Proceedings of the 2002 Winter Simulation Conference
E. Yücesan, C.-H. Chen, J. L. Snowdon, and J. M. Charnes, eds.

PANEL: USING SIMULATION TO TEACH PROBABILITY
SESSION 1: WORDS
SESSION 2: DEEDS

Matthew Rosenshine (Organizer and Moderator)

Department of Industrial and Manufacturing Engineering
The Pennsylvania State University
University Park, PA 16802, U.S.A.

Russell R. Barton (Panelist)

The Smeal College of Business Administration
The Pennsylvania State University
University Park, PA 16802, U.S.A.

David Goldsman (Panelist)

School of Industrial and Systems Engineering
Georgia Institute of Technology
Atlanta, GA 30332-0205, U.S.A.

Lawrence M. Leemis (Panelist)

Department of Mathematics
College of William and Mary
P.O. Box 8795
Williamsburg, VA 23817-8795, U.S.A.

Barry L. Nelson (Panelist)

Department of Industrial Engineering & Management Sciences
Northwestern University
Evanston, IL 60208-3119, U.S.A.

ABSTRACT

This panel has been put together to promote the use of simulation as a teaching tool to expedite the learning and, more importantly, the understanding of probability theory. "In a nutshell," the thesis upon which this panel is based is that the simulation approach is more effective than a mathematical approach on a stand-alone basis. It also dominates any statistical approach as a pedagogical tool.

1 INTRODUCTION (MATTHEW ROSENSHINE)

Elementary probability theory is usually taught as a one-semester course. With the rise in importance of probability in scientific, technical, and business areas, it is likely to remain at this level for the foreseeable future. As the importance of probability rose, the variation in the rigor of mathematical preparation of the students taking this type of course has increased while the average level of rigor has decreased.

The response to the decrease in mathematical rigor has been basically non-existent but fortunately the increase in variation along with the necessity to be more inclusive led to a recognition that axiomatic probability needed some help. Unfortunately, some of the help did not help. The use of statistics to provide an introduction to the study of probability was well-intentioned but confusing. Even worse, many of the confused students did not know that they were confused.

The replacement of many derivations and proofs with discussions and less rigorous proofs was helpful. The elimination of some proofs entirely was also helpful. A proof of the central limit theorem is of little use to a student who does not understand what the sum of random variables means. Unless the proof provides understanding, which for almost all students it does not, it is of little use except as a mathematical exercise—albeit elegant.

So here we are with what appears to be a good idea—use simulation to teach probability. Why is it a good idea? Let me offer a few reasons, each somewhat convincing in its own right. Collectively, their appeal soars!

1. It occurred to me during a ten-plus year period during which I have been trying to teach middle school and high school teachers to teach probability. After it dawned on me, I used it to teach eighth and ninth graders.

2. It has occurred to others independently, I assume, since I have never met or spoken to some of the people who have written in recent years about using simulation as a teaching aid for probability.
3. Simulation has only recently become a feasible teaching tool for probability with the advent of high-speed desk top computers.
4. No one to whom I have spoken about it has expressed a negative opinion about the idea. Many have said that they are, in fact, doing this to some extent in the courses they teach.
5. Probably the most convincing argument for using simulation as a pedagogical tool lies in the response I received long ago when I questioned why simulation was being proposed to obtain a result that could be obtained analytically—"Any idiot can understand simulation."

With the preceding comments as background, this panel of educators will attempt through the spoken word (Session 1) to provide contexts in which simulation can be a valuable asset in teaching probability, and deeds (Session 2) to provide demonstrations of simulations that provide concrete back-up for the words. Panel-member statements are provided in the following sections to facilitate the panel discussion and audience interaction with the panelists.

2 RUSSELL R. BARTON

My remarks focus on the use of simulation to understand and compute conditional probabilities. Many real-world situations are described by conditional probabilities. For example, conditional probabilities determine the fraction of defective products reaching consumers, given the fraction defective that are produced and the type I and type II error probabilities of quality tests performed before shipping the product. Conditional probability calculations can be difficult to teach students to perform correctly, in part because they can violate intuition. During the late 1980s the "Let's Make a Deal" door-choice problem led to conflicting opinions about the correct calculation of conditional probabilities. After the contestant chose an unseen prize behind Door 1, 2, or 3, the announcer would reveal one of the remaining doors (never the Grand Prize door), and allow the contestant to switch to the remaining unopened door. For this scenario, the contestant's chances of winning the Grand Prize would be improved by switching, regardless of what was revealed.

This counterintuitive result was debated in professional periodicals for a number of months; the correct solution was not obvious, even to some folks with significant exposure to probability calculations! A convincing case was provided by simulating the 'always switch' and 'never switch' options, which showed in a concrete fashion the superiority of the switching strategy.

This experience motivated me to teach my students to use simulation to compute conditional probability. I begin by describing calculations for parallel and serial systems assuming independence of failure probabilities for the components. Then I describe the concept of conditional probability, give a graphical representation of the ratio of measures, and then define independence as the equality of conditional and unconditional probabilities. After calculating conditional probabilities in several simple examples, I describe the result for the "Let's Make a Deal" decision. Then I ask the students to compute probabilities for a reliability problem from Kolarik (1995).

2.1 The Car-Trip Example

The students are asked the following question. *D. Event owns two cars, one old, the other older. Each day, he uses a car to get to school and back. Car 1 is old and has probability 0.79 of starting on any day, while Car 2, even older, starts with probability 0.71. For either car, once it is started, the chance that it completes a trip to or from school is 0.95. Compute the probability that D. Event will make it to school and back on any particular day.*

Most students assume that the reliability system consists of two cars in parallel, and they perform calculations based on the parallel system shown in Figure 1. The computed system reliability is 1 – [(1 – 0.56325025)(1 – 0.45495025)], approximately 0.762. This value is greater than the correct probability by more than 0.10. After the students present this solution, I ask them whether D. Event can start both cars at the same time, and, assuming both start, drive both to school simultaneously. At this point the students realize the error in applying a parallel system reliability model, but they are uncertain about a solution. When I ask them which car Event will try first, they do not realize that i) a vehicle-selection rule must be decided before the probability can be calculated, and ii) the vehicle-selection rule (try Car 1 first, try Car 2 first, or randomly select a car to try) affects the value. Event has a better chance of making the round trip by trying Car 1 first every day.

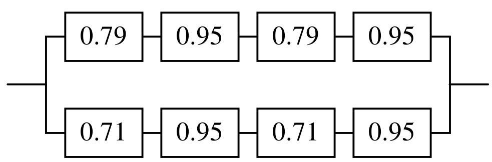

Figure 1: Typical Incorrect System Reliability Formulation for the Car-Trip Example

To compute the probability, I build a simulation model together with the students. Figure 2 shows the car-trip example modeled in Arena. The model takes only a few minutes to construct, and follows from a description of D. Event's actions. Deciding which car to start first occurs naturally in this process, and the students are not tempted

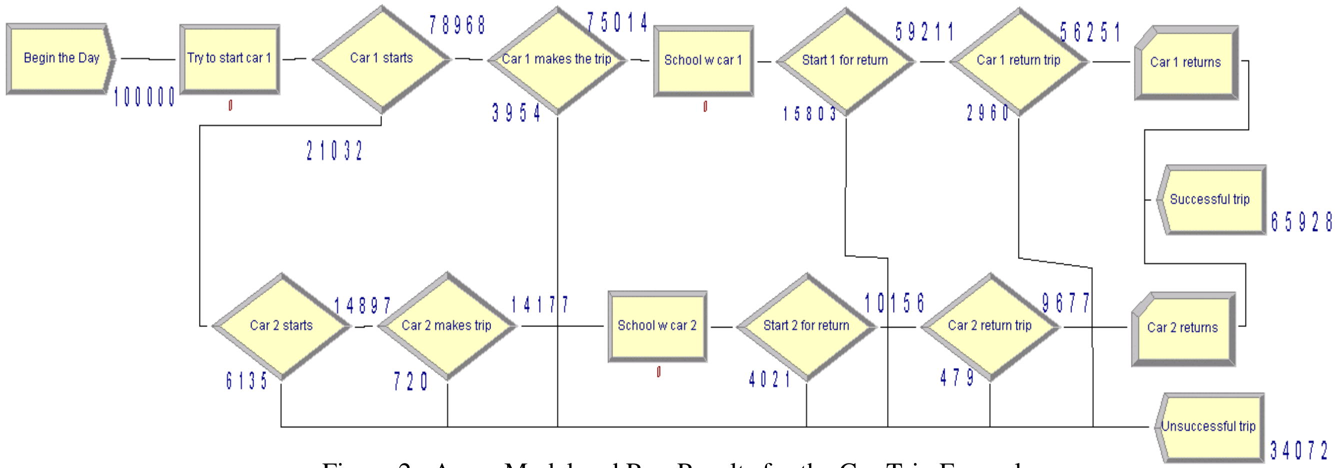

Figure 2: Arena Model and Run Results for the Car-Trip Example

to 'split' the driver in two in order to try both cars at the same time. Run time for 100,000 repetitions (days) takes less that a minute (in Fast Forward mode) on my 1GHz Pentium III laptop. The conditional probabilities and unconditional probabilities can be constructed using the event counts that are provided at each block automatically by the software. One such run provided the numbers in Figure 2, which were used to construct the results shown in Table 1.

Table 1. Simulation-Based and Conditional Probability Calculations for the Example

Quantity	Number in 100,000	Estimated Probability	Exact Probability
Car 1 makes round trip	56251	0.56251	0.56325025
Car 2 makes round trip	9677	0.09677	0.0955395525
Successful round trip	65928	0.65928	0.6587898025

For comparison, the conditional probability calculations (assuming that the starting probabilities of the two cars are independent) are

P(Successful round trip)
= P(Car 1 makes round trip | Car 1 starts)P(Car 1 starts)
+ P(Car 2 makes round trip | Car 1 fails and Car 2 starts)
× P(Car 1 fails and Car 2 starts)
$= (0.95)(0.79)(0.95)(0.79) + (0.95)(0.71)(0.95)(0.21)(0.71)$
$= 0.56325025 + 0.0955395525$
$= 0.6587898025.$

The results of direct conditional probability calculations are shown in the last column of Table 1. The students see that the probabilities estimated by simulation provide results that are close to the actual probabilities.

2.2 Experience

My students found it easy to understand the conditional-probability calculations for the car-trip example when they were described by a flow diagram. Today's simulation packages make it easy to convert a flow diagram into a discrete-event simulation model, and run times are short for approximations accurate to two decimal places. Further, the built-in animation capability of many simulation packages makes it more likely that the student will model the situation correctly. The car-trip exercise convinces students of the value of discrete-event simulation for the calculation of complex probabilities, and piques their interest in discrete-event simulation software. I have more confidence that students will calculate conditional probabilities correctly when they build and view animated simulation models that represent real situations.

3 DAVID GOLDSMAN

Simulation certainly helps to motivate concepts and to answer interesting questions in the probability-classroom environment. In this section, we will discuss relevant several examples that one can incorporate into probability lectures.

3.1 The Birthday Problem

For instance, we can easily use simulation to supplement the discussion of the following classic combinatorial problem. Suppose we have n people in a room. What is the probability that at least two will have the same birthday? To keep things reasonable, we shall assume that all 365 birthdays have equal probability (sorry, February 29). As is well known, the surprisingly low value of $n = 23$ yields a (slightly greater than) 50-50 chance that there will be a match. Figure 3 illustrates a realization of the birthday problem in which we sequentially sample simulated people in a room until a match has been achieved. In the current realization, we see that a sample of 24 was required before a match finally occurred. Students can use this example to run multiple realizations, and can quickly get a feel for the variation of the results between runs.

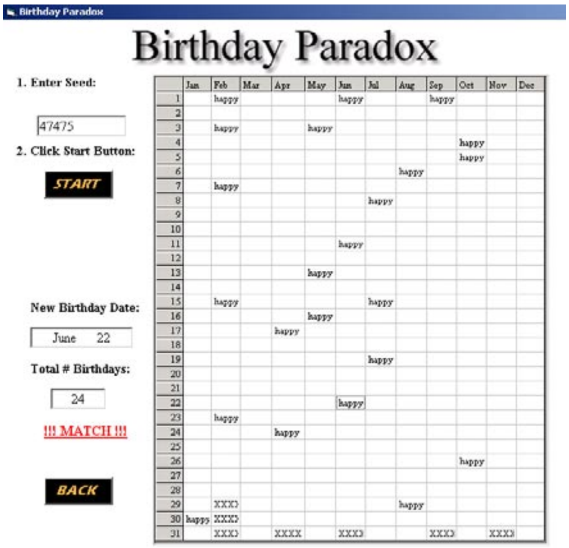

Figure 3: Simulating the Birthday Problem

3.2 Estimating π

The next example shows that we can use simulation to enhance the discussion of an elementary probability calculation. We will now estimate π using Monte Carlo simulation in conjunction with a basic geometric relation. Referring to Figure 4, consider a unit square with an inscribed circle, both centered at (1/2, 1/2). If one were to throw darts randomly at the square, the probability that a particular dart will land in the circle is $\pi/4$, the ratio of the circle's area to that of the square. How can we use this simple fact to estimate π? We shall use Monte Carlo simulation to throw many darts at the square. Specifically, generate independent pairs of independent uniform (0,1) random variables, (U_{11}, U_{12}), (U_{21}, U_{22}), …, so that these pairs will fall randomly on the square. If, for pair i, it happens that

$$(U_{i1} - 1/2)^2 + (U_{i2} - 1/2)^2 < 1/4$$

then that pair will also fall within the circle. Suppose we run the experiment for n pairs (darts). Let $X_i = 1$ if pair i satisfies the above inequality, i.e., if the i^{th} dart falls in the circle; otherwise, let $X_i = 0$. Now count up the number of darts $X = X_1 + \cdots + X_n$ falling in the circle. Clearly, X has the binomial distribution with parameters n and $p = \pi/4$. Then the proportion $p' = X/n$ is the maximum-likelihood estimate for $p = \pi/4$; and so the maximum-likelihood estimator for π is just $\pi' = 4p'$. If, for instance, we conducted $n = 500$ trials and observed $X = 397$ darts in the circle, as in Figure 4, our estimate would be $\pi' = 3.176$. We usually run this example in class with at least 10,000 darts; it most often happens that the estimator is even closer to the true value of π.

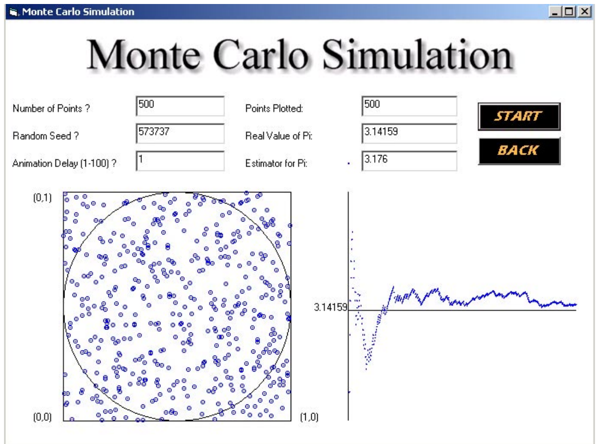

Figure 4: Throwing Darts to Estimate π

3.3 Monte Carlo Integration

Here, we show the students how one can use probability in a particularly novel way, viz., to conduct an integration exercise. To this end, consider the integral

$$I = \int_a^b f(x)dx = (b-a)\int_0^1 f(a+(b-a)u)du.$$

As described in Figure 5, we shall estimate the value of this integral by summing up n rectangles, each of width $1/n$, centered randomly at point U_i on [0, 1], and of height $f(a + (b - a)U_i)$. Then an estimate for I is

$$I' = \frac{b-a}{n}\sum_{i=1}^{n} f(a+(b-a)U_i)\cdot$$

In fact, it turns out that I' is an unbiased estimator for I, i.e., $E[I'] = I$ for all n. This makes I' an intuitive and attractive estimator—one that probability students will find easy to understand. Figure 5 shows how one could use simulation to carry out the integration of $\sin(\pi x)$ with $n = 64$ uniform samples. Although the estimate (0.5886) differs from the actual integral value (0.6366), a larger sample size will invariably do better, a fact that the students can easily discover.

3.4 A Single-Server Queue

As a final motivational example, we simulate the behavior of a single-server queueing system—to show how one can combine basic probability and simulation techniques to

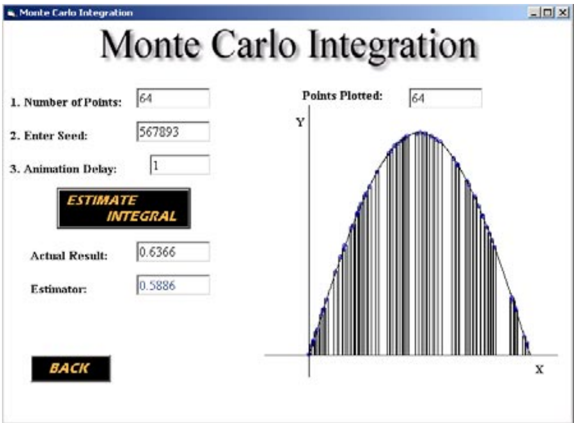

Figure 5: Monte Carlo Integration to Evaluate $\int_0^1 \sin(\pi x)dx$

study a "real-life" system. Suppose that customers arrive at a bank one-at-a-time, and queue up in front of a single teller to be processed sequentially in a first-come-first-served manner. Figure 6 traces the evolution of the system as time progresses. The associated table keeps track of the times at which customers arrive, begin service, and leave. The graphs keep track of the status of the system as a function of time; in particular, we find plots of the queue length and server utilization. We have found that students understand this example with little trouble.

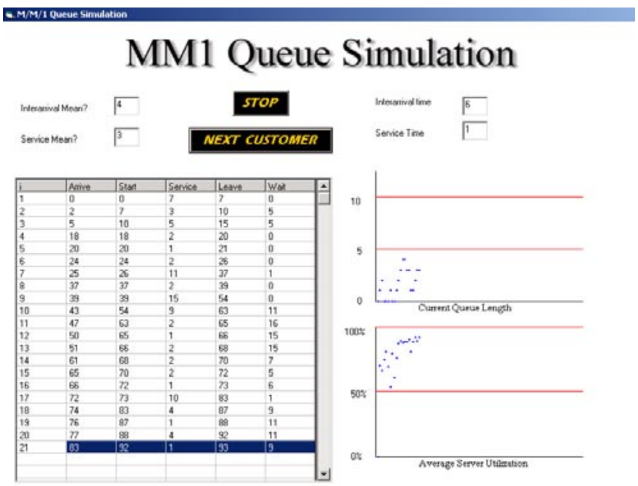

Figure 6: Simulating a Single-Server Queue

4 LAWRENCE M. LEEMIS

4.1 Axiomatic vs. Simulation Approaches

Monte Carlo and discrete-event simulation are reasonable approaches to introducing students to probability. Monte Carlo can be used to cover the topics in a traditional calculus-based probability class, and discrete-event simulation can be used to cover the topics in a traditional stochastic processes class. I will refer to the standard calculus-based approach to teaching probability as the "axiomatic approach."

Here are five advantages to the simulation approach to teaching probability:

1. The relative-frequency approach to determining probabilities is very intuitive to beginning students. All of us have used the experimental approach to determining the likelihood of events from a very early age, so simulating events on the computer comes as a logical next step.
2. Random variables are easily introduced after the relative-frequency approach is well understood. Once the binary aspect of an event's occurring or not occurring has been established using coins or dice using the simulation approach, it is then reasonable to consider the number of spots that appear on the up face when a die is cast, which leads to the introduction of discrete random variables. Continuous random variables and other quantities (e.g., expected values) come along next.
3. There is almost no limit to the complexity of the problems that can be addressed. This is certainly not the case in the axiomatic approach. One seemingly minor twist to the assumptions in a particular problem can sink the axiomatic approach.
4. A simulation-based approach to teaching probability prepares students for statistics in several ways that are not possible using the axiomatic approach. Here are three examples:
   a. Students completing a simulation-based probability class will have an intuitive notion of sampling variability since they have seen it occur in every simulation that they have run. The notion of observed data also having sampling variability follows.
   b. Point and interval estimation will also be familiar after a simulation-based class since these two concepts must be included in any serious Monte Carlo analysis of a problem.
   c. Simulation allows the assimilation of the bootstrapping approach (Efron and Tibshirani 1993) when a student moves from probability to statistics. This transition is much more awkward when the student comes from the axiomatic approach.
5. The simulation approach forces students to program. This means that a discussion of a random-number generator is appropriate, along with a discussion of random-variate generation. The latter is a bit tricky without a formal definition of a random variable, which is part of the axiomatic approach.

The last point concerning programming points to an important side-issue: what platform should be used? One extreme is to use a standard algorithmic language such as C or FORTRAN. Two advantages of this choice are the programming flexibility and the ability to move seamlessly to discrete-event simulation. The disadvantage of this choice is that the class may begin to look more like a programming course than a probability course. The other extreme is to use a statistical language (such as Splus) or a computer algebra system (such as Maple). In both cases, the programming time is cut considerably, but at the cost of flexibility. Consider the simple problem of generating a *p*-value associated with the Kolmogorov-Smirnov goodness-of-fit test for six U(0,1) random variates. This can be done with the rather cryptic Splus command:

```
ks.gof(runif(6),  distribution  =  "uniform",
min = 0, max = 1)$p.value
```

or coded in C. The former approach keeps the discussion at a high level, while the latter forces a student to dive into the details.

There are three disadvantages to the simulation approach to teaching probability that also must be considered:

1. There are times when the axiomatic approach is faster and more appropriate. I would certainly not want a student to begin programming when asked for the probability of exactly two heads appearing in three coin flips. There must be a mix of the axiomatic and simulation approaches. In a comprehensive first probability course, a student would know the mathematical/axiomatic approach to the sample problem stated above, as well as knowing that the Splus statement

   ```
   dbinom(2, 3, 0.5)
   ```

   gives the analytic solution and that the statement

   ```
   sum(rbinom(1000, 3, 0.5) == 2) / 1000
   ```

   gives the Monte Carlo point estimate of the probability using 1000 replications of the experiment.
2. To my knowledge, there is no textbook available that integrates these two approaches.
3. Simulation requires a 100-fold increase in the number of replications in order to get another digit of accuracy. There are going to be applications where getting exact results is appropriate, and only the axiomatic approach can deliver.

I close with mentioning that there is still another way of teaching probability that minimizes the reliance on the axiomatic approach. Rose and Smith (2002) and Glen, Evans, and Leemis (2001) have developed languages (MathStatica, which is Mathematica-based, and APPL, which is Maple-based) that are capable of determining exact probability results by manipulating random variables.

4.2 Experience

We teach a C-based discrete-event simulation course at William & Mary, CSCI 426, annually to about 30 students, mainly computer-science majors. Since the computer-science curriculum is very crowded, most of the students entering this class have had minimal exposure to probability. Although discrete-event simulation is the emphasis, two or three lectures are spent on random-number generation (emphasizing Lehmer generators), and two or three lectures are spent on Monte Carlo simulation.

In addition, I have taught the introductory calculus-based probability class, Math 401, to primarily mathematics majors. I often use the Monte Carlo simulation approach to verify analytic solutions, then show how a small change in the assumptions to the problems makes that axiomatic approach intractable, yet the Monte Carlo approach remains viable. Approaching simulation as a way to check analytic results and estimate the solution to a difficult problem gives the students a healthy view of simulation—it should be relied on when appropriate.

5 BARRY L. NELSON

I do not teach introductory probability, but for many years I have taught the course that follows it in most Industrial Engineering programs: stochastic processes. I am a firm believer in integrating simulation into such a course, and I wrote a textbook supporting that approach. Here are my top five reasons why:

5.1 Visualization

I had already completed my Ph.D. before I developed any intuition about how a Markovian queueing process would actually look. To gain that sense I coded up an animated *M*/*M*/1 queue on a Commodore Vic 20 computer in Basic. I distinctly remember being surprised at how bursty the process was, and finally understanding what "bursty" meant. This sort of intuition is critical for engineering students because they need to be able to recognize when particular models are appropriate, and no amount of talking about stochastic processes will develop this sense.

5.2 Algorithmic Representation of Probability

Here are some essential concepts in probability that many students fail to grasp, and how I think simulation can help them:

a. A random variable is a function from the sample space to the real numbers: Simulations map random numbers into sample paths.

b. The random variable X has probability distribution F: We can simulate observations of X from $X = F^{-1}(U)$, where U is a Uniform(0,1) variate.
c. The random variables are dependent: The random variables are functions of some of the same random numbers.
d. The random variables are identically distributed: The random variables are the same function of different random numbers.

5.3 Sensitivity or Insensitivity of Results to Assumptions

How much does it matter if we assume a Poisson arrival process? Or Markovian state changes? Is a model useless if the assumptions behind it are not rigorously satisfied? Do models really give better estimates of relative differences in performance than they do of actual performance? The best way for students to obtain some idea of the answers to these questions is to let them use simulation to test and discover.

5.4 Connecting Probability to Statistics

Because simulation generates data from a probability model, an understanding of simulation makes it easier to understand the reverse process of using data to infer something about an underlying model.

5.5 Integrating Probability and Simulation Supports a Unified Treatment of Stochastic Modeling and Analysis

Last year I wrote a panel piece that included the following argument against having computer simulation courses for undergraduate engineering students (Altiok et al. 2001):

> "I am a proponent of generic courses in stochastic modeling and analysis, in which mathematical, numerical and simulation solution techniques all appear. I have been teaching a two-quarter (20-week) sequence in this way for over six years, and I am convinced that there are at least two features that are critical to making it work:
>
> For every stochastic modeling problem, start by thinking about how to simulate it. Simulation (inputs, events, states, etc.) provides the formulation language, much like the decision-variable, objective-function and constraint concepts do for optimization. Simulation is also intuitive. We then teach students to recognize those situations in which a mathematical or numerical solution is possible or appropriate.
>
> When a large-scale simulation is required, force students to do a rough-cut model prior to simulating. (I am pretty sure I stole this idea from Lee Schruben.) Sometimes the rough-cut model is just plugging in mean values for all the stochastic stuff, or deriving best-case and worst-case bounds. More often it involves using some sort of simplified model, such as an M/M-type queue. This allows students to understand that both approaches apply to the same types of problems. They also see that the numbers that come out of the simulation typically do not match the rough-cut model—demonstrating that there is a reason for simulation—but they also see that the best solution, as determined by the rough-cut model, is often identical to the one indicated by the far more detailed simulation."

REFERENCES

Altiok, T., W.D. Kelton, P. L'Ecuyer, B.L. Nelson, B.W. Schmeiser, T.J. Schriber, L.W. Schruben, and J.R. Wilson. 2001. Various ways academics teach simulation: are they all appropriate? In *Proceedings of the 2001 Winter Simulation Conference*, ed. B.A. Peters, J.S. Smith, D.J. Medeiros, and M.W. Rohrer, 1580-1591. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers.

Efron, B., and R.J. Tibshirani. 1993. *An introduction to the bootstrap*. New York: Chapman & Hall.

Glen, A.G., D.L. Evans, and L.M. Leemis. 2001. APPL: a probability programming language. *The American Statistician* 55: 156-166.

Kolarik, W.J. 1995. *Creating quality: concepts, systems, strategies and tools*. New York: McGraw Hill.

Rose, C., and M. Smith. 2002. *Mathematical statistics with Mathematica*. New York: Springer.

AUTHOR BIOGRAPHIES

MATTHEW ROSENSHINE is Professor Emeritus of Industrial Engineering at the Pennsylvania State University. His interests are now centered on teaching probability at the middle-school and high-school levels. He is active in INFORMS outreach programs that seek to encourage early exposure of students to "mathematics of operations research and management science." His email address is <mxr2@psu.edu>.

RUSSELL R. BARTON is a Professor and Associate Dean for Research and Ph.D./M.S. Programs in the Smeal College of Business Administration at the Pennsylvania State University. He has a B.S. degree in Electrical Engineering from Princeton and M.S. and Ph.D. degrees in Operations Research from Cornell. Before entering academia, he spent twelve years in industry. He is the past Secretary/Treasurer for the INFORMS College on Simulation.

His research interests include applications of statistical and simulation methods to system design and to product design, manufacturing and delivery. His email address is <rbarton@psu.edu>.

DAVID GOLDSMAN is a Professor in the School of Industrial and Systems Engineering at the Georgia Institute of Technology. His research interests include simulation output analysis and ranking and selection. He is the Simulation Department Editor for *IIE Transactions*, and an Associate Editor for *Operations Research Letters*. He is an active participant in the Winter Simulation Conference, and is currently on the WSC Board. His e-mail address is <sman@isye.gatech.edu>, and his web page is <www.isye.gatech.edu/~sman>.

LAWRENCE M. LEEMIS is a Professor of the Mathematics Department at the College of William & Mary. He received his BS and MS degrees in Mathematics and his Ph.D. in Industrial Engineering from Purdue University. He has also taught at Baylor University, The University of Oklahoma, and Purdue University. His consulting, short-course, and research-contract work includes contracts with AT&T, NASA/Langley Research Center, Delco Electronics, Department of Defense (Army, Navy), Air Logistic Command, ICASE, Komag, Federal Aviation Administration, Tinker Air Force Base, Woodmizer, Magnetic Peripherals, and Argonne National Laboratory. His research and teaching interests are in reliability and simulation. He is a member of ASA, IIE, and INFORMS. His email and web addresses are <leemis@math.wm.edu> and <www.math.wm.edu/~leemis>.

BARRY L. NELSON is a Professor in the Department of Industrial Engineering and Management Sciences at Northwestern University, and is Director of the Master of Engineering Management Program there. His research centers on the design and analysis of computer simulation experiments on models of stochastic systems. He has published numerous papers and two books. Nelson has served the profession as the Simulation Area Editor of *Operations Research* and President of the INFORMS (then TIMS) College on Simulation. He has held many positions for the Winter Simulation Conference, including Program Chair in 1997 and current membership on the Board of Directors. His e-mail and web addresses are <nelsonb@northwestern.edu> and <www.iems.northwestern.edu/~nelsonb/>.