

Pick a Sample

By IVARS PETERSON

Few college students escape taking an introductory course in statistics. But for many of these students, the lessons don't seem to stick. They remember the pain but not the substance.

Such initial courses tend to turn students off, says Barbara A. Bailar, executive director of the American Statistical Association in Alexandria, Va. To remedy the situation, she says, "a lot of people are now trying to restructure that first course in statistics."

"We're trying to reform what we do so that we're not always lecturing and the students just sitting there trying to transcribe what we say into their notebooks," says David S. Moore of Purdue University in West Lafayette, Ind., author of a number of widely used statistics textbooks. "Statistics teaching should get students actively involved in working with data to discover principles for themselves."

The computer revolution has already wrought some salutary changes. New computer-based techniques, collectively known as "resampling," have provided statisticians with innovative and powerful methods of handling data. At the same time, computers have opened up novel approaches to teaching the basic concepts of statistics. These techniques promise to eliminate many of the shoals encountered by novices venturing into statistical waters.

One pioneer in computer-based approaches to statistical education is not a statistician but an economist. In 1967, Julian L. Simon — now a professor of business administration at

the University of Maryland in College Park — became convinced there had to be a better way to teach basic statistics. This realization was triggered by his discovery that all four graduate students taking his course on research methods in business had used "wildly wrong" statistical tests to analyze data in their class projects.

"The students had swallowed but not digested a bundle of statistical ideas — which now misled them — taught by professors who valued fancy mathematics even if useless or wrong," Simon says.

The experience led Simon to develop an alternative approach to teaching statistics. His method emphasizes intuitive reasoning and involves using a computer to perform experiments on the available data — a process known as resampling — to get meaningful answers to statistical problems. In his scheme, students avoid having to tangle with mysterious formulas, cryptic tables and other forms of mathematical magic to get their results.

"The theory that's usually taught to elementary [statistics] students is a wa-

expensive and had to be avoided as much as possible.

A resampling experiment uses a mechanism such as a coin, a die, a deck of cards or a computer's stock of random numbers to generate sets of data that represent samples taken from some population. The availability of such "artificial" samples permits students — as well as researchers — to experiment systematically with data in order to estimate probabilities and make statistical inferences.

Consider the following question: What are the chances that basketball player Magic Johnson, who averages 47 percent success in shooting, will miss eight of his next 10 shots?

You could try to answer it by looking up a formula of some kind and plugging in numbers to calculate a probability. In Simon's approach, the answer emerges instead from a numerical experiment.

The computer generates 10 random numbers between 1 and 100. It counts any numbers that fall between 1 and 53 as missed shots, then generates a second set of 10 random numbers, and so on, each

Learning statistics — minus the frustrations

tered-down version of a very complicated theory that was developed in order to avoid a great deal of numerical calculation," says statistician Bradley Efron of Stanford University. "It's really quite a hard theory, and should be taught second, not first."

Simon's resampling approach, refined over many years and tested on a variety of students, is now starting to catch the attention of the statistics and education communities.

"With six or nine hours of instruction, students are generally able to handle problems usually dealt with only in advanced university courses," Simon says. Moreover, classroom surveys show that when his resampling techniques replace conventional methods, students enjoy the subject much more.

In one sense, the resampling approach brings statistics back to its roots.

More than 300 years ago, before the invention of probability theory, gamblers had to rely on experimentation — by dealing hands or throwing dice — to estimate odds. The computer has returned this kind of empirical approach to the center of statistics, shifting emphasis away from the mathematical statistical methods developed between 1800 and 1930, when computation was slow and

time counting up the number of misses. After repeating this process, say, 100 times, the computer determines the number of trials in which the player misses eight or more shots. This figure establishes the likelihood of the player suffering a streak of misses, 6.5 percent in this example.

The answer obtained through this approach is as valid as that obtained through conventional methods, Simon says. Moreover, students learn important lessons about statistical inference from such an exercise.

Stripped of the mathematical trappings associated with the standard approach to statistics, the issue confronting a student or researcher shifts from one of simply calculating an answer to one of statistical inference. Instead of asking which formulas they should use, students begin tackling such questions as what makes certain results statistically "significant."

"It becomes mostly a matter of clear thinking," Simon says.

Consider the results that show Magic Johnson will miss eight or more shots out of 10 tries about 6.5 percent of the time. Because of random variation, such streaks of misses occur naturally, even when a player is not in a slump. Hence, when a good player misses eight of 10 shots, it doesn't necessarily mean that he

or she must be replaced or stopped from shooting. The streak could quite easily occur by chance.

The resampling approach addresses a key problem in statistics: how to infer the "truth" from a sample of data that may be incomplete or drawn from an ill-defined population. For example, suppose 10 measurements of some quantity yield 10 slightly different values. What is the best estimate of the true value?

Such problems can be solved not only by the use of randomly generated numbers but also by the repeated use of a given set of data. Consider, for instance, the dilemma faced by a company investigating the effectiveness of an additive that seems to increase the lifetime of a battery. One set of experiments shows that 10 batteries with the additive lasted an average of 29.5 weeks, and that 10 batteries without the additive lasted an average of 28.2 weeks (see table). Should the company take seriously the difference between the averages and invest substantial funds in the development of additive-laced batteries?

To test the hypothesis that the observed difference may be simply a matter of chance and that the additive really makes no difference, the first step involves combining the two samples, each containing 10 results, into one group. Next, the combined data are thoroughly mixed, and two new samples of 10 each are drawn from the shuffled data. The computer calculates the new averages and determines their difference, then repeats the process 100 or more times.

The company's researchers can examine the results of this computer simulation to get an idea of how often a difference of 1.3 weeks comes up when variations are due entirely to chance. That information enables them to evaluate the additive's "true" efficacy.

"The resampling approach . . . helps clarify the problem," Simon and colleague Peter C. Bruce write in the winter 1991 CHANCE, released in late June. "Because there are no formulae to fall back upon, you are forced to think hard about how best to proceed."

In particular, "you get a much clearer idea of what variability means," Efron says. By seeing the possible variation from sample to sample, you learn to recognize how variable the data can be.

In 1977, Efron independently invented a resampling technique for handling a variety of complex statistical problems. The development and application of that technique, called the "bootstrap," has given Simon's approach a strong theoretical underpinning.

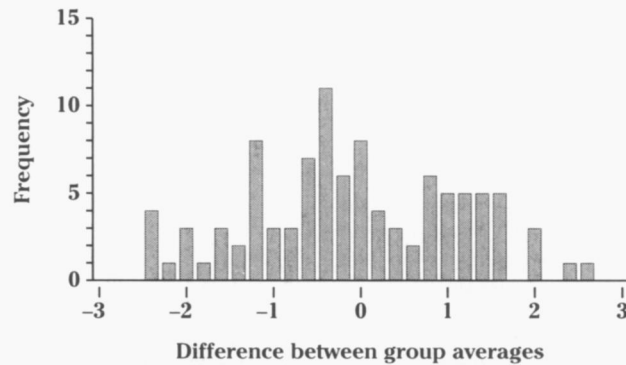
In essence, the bootstrap method substitutes a huge number of simple calcula-

BATTERY LIFETIMES (weeks)										AVERAGE	
WITH ADDITIVE:	33	32	31	28	31	29	29	24	30	31	29.5
WITHOUT ADDITIVE:	28	28	32	31	24	23	31	27	27	31	28.2

A set of tests reveals that batteries with a certain additive last an average of about 1.3 weeks longer than batteries without the additive (above). How strong is the evidence in favor of the additive? Might the difference between the averages have occurred by chance?

One can find out if a difference of 1.3 is unusual by mixing together all the observations, then repeatedly drawing two new groups of 10 numbers each. If this procedure routinely produces differences between the group averages greater than 1.3, then one can't rule out chance as an explanation.

The histogram (below) displays the differences resulting from 100 trials. In this sample, a difference of 1.3 or greater occurs about 27 percent of the time, which suggests that chance can't be ruled out.



Simon & Bruce

tions, performed by a computer, for the complicated mathematical formulas that play key roles in conventional statistical theory. It provides a way of using the data contained in a single sample to estimate the accuracy of the average or some other statistical measure.

The idea is to extract as much information as possible from the data on hand. Suppose a researcher has only a single data set, consisting of 15 numbers. In the bootstrap procedure, the computer copies each of these values, say, a billion times, mixes them thoroughly, then selects samples consisting of 15 values each. These samples can then be used to estimate, for example, the "true" average and to determine the degree of variation in the data.

The term "bootstrap" – derived from the old saying about pulling yourself up by your own bootstraps – reflects the fact that the one available sample gives rise to all the others, Efron says.

Such a computationally intensive approach provides freedom from two limiting factors that dominate traditional statistical theory: the assumption that the data conform to a bell-shaped curve, called the normal distribution; and the need to focus on statistical measures whose theoretical properties can be analyzed mathematically.

"Part of my work is just trying to break down the prejudice that the classical way is the only way to go," Efron says. "Very simple theoretical ideas can get you a

long way. You can figure out a pretty good answer without going into the depths of [classical] theory."

Efron and Robert Tibshirani of the University of Toronto describe the application of the bootstrap procedure and other computationally intensive techniques to a variety of problems, both old and new, in the July 26 SCIENCE.

"Efron's work has had a tremendous influence on the route statistics is taking," Bailar says. "It's a very powerful technique – a major contribution to the theory and practice of statistics. We're going to see more and more applications of it."

But the method also has its critics. "Not everyone is enamored with these [resampling] techniques," says Stephen E. Fienberg of York University in Toronto. The skeptics argue that "you're trying to get something for nothing," he says. "You use the same numbers over and over again until you get an answer that you can't get any other way. In order to do that, you have to assume something, and you may live to regret that hidden assumption later on."

Efron and others who study and use the bootstrap method concede that it doesn't work for every statistical problem. Nonetheless, Efron says, it works remarkably well in many different situations, and further theoretical work will help clarify its limits.

"The potential dangers for misuse are small compared to the amount of good it

does," says mathematician Persi Diaconis of Harvard University. "I think the bootstrap is one of the most important developments [in statistics] of the last 10 years."

The bootstrap's success among statisticians has helped validate Simon's use of resampling techniques to teach statistics and to solve real problems in everyday life. "Simon is onto something that's very good," Bailar says. "I think it's the wave of the future."

"Simon's approach is very consistent with a major theme in statistical education," says Purdue's Moore. "Statistics education ought to concentrate on understanding data and variation. There's a movement away from black-box calculation and toward more emphasis on reasoning—more emphasis on actual experience with data and automating the calculations so that the students have more time to think about what they're doing."

But bringing innovative ideas into the statistics classroom remains a slow process. Part of the problem, says Moore, is that many statistics courses are taught not by statisticians but by engineers, mathematicians and others who learned statistics the old way and haven't seen firsthand how the computer has revolutionized statistical analysis.

"Statisticians work with data all the time," Moore says. "The calculations are so messy and you have to do so many of them that the professional practice of statistics has always relied on automated calculation."

Hence, experienced statisticians are probably more inclined to accept teaching methods that rely on large amounts of computation rather than on manipulation of standard formulas. "There's been a major movement to get computing into the introductory classroom, and to have it revolve around real examples where there's some payoff," Fienberg says.

Simon, who calls his approach "statistics without terror," is convinced this is the way to go. About two years ago, he launched "The Resampling Project" to help spread the word. The project, conducted in association with the University of Maryland, handles such functions as scheduling courses in resampling statistics and providing written material and other help to schools interested in trying these methods.

Simon and Peter Bruce, who heads the project office, videotaped a series of lectures this spring to demonstrate the approach. The University of Maryland Instructional Television System now distributes those tapes. Simon also has developed a special computer language

and software package that enables a user to perform experimental trials on a wide variety of data. With this package, students can learn how to develop an abstract model of a real-life situation, to write a computer program simulating that model and to draw statistical inferences from the resulting data. The package works equally well in the high school or college classroom and in the laboratory or workplace, Simon says.

"The resampling method enables people to obtain the benefits of statistics and probability theory without the shortcomings of conventional methods because it is free of mathematical formulas and is easy to understand and use," Simon and Bruce contend in their CHANCE article. "Our interest is in providing a tool that researchers and decision-makers, rather than statisticians, can use with less opportunity of error and with sound understanding of the process."

Simon has gone so far as to issue the following challenge in a number of publications:

"I will stake \$5,000 in a contest against any teacher of conventional statistics, with the winner to be decided by whose students get the larger number of both simple and complex realistic numerical problems correct, when teaching similar groups of students for a limited number of class hours."

No one has yet accepted the wager. □

More mathematical delights from the creator of the popular Mathematics Calendars!

Theoni Pappas views mathematics as a beautiful collage whose elements are reflected in every aspect of our universe, from the curve of a palm frond to the crest of a wave. To recreate this collage in her books, she has packed their pages full of randomly arranged mathematical puzzles, ideas, games, history, graphics and breakthroughs, each brief morsel designed to pique the reader's curiosity and encourage further exploration. Whether your interest is the evolution of mathematics, geometry in nature, proofs of mathematical theorems or just mathematical trivia in general, there is plenty here to entertain and inform.

Science News Books
1719 N Street, NW, Washington, DC 20036

JoyMath
MoreJoyMath
SetJoyMath

Please send me the book(s) marked below. I enclose the price of the book(s) plus \$2.00 for postage and handling per book (maximum \$4.00 charge). Domestic orders only.

___ Joy of Mathematics, \$10.95 ___ Set of both, \$19.95
___ More Joy of Mathematics, \$10.95

Name _____

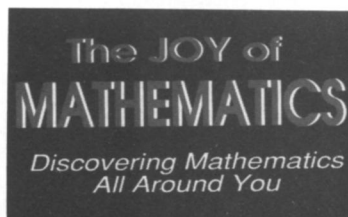
Address _____

City _____ State _____ Zip _____

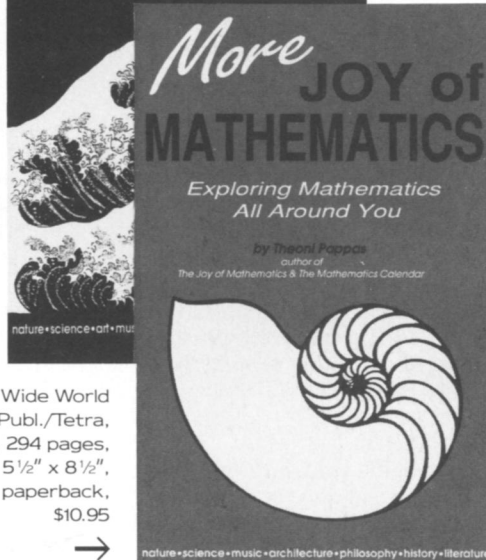
Daytime Phone _____

(used only for problems with order)

RB1464



Wide World Publ./Tetra, 1989, 237 pages, 5 1/2" x 8 1/2", paperback, \$10.95



Wide World Publ./Tetra, 1991, 294 pages, 5 1/2" x 8 1/2", paperback, \$10.95



To order by phone, call 1-800-544-4565 (Visa or MasterCard Only)