PREFACE

A Look Back and A Look Ahead

Brief History of the Resampling Method
Brief History of Statistics
A Look Ahead at the Book

Brief history of the resampling method

This book describes a revolutionary—but now fully-accepted—approach to probability and statistics. Monte Carlo resampling simulation takes the mumbo-jumbo out of statistics and enables even beginning students to understand completely everything that is done.

Before we go further, let's make the discussion more concrete with an example. Ask a class: What are the chances that three of a family's first four children will be girls? After various entertaining class suggestions about procreating four babies, or surveying families with four children, someone in the group always suggests flipping a coin. This leads to valuable student discussion about whether the probability of a girl is exactly half (there are about 105 males born for each 100 females), whether .5 is a satisfactory approximation, whether four coins flipped once give the same answer as one coin flipped four times, and so on. Soon the class decides to take actual samples of coin flips. And students see that this method quickly arrives at estimates that are accurate enough for most purposes. Discussion of what is "accurate enough" also comes up, and that discussion is valuable, too.

The Monte Carlo method itself is not new. Near the end of World War II, a group of physicists at the Rand Corp. began to use random-number simulations to study processes too complex to handle with formulas. The name "Monte Carlo" came from the analogy to the gambling houses on the French Riviera. The application of Monte Carlo methods in *teaching* statistics also is not new. Simulations have often been used to illustrate basic concepts. What *is* new and radical is using Monte Carlo methods routinely as problem-solving tools for everyday problems in probability and statistics.

From here on, the related term *resampling* will be used throughout the book. Resampling refers to the use of the observed data or of a data generating mechanism (such as a die) to produce new hypothetical samples, the results of which can then be analyzed. The term *computer-intensive methods* also is frequently used to refer to techniques such as these.

The history of resampling is as follows: In the mid-1960's, I noticed that most graduate students—among them many who had had several advanced courses in statistics—were unable to apply statistical methods correctly in their social science research. I sympathized with them. Even many experts are unable to understand intuitively the formal mathematical approach to the subject. Clearly, we need a method free of the formulas that bewilder almost everyone.

The solution is as follows: Beneath the logic of a statistical inference there necessarily lies a physical process. The resampling methods described in this book allow us to work directly with the underlying physical model by simulating it, rather than describing it with formulae. This general insight is also the heart of the specific technique Bradley Efron felicitously labeled 'the bootstrap' (Efron, 1982), a device I introduced in 1969 that is now the most commonly used, and best known, resampling method.

The resampling approach was first tried with graduate students in 1966, and it worked exceedingly well. Next, under the auspices of the father of the "new math," Max Beberman, I "taught" the method to a class of high school seniors in 1967. The word "taught" is in quotation marks because the pedagogical essence of the resampling approach is that the students discover the method for themselves with a minimum of explicit instruction from the teacher.

The first classes were a success and the results were published in 1969. Three PhD experiments were then conducted under Kenneth Travers' supervision, and they all showed overwhelming superiority for the resampling method (Simon, Atkinson, and Shevokas, 1976). Subsequent research has confirmed this success.

The method was first presented at some length in the 1969 edition of my book *Basic Research Methods in Social Science* (third edition with Paul Burstein, 1985).

For some years, the resampling method failed to ignite interest among statisticians. While many factors (including the accumulated intellectual and emotional investment in existing methods) impede the adoption of any new technique, the lack of readily available computing power and tools was an obstacle. (The advent of the personal computer in the 1980s changed that, of course.)

Then in the late 1970s, Efron began to publish formal analyses of the bootstrap—an important resampling application. Interest among statisticians has exploded since then, in conjunction with the availability of easy, fast, and inexpensive computer simulations. The bootstrap has been the most widely used, but across-the-board application of computer intensive methods now seems at hand. As Noreen (1989) noted, "there is a computer-intensive alternative to just about every conventional parametric and nonparametric test." And the bootstrap method has now been hailed by an official *American Statistical Association* volume as the only "great breakthrough" in statistics since 1970 (Kotz and Johnson, 1992).

It seems appropriate now to offer the resampling method as the technique of choice for beginning students as well as for the advanced practitioners who have been exploring and applying the method.

An early difficulty with the method had been that flipping coins and using a random-number table to simulate the resampling trials are laborious. Therefore, in 1973 Dan Weidenfeld and I developed the computer language called RESAMPLING STATS (earlier called SIMPLE STATS). Each command in the language executes an operation that simulates what the hand does when working with cards or dice. For example, instead of flipping coins in the problem described above, we first give the command GENERATE 4 1,2 A, which tells the computer to choose four '1s' and '2s' randomly and store the result in location A. Then we command COUNT a = 1 K, which tells the computer to count the number of '1s' (girls) that were selected in the previous operation and put the result in "K." Next we place the command REPEAT before the coin-flipping operations; this instructs the computer to perform (say) a thousand trials. After commanding END to complete the 1000 repetitions, we command COUNT z = 3 J, which tells the computer to count the number of trials where the outcome is three girls. That's all it takes to obtain the answer we seek.

The RESAMPLING STATS computer program is a simple language requiring absolutely no experience with computers. Yet it also provides a painless introduction to the power of computers. RESAMPLING STATS is available for both the

MacIntosh and Windows computers.

Though the term "computer-intensive methods" is nowadays used to describe the techniques elaborated here, this book can be read either with or without the accompanying use of the computer. However, as a practical matter, users of these methods are unlikely to be content with manual simulations if a quick and simple computer-program alternative such as RESAMPLING STATS is available.

The ultimate test of the resampling method is how well you, the reader, learn it and like it. But knowing about the experiences of others may help beginners as well as experienced statisticians approach the scary subject of statistics with a good attitude. Students as early as junior high school, taught by a variety of instructors and in other languages as well as English, have—in a matter of 6 or 12 short hours—learned how to handle problems that students taught conventionally do not learn until advanced university courses. And several controlled experimental studies show that, on average, students who learn this method are more likely to arrive at correct solutions than are students who are taught conventional methods.

Best of all, the experiments comparing the resampling method against conventional methods show that students *enjoy* learning statistics and probability this way, and they don't suffer statistics panic. This experience contrasts sharply with the reactions of students learning by conventional methods. (This is true even when the same teachers teach both methods as part of an experiment.)

A public offer: The intellectual history of probability and statistics began with gambling games and betting. Therefore, perhaps a lighthearted but very serious offer would not seem inappropriate here: I hereby publicly offer to 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 simple and complex numerical problems correct, when teaching similar groups of students for a limited number of class hours—say, six or ten. And if I should win, as I am confident that I will, I will contribute the winnings to the effort to promulgate this teaching method. (Here it should be noted that I am far from being the world's most skillful or charming teacher. It is the subject matter that does the job, not the teacher's excellence.) This offer has been in print for many years now, but no one has accepted it.

The early chapters of the book contain considerable discussion of the resampling method, and of ways to teach it. This material is intended mainly for the instructor; because the method is new and revolutionary, many instructors appreciate this guidance. But this didactic material is also intended to help the student get actively involved in the learning process rather than just sitting like a baby bird with its beak open waiting for the mother bird to drop morsels into its mouth. You may skip this didactic material, of course, and I hope that it does not get in your way. But all things considered, I decided it was better to include this material early on rather than to put it in the back or in a separate publication where it might be overlooked.

Brief history of statistics

In ancient times, mathematics developed from the needs of governments and rich men to number armies, flocks, and especially to count the taxpayers and their possessions. Up until the beginning of the 20th century, the term *statistic* meant the number of something—soldiers, births, taxes, or what-have-you. In many cases, the term *statistic* still means the number of something; the most important statistics for the United States are in the *Statistical Abstract of the United States*. These numbers are now known as descriptive statistics. This book will not deal at all with the making or interpretation of descriptive statistics, because the topic is handled very well in most conventional statistics texts.

Another stream of thought entered the field of probability and statistics in the 17th century by way of gambling in France. Throughout history people had learned about the odds in gambling games by repeated plays of the game. But in the year 1654, the French nobleman Chevalier de Mere asked the great mathematician and philosopher Pascal to help him develop correct odds for some gambling games. Pascal, the famous Fermat, and others went on to develop modern probability theory.

Later these two streams of thought came together. Researchers wanted to know how accurate their descriptive statistics were—not only the descriptive statistics originating from sample surveys, but also the numbers arising from experiments. Statisticians began to apply the theory of probability to the accuracy of the data arising from sample surveys and experiments, and that became the theory of *inferential statistics*.

Here we find a guidepost: probability theory and statistics are relevant whenever there is uncertainty about events occurring in the world, or in the numbers describing those events.

Later, probability theory was also applied to another context in which there is uncertainty—decision-making situations. Descriptive statistics like those gathered by insurance companies—for example, the number of people per thousand in each age bracket who die in a five-year period—have been used for a long time in making decisions such as how much to charge for insurance policies. But in the modern probabilistic theory of decision-making in business, politics and war, the emphasis is different; in such situations the emphasis is on methods of *combining* estimates of probabilities that depend upon each other in complicated ways in order to arrive at the best decision. This is a return to the gambling origins of probability and statistics. In contrast, in standard insurance situations (not including war insurance or insurance on a dancer's legs) the probabilities can be estimated with good precision without complex calculation, on the basis of a great many observations, and the main statistical task is gathering the information. In business and political decision-making situations, however, one often works with probabilities based on very limited information—often little better than guesses. There the task is how best to combine these guesses about various probabilities into an overall probability estimate.

Estimating probabilities with conventional mathematical methods is often so complex that the process scares many people. And properly so, because its difficulty leads to errors. The statistics profession worries greatly about the widespread use of conventional tests whose foundations are poorly understood. The wide availability of statistical computer packages that can easily perform these tests with a single command, regardless of whether the user understands what is going on or whether the test is appropriate, has exacerbated this problem. This led John Tukey to turn the field toward descriptive statistics with his techniques of "exploratory data analysis" (Tukey, 1977). These descriptive methods are well described in many texts.

Probabilistic analysis also is crucial, however. Judgments about whether the government should allow a new medicine on the market, or whether an operator should adjust a screw machine, require more than eyeball inspection of data to assess the chance variability. But until now the teaching of probabilistic statistics, with its abstruse structure of mathematical formulas, mysterious tables of calculations, and restrictive assumptions concerning data distributions—all of which separate the student from the actual data or physical process under consideration—have been an insurmountable obstacle to intuitive understanding.

Now, however, the resampling method enables researchers and decision-makers in all walks of life to obtain the benefits of statistics and predictability without the shortcomings of conventional methods, free of mathematical formulas and restrictive assumptions. Resampling's repeated experimental trials on the computer enable the data (or a data-generating mechanism representing a hypothesis) to express their own properties, without difficult and misleading assumptions.

A look ahead at the book

Now for a brief look ahead at the book. The Introduction and Chapter 1 introduce you to probability and statistics, and to the resampling method of estimating the probabilities of events. If you have studied statistics before, you can skip these chapters. Chapters 2 and 3 briefly discuss some of the basic concepts which statisticians use in their thinking. You might choose to move right on to the method itself starting in Chapter 4, and return later to read about these concepts after you have experienced them in practice. Chapters 4 through 10 get to work applying the Monte Carlo method to simple problems in estimating the probability of compound events, such as the chance of having three girls in four children. Chapter 11 begins the study of statistical inference.

So—good luck. I hope that you enjoy the book and profit from it.