise to them. "To step backward from what you've seen to what might have given rise to it is logically, mathematically, and actually difficult," Efron remarks. "We statisticians think deduction is child's play. In a sense, statistics is the most ambitious intellectual attack."

For Efron, the decision to become a statistician came only gradually and after he realized what he believes are his limitations as a mathematician. He always wanted to be a mathematician, he says, but he had a problem. "I was a terrific 19th-century mathematician. Give me a calculus problem and I could knock it dead. But I was not a very good twentieth century mathematician. I like to compute things. Modern mathematicians don't compute. They organize their ideas to another level of abstraction beyond calculations. I was terrible at things like modern abstract algebra. I have no mind for it at all."

Efron grew up in St. Paul, Minnesota, the son of a truck driver who was also an amateur mathematician. He learned from his father how to do calculations in his head. Set on becoming a mathematician, he majored in math at the California Institute of Technology, where he graduated second in his class. Then he started graduate school at Stanford, still majoring in math. But he was suspended from Stanford when, as editor of the school humor magazine, he published an article poking fun at religion. When he returned to Stanford, he returned to the statistics department.

Efron jokes that his suspension from Stanford will haunt him to his grave. "I often say that if I cure cancer, the Stanford newspaper story will begin, 'Bradley Efron, who once was kicked out of Stanford, today discovered a cure for cancer.' " But his reentry into statistics proved providential. Here was a field after his own heart—a field where computations reign supreme and where he could invent methods that would make statisticians compute even more.

Reflecting the interplay between logic and mathematics, there are what Efron calls "two opposing currents" in statistics. One approach, initiated in the early part of this century by the British statistician Sir Ronald Fisher, is based on the idea that the way to solve statistical problems is to get at their logical basis. The problem is that of uncertain inference. For example, if you say, "John is a man and men live less long than women," does that mean that you can conclude that John will live less long than his wife?

Efron gives another example of uncertain inference. "Suppose you say that penicillin is better than sulfa drugs for treating pneumonia. Well, how do you prove something like that? What do you mean? Do you mean that every patient will do better if you give them penicillin? If so, you will be sadly surprised when you do an experiment." In the real world, most things are not always true 100 percent of the time.

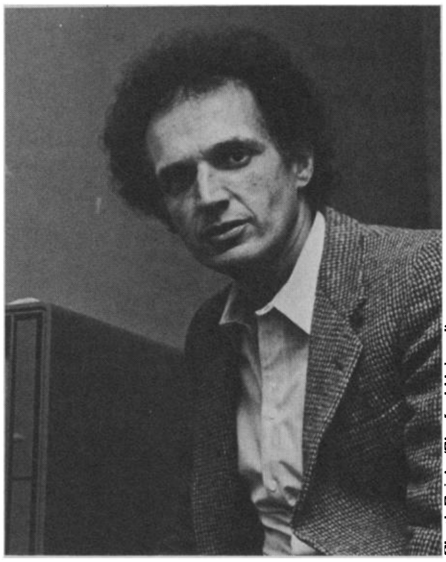
The second approach to statistics, which was initiated by Jerzy Neyman, who was at the University of California at Berkeley for much of his career, is to say that the problem of solving statistics problems is basically a mathematical optimization problem. Efron explains, "Usually every possible method gives some probability of making an incorrect statement when you interpret your data. Neyman's theory says that if you can find a method of interpretation that has the smallest possible chance of making an incorrect statement, you will have done the correct thing." The idea of optimization sounds appealing but, according to Efron, it has been carried out successfully only for very simple problems.

Both of these schools of thought contribute to statistics, and, in one of the major achievements of the field, both helped demonstrate the value of taking an average. This was work that went on for nearly 50 years, concluding around 1950. The question was, What should you do if you have accumulated data that are distributed as a bell shaped curve? "An obvious thing to do is to take an average," Efron remarks. "The triumph of the logical school was its demonstration that once you have the average you can throw away the rest of the data. You might as well use the average as use all of the data. Then the optimality people finished off the problem by showing that the average really does give you the best answer."

"When I was a student, the optimality school reigned supreme [at Stanford]," Efron says. "But the biostatisticians mostly used the results of the inference group. You'd be amazed at what contention it [the debate between the two philosophies] has caused over the years."

Statistics, Efron says, is "a slow entry field." It takes time to develop a feel for it. "There has never been a great 19-year-old statistics genius," Efron remarks. "It took me a good part of 15 years to get straight in my mind what I should work on."

Efron's most recent work has been on a general question that often plagues



Chuck Painter/Stanford University

Bradley Efron

Working on the border between philosophy and mathematics.

scientists: What would be seen if there were a lot more data? He tells, as an example, of an experiment he is analyzing for a researcher in Stanford's medical school who examined 120 human-mouse hybrid cells, 40 of which made a protein of interest. She wants to know how variable her results are. Might they have occurred by chance? If she had had 10 times as many hybrid cells, would there still be about one-third of them making the protein? In this case, Efron can use well-known methods to get the standard error and answer her question. But, he says, "This is an extremely simple situation. It is common to have 1000 pieces of information and 100 unknown parameters. Then you ask how accurate is your determination of one of those parameters. You never get to see the 'real' answer. The whole point of the analysis is to know what the range of possibilities would be if you took a lot more data."

Fisher, founder of the inference school of statistics, was a leader in developing a method, called the "maximum likelihood method," that provides estimates of averages and standard deviations for general problems. But, says Efron, "The theory becomes undependable in complicated situations, particularly if there are lots of unknown parameters. The standard error it gives can be quite a bad approximation and the worst thing is that you can't always calculate it. It involves putting together a probability model assuming such things as a bell shaped curve. What I've been trying to do is to develop automated methods that don't involve such assumptions. I noticed that a number of calculations could be automated and you were barely using the model."

Efron calls his method the bootstrap because "You use the data to estimate probabilities and then you pick yourself up by your bootstraps and see how variable the data are in that framework." As an example of how the method works, Efron tells of a study he and his Stanford colleague Persi Diaconis did of correlations between grade point averages (GPA) and scores on the Law School Admission Test (LSAT). The example is a simple one and the results that Diaconis and Efron obtain with the bootstrap could easily have been obtained with classical methods. But it demonstrates the bootstrap procedures in a straightforward way.

For 15 schools in 1973, the correlation between the average LSAT score and the average GPA at each school was 0.776, which means that the LSAT scores and GPA's were highly correlated. But does that mean that LSAT's and GPA's are highly correlated for *all* law schools? The bootstrap method gives a way to find out.

The first step is to copy the data for each school in the sample an enormous number of times—a billion times, for example. This creates a universe of 15 billion data points, 1 billion for each school. The computer selects from this universe random samples of 15 data points, called bootstrap samples, and calculates the correlation coefficient for each such sample. (In practice, Efron points out, the bootstrap samples are obtained without actually creating the 15 billion point universe by using random number generators to select among the original 15 points.)

When Diaconis and Efron looked at 1000 of these computer-generated correlation coefficients, they found that 68 percent of them were between 0.654 and 0.908. The conclusion of this bootstrap analysis is that the observed value of the correlation coefficient from a random sample of 15 bootstrap samples varies from the true value by 0.127.

Since they had data from all of the law schools in this country in 1973, Efron and Diaconis were able to test the predictions of their bootstrap model. The true correlation coefficient was 0.761, with a true variability of 0.135 in samples of size 15. Since the bootstrap variability estimate was 0.127, the bootstrap was quite accurate in this case.

Efron's method seems like magic, like a sleight of hand trick. And many statisticians instinctively distrusted it. "When I presented it to people they said it wouldn't work. Some said it was too simple. Others said it was too complicated," Efron recalls. David Freedman of the University of California at Berkeley, who also has lectured on the bootstrap, says he got the same sort of reactions. "In my field, a lot of people come from Missouri and want to be shown. Some you would even call curmudgeons," Freedman says. "I never generated quite so much opposition as when I was talking about the bootstrap. People were afraid it was all done with mirrors." Frederick Mosteller of Harvard University, who says he thinks the bootstrap "is a very good idea," nonetheless sympathizes with those who tend to doubt it works. "The bootstrap is a little hard to believe," he says. "It seems incestuous. You are trying to learn about the sample error by sampling the sample." Statisticians, Mosteller remarks, "are not ordinarily involved with something as anti-intuitive as this."

But, says Freedman, the whole point about the bootstrap is that it is not done with mirrors. "There are no free lunches in statistics. To draw conclusions about data, you have to make assumptions about processes that generated them. You're really using these assumptions as well as the data when you use the bootstrap."

The bootstrap is gradually coming into use, thanks partly to the efforts of Freedman and Peter Bickel of the University of California at Berkeley who helped establish its theoretical underpinnings. "I think it's a powerful tool," says Freedman. "I see it becoming one of the standard techniques in statistics. It will have a big influence on the field."

But the bootstrap is not perfect—no statistical method is. Freedman, Bickel, Jeffry Wu of the University of Wisconsin in Madison, and others have shown that there are situations in which the bootstrap does not work. "It can give drastically wrong answers and it is hard to say in advance when it will work, although we're beginning to get some pretty good guidelines," says Freedman. A number of statisticians, including Efron, are now trying to pinpoint when the method will work and when it will not.

Richard Olshen of the University of California at San Diego gives two examples of medical statistics problems in which he used the bootstrap. In one case, the method was perfect. In the other case, it had to be slightly modified.

The first example is in a study of the evolution of gait in children. Olshen, a statistician, together with orthopedic surgeon David Sutherland, engineer Edmund Bidden, and physical therapist Marilyn Wyatt, all at Children's Hospital and Health Center in San Diego, want to develop curves to establish angles of rotation of the legs, hips, and ankles in normal children from age 1 to age 7, by which time gait is established. These will be curves much like the height and weight curves that pediatricians use to determine whether children's growth is within the normal range. But the gait curves are much more difficult to develop.

"We statisticians think deduction is child's play. In a sense, statistics is the most ambitious intellectual attack."

The gait curves and the percentiles showing the distributions of normal gaits will be particularly valuable, Olshen says, in assessing in a noninvasive way the physical development of children with such conditions as mild cerebral palsy or muscular dystrophy. To develop the percentiles, Olshen had data on more than 400 normal children. He recalls, "Getting the averages of the curves and the shapes of the standard deviation curves was no big deal. But getting the percentiles theoretically is a problem that I don't know how to do. It took me a long time to see the obvious—that I should use the bootstrap." Efron's method, Olshen remarks, "was tailor made for this problem."

The bootstrap's limitations were apparent in a problem Olshen worked on with Lee Goldman and Harvey Fineberg of Harvard Medical School and their associates involving diagnosis of heart attack. When patients come into an emergency room complaining of chest pain, how can you quickly determine whether they had a heart attack? The investigators had data on emergency room patients from the Yale-New Haven Hospital and they wanted to develop a set of diagnostic criteria that they then would test on patients at Peter Bent Brigham Hospital in Boston. The goal was to make a "decision tree" that would tell the physicians at each step what decisions to make in diagnosing heart attacks.

The problem, says Olshen, is that "in diagnostic techniques you can have huge biases in the estimation of error rates. The bootstrap is very good for small biases but not as good for large biases." This is not particularly surprising, Olshen notes, because the bootstrap has little variability. There is, he says, "an uncertainty principle in statistics. Many techniques with little bias have much variability and vice versa."

In the heart attack study, Olshen was able to get around the limitation of the bootstrap by doing what is called "bias adjusting." He used the subjective knowledge of the physicians to "prune" the decision tree beyond what the bootstrap suggested. The resulting tree was so good that the investigators using it did better than physicians who used their medical judgment alone to determine whether patients with chest pain had heart attacks.

The problem of determining the limitations of the bootstrap and how and when to augment it with other methods is inordinately difficult. "One of the troubles with statistics as a field is that it is very difficult to prove that things do or do not work," Efron says. "There isn't exactly a 'real world' of data sets. You want methods to work for drug companies, econometrics, all data sets. Trying to say what that means is the subject of quite bitter discussions."

In the meantime, Efron is working on a new aspect of the bootstrap. He believes that the method can give better confidence intervals than more traditional techniques and, to show that, he has been laboriously working away on the desk-size pad of paper he uses for his theoretical statistics work. "I hate working on this," he remarks. "It's difficult and it's very slow going." Efron says the pad of paper is, to him, "torture." It is much more exciting to calculate.

But, in the end, it is often the results of such torturous theoretical work that statisticians find convincing. And Efron would very much like to win the entire statistics community over to his view of large-scale computing as the wave of the future. Still, he says, "I've taken a tremendous amount of guff. Statisticians are hard to convince. They tend to be very conservative in practice. And they should be. This stuff is serious. People use it."—**Gina Kolata**