Proceedings of the 2001 Winter Simulation Conference
B. A. Peters, J. S. Smith, D. J. Medeiros, and M. W. Rohrer, eds.

RESAMPLING METHODS FOR INPUT MODELING

Russell R. Barton

The Harold and Inge Marcus Department of Industrial and Manufacturing Engineering
The Pennsylvania State University
310 Leonhard Building
University Park, PA 16803, U.S.A.

Lee W. Schruben

Department of Industrial Engineering and Operations Research
4117 Etcheverry Hall
The University of California, Berkeley
Berkeley, CA 94720-1777, U.S.A.

ABSTRACT

Stochastic simulation models are used to predict the behavior of real systems whose components have random variation. The simulation model generates artificial random quantities based on the nature of the random variation in the real system. Very often, the probability distributions occurring in the real system are unknown, and must be estimated using finite samples. This paper shows three methods for incorporating the error due to input distributions that are based on finite samples, when calculating confidence intervals for output parameters.

1 INTRODUCTION

Stochastic simulation models are used to predict the behavior of real systems whose components have random variation. The simulation model generates artificial random quantities based on probability distributions that represent the nature of the random variation in the real system. Very often, the probability distributions occurring in the real system are unknown, and must be estimated using finite samples. Any finite sample estimates the true distribution with some error. The nature of the error in the empirical distribution's approximation to the true distribution function is well understood, yet this error is typically ignored in the analysis of simulation output (e.g., in determining confidence intervals for the mean value of some system performance measure).

This paper extends the discussion in Barton and Schruben (1993). Three possible approaches not currently used in simulation practice are discussed: direct system resampling, bootstrap resampling of empirical distribution functions, and uniformly randomizing empirical distribution functions. Direct resampling uses new samples of empirical data for each simulation replication. Bootstrap resampling constructs artificial new samples for each replication by sampling (with replacement) from the original sample. Randomizing empirical distribution functions makes (random) changes to the increments of the empirical distribution functions for each simulation replication. Only direct resampling explicitly accounts for the uncertainty in the input but it may require much more system data than is readily available. In our example, bootstrap and uniform randomization schemes provide greatly improved confidence interval coverage without requiring additional system data. Neither scheme is a panacea since they involve additional computation. However, the alternative may be invalid results. The methods focus on cases where the input distributions are empirical rather than parametric, but these methods can be extended to the case of parametric input distributions whose parameters are estimated from (finite) data sets.

The next section presents a review of input modeling issues and the form of empirical distributions that we will use in this paper. Section 3 illustrates how a simple confidence interval calculation can provide erroneous results when the input distributions are based on finite empirical samples.

Section 4 presents the three candidate analysis methods, along with a discussion of the difficulties with these approaches. Section 5 shows the performance of these methods for some simple capacitated queueing models. Findings and future work are presented in Section 6.

2 INPUT MODELS

Approaches to modeling random processes range from not assuming anything to assuming everything with several intermediate positions. At one extreme is the conservative approach of using actual system performance records to drive a simulation run. This is called a trace driven simulation experiment. At the other extreme, the simulation results can be viewed as conditional on specific scenarios that are concocted with all the probabilistic and/or deterministic characteristics of the input processes specified.

2.1 Parametric Input Distributions

Driving simulations with artificially generated samples from parametric probability distributions fitted to system data is probably the most commonly used method in practice. This method presumes the availability of a sample of independent observations of system random variables. There are several software packages on the market that take a data set as input and suggest several common parametric probability models that may fit the data. Widely available algorithms for generating artificial samples from the fitted distributions are then used to drive the simulation. This method for generating random simulation input is inexpensive and fast and it facilitates easy input sensitivity analysis, replication, and time compression. Furthermore, the results of the simulation study can be generalized to provide answers to "what if" questions.

The major disadvantages of this method include model selection error, parameter estimation error, loss of important serial and cross dependencies in the data, and difficulty in convincing others of a model's validity. Shanker and Kelton (1991) demonstrate the consequences of generating data from an incorrect distribution using the Pollaczek-Khintchine formula for the M/G/1 queue.

2.2 Empirical Input Distributions

Difficulties in validating parametric probability distribution assumptions have, in part, motivated simulation modelers and software developers to use empirical distributions based on samples of data from the real system. There are many software packages on the market for doing this. In fact, the first simulation language in widespread use, GPSS, initially facilitated *only* this method of pseudo-random input generation using tabled functions. The nature of the error in the empirical distribution's approximation to the true distribution function is well understood, yet this error is typically ignored in the analysis of simulation output.

We denote the observed real system data, ordered from smallest to largest, as $x_1 \le x_2 \le \ldots \le x_n$. For convenience, we consider estimating the distribution function as a two step process: first we select point estimates for the distribution function at the observed real system values, say $\hat{F}^{(p)}(x_i)$, then we interpolate the distribution function between these observed values. Several common point estimates for the first step are:

$$\hat{F}^{(p)}(x_i) = i/n, \tag{1-a}$$

$$\hat{F}^{(p)}(x_i) = i/(n+1), \text{ and} \tag{1-b}$$

$$\hat{F}^{(p)}(x_i) = (i - .5)/n. \tag{1-c}$$

Choosing among these approximations (or others) amounts to choosing the total probability that is to be assigned to each of the following intervals: $I_1 = (\infty < x \le x_1)$; $I_2 = (x_1 < x \le x_2)$; … $I_{n+1} = (x_n < x < \infty)$. For the approximations above: (1-a) assigns $1/n$ to intervals I_1 through I_n and 0 to interval I_{n+1}; (1-b) assigns $1/(n+1)$ to every interval; (1-c) assigns $.5/n$ to the two boundary intervals I_1 and I_{n+1}, and assigns a probability of $1/n$ to the other intervals.

Given point estimates for the distribution function, $\hat{F}^{(p)}(x_i)$, at the real system observations, there are many choices for interpolation to other values of x, including linear interpolation and kernel smoothing. The conclusions in this paper should hold for many if not all of these techniques. Here we will describe two natural interpolations that are the most commonly used in simulation. Let $\lambda(x)$ denote the index of the nearest real system observation at or below an arbitrary constant, x. Explicitly: $\lambda(x) = \max(i \mid x_i \le x)$, which we will denote simply as λ since the value of x will be obvious. For discrete random variables in the set $\{x_i\}$, our interpolated EDF will be the usual empirical distribution function,

$$\begin{aligned} \hat{F}(x) &= \hat{F}^{(p)}(x_\lambda) \quad \text{for } x_1 \le x \le x_n \\ &= 0 \qquad\qquad\quad x < x_1 \\ &= 1 \qquad\qquad\quad x > x_n . \end{aligned} \tag{2-a}$$

For continuous random variables on the interval (x_0, x_{n+1}), we will linearly interpolate values for our EDF between the point estimates, $\hat{F}^{(p)}$, at the real system observations.

$$\hat{F}(x) = \alpha \hat{F}^{(p)}(x_\lambda) + (1-\alpha)\hat{F}^{(p)}(x_{\lambda+1}) \tag{2-b}$$

with

$$\alpha = (x_{\lambda+1} - x)/(x_{\lambda+1} - x_\lambda),$$

and the two added points,

$$\hat{F}(x_0) = 0, \ \hat{F}(x_{n+1}) = 1.$$

Another way of smoothing the empirical distribution function is to fit Bézier curves to the EDF (Wagner and Wilson 1995). The general conclusions in this paper should also apply to this approach to modeling the input distributions for simulating real systems. For the examples in Section 5, we will use EDF (1-b) smoothed with approximation (2-b) with $x_0 = 0$ and $x_{n+1} = x_n + (x_n - x_{n-1})$.

3 INFERENCE ERRORS FOR SIMULATIONS USING INPUT DISTRIBUTIONS BASED ON FINITE SAMPLES

Consider the following scenario. A simulation model has been constructed for a capacitated (capacity = 10) single-server queue. The model uses empirical distributions for the interarrival and service distributions. Samples of size n = 100 are used to construct the empirical distributions. Suppose in this scenario that the true probability models generating the data (unknown to the simulationist) are exponential with rates 1.0 and 0.7 for service and interarrival, respectively. For this scenario, the value for E(W) is approximately 3.04.

The simulation model will be used to develop a 90% confidence interval for E(W), the average time in the system. The conventional strategy (*nr*: *n*o *r*esampling, Student t interval) is used to construct a confidence interval for W, based on an experiment consisting of ten replications, with 5000 customers per replication. Sample means do not include the first 1000 customers, in order to reduce initialization bias. In this scenario, if the modeler knew the actual form of the interarrival and service distributions, then the simulation runs could use these true distributions for generating random arrival and service times. The conventional confidence interval for W, after deleting the initial transient of c jobs (or beginning in steady state), would be

$$[\overline{W} \pm t_{1-\alpha/2,r-1} S_W / \sqrt{r}], \qquad (3)$$

where $\overline{W} = \sum_{i=1}^{r} W_i / r$ is the average over all r (= 10) replications with W_i the average time a job spends in the system for the i^{th} simulation replication,

$$W_i = \frac{1}{d - c} \sum_{j=c+1}^{d} W_{i,j} \text{ , for } i = 1, ..., r$$

where $W_{i,j}$ is the time in the system for the j^{th} job in the i^{th} simulation run (replication), and S_W is the sample standard deviation of the r W_i values, $S_W = \sqrt{\sum_{i=1}^{r} (W_i - \overline{W})^2 / (r-1)}$.

Figure 1 shows the results of one such experiment. Each point, W_i, corresponds to the average time in the system for the last 4000 customers in the i^{th} replication. The vertical line indicates the true value of W based on the underlying exponential distributions. For this example, the confidence interval for E(W) computed from the ten *nr* replications is [3.66, 3.88], which fails to cover the true value. Of course, a 90% confidence interval will fail to cover the true value occasionally: on average, 10% of such intervals will fail to cover the true value. We will see in Section 5, however, that if this experiment were repeated many times, the intervals would fail to cover the true value approximately 79% of the time, not 10%. Making replication runs of the simulation model using the same sample data to set the EDF values (Barton and Schruben 1993 and Cheng 1994) risks significantly underestimating the variability in the response due to the finite sample size.

Plotted on the same figure are two other sets of data. The set labeled *tm* is the set of run averages that would be obtained if the *t*rue *m*odels (exponential with parameters 1.0 and 0.7) had been substituted for the empirical distributions. The confidence interval computed from this set of W_i values is [2.89, 3.06] which does cover the true value of E(W), and if the experiment were repeated many times, only 10% of these confidence intervals would fail to cover the true value.

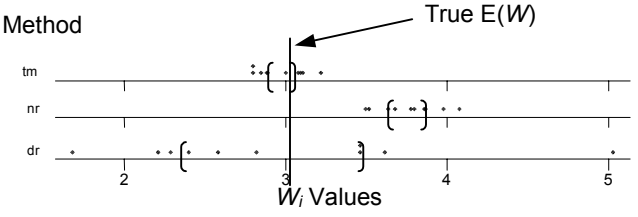

Figure 1: Replication Results and Confidence Intervals for Three Methods

Unfortunately, the simulationist does not know the true probability model, and so one might be tempted restrict any inference about system performance to be *conditional on the chosen input distributions.* It turns out that this restriction is not necessary. The data set labeled *dr* was created by *d*irectly *r*esampling a *new* sample of 100 data points for each of the ten replications, thus requiring 1000 data points in all. The confidence interval computed from this set of W_i values is [2.40, 3.51] which does cover the true value. Without knowledge of the true distributions, this resampling strategy provides much better coverage than the no-resampling strategy. The larger size of the interval reflects the uncertainty cause by constructing input distributions from less-than-infinite sample sizes. Based on the Monte-Carlo studies described in Section 5, approximately 25% of such intervals would fail to cover the true value.

Despite the low coverage and larger interval, this strategy requires ten times as much input data as the *nr* strategy. The bootstrap and uniform randomization resampling methods described in the next section provide approximately correct coverage without requiring additional empirical data. The cost, however, is even larger confidence intervals.

4 RESAMPLING METHODS

The poor coverage of the conventional confidence interval described above is due to neglecting the uncertainty introduced by replacing the true interarrival and service distributions with finite-sample empirical approximations. In this section, we examine three ways to generate input random variables for simulation experiments that attempt to account for the uncertainty in the probability models used for the simulation input. We call these methods *direct (EDF) resampling*, *bootstrap (EDF) resampling*, and *uniformly randomized (EDF) resampling*. In direct resampling, a different real world data set is used to create a new EDF for every run of the simulation. This can be done by partitioning a real world data set into r exclusive subsets and using a different subset of data to create a new set of EDF's for every run of the simulation. This is equivalent to collecting a new set of real system observations for each run. The central idea in the bootstrap and uniform randomization techniques presented in this section is to *simulate* the collection of a new set of actual system data for each run of the simulation.

4.1 Direct Resampling

As shown in Figure 1, confidence intervals for our queueing example that account for input uncertainty can be constructed by using a completely new sample of n empirical interarrival and n empirical service times to create a new EDF for each replications in the experiment. But this requires a factor of r increase in the amount of data available for fitting input models. Alternatively, the set of n data values can be divided into r subsets (assume that n is a multiple of r). One can view this approach as direct resampling with sample sizes of n/r.

Unfortunately, there is little justification for constructing Student t confidence intervals for E(W) for the system based on W_i values from direct resampling replications. Each replication uses different input distributions. If this variability dominates the variability due to the finite length (e.g. 5000 customers) of the simulation, then the Central Limit Theorem need not apply. Further, the average of the W_i values need not tend to E(W). Figure 2 shows a histogram of 500 W_i values for direct resampling with subset sample sizes of 10, for a capacity 10 single server system with service rate 1.0 and arrival rate 0.7 (the parent distributions). The W_i values are right-skewed. The median of the W_i values (3.04) provides a better estimate for E(W) (3.04) than the mean of the W_i values (4.54). This difficulty is less severe for larger subsample sizes and smaller numbers of replications.

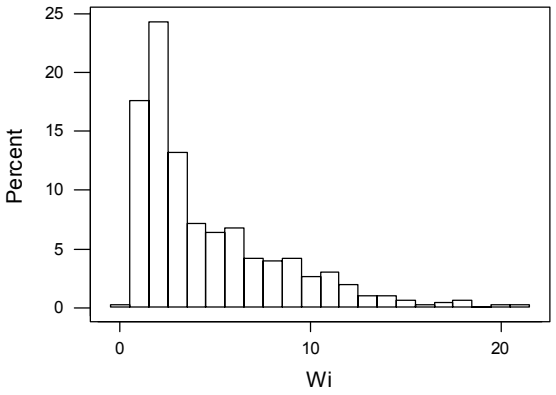

Figure 2: Histogram of 500 W_i Values for Empirical Service and Arrival Distributions Based on 10 Samples Each

4.2 Bootstrap Resampling

One way to capture the uncertainty in the input distributions without resorting to collecting additional (or partitioning) real world samples is to apply the bootstrap technique to create a new EDF for use in each simulation run. The bootstrap technique has been used primarily to characterize the sampling distribution of complex statistics (Efron 1982, Babu and Rao 1993). Before each simulation replication, a new bootstrap sample is drawn by selecting, randomly *with replacement*, n values from the real world data set. This sample is used to construct an EDF to be used in the run. Bootstrap resampling has been applied in simulation settings by a number of authors. See for example Cheng and Jones (2000), and Kleijnen, Cheng and Bettonvil (2000) and the references therein.

Resampling is done for each distribution for each random variable used in the simulation. Since resampling is done with replacement, some of the observations in the original data set may appear more the once or might disappear altogether from the bootstrap sample.

In a simulation experiment, the bootstrap EDF resampling scheme is implemented as follows:

1. Let $i = 1$. For each of the random quantities (indexed by q) modeled with an empirical input distribution, sample n values from the observed data set with replacement. Call the ordered resampled values $v^i_{q(1)}, v^i_{q(2)}, ..., v^i_{q(n)}$.
2. Construct an empirical distribution for each random quantity based on $\{v^i_{q(1)}, v^i_{q(2)}, ..., v^i_{q(n)}\}$. The experiments in this paper used (1-b) and (2-b) to construct the empirical distributions.
3. Conduct a simulation run using these new input distributions.
4. Repeat steps 1 - 3 r times, for $i = 2, 3, ..., r$. Construct confidence intervals using the outputs of the r replications.

Conventional protocol for bootstrapping methodology is to use percentile confidence intervals to emulate the classical t-intervals in (3). See Efron and Tibshirani (1986, 1993) and Babu and Rao (1993) for discussions and justification.

The bootstrap percentile interval calculation in our setting is not strictly valid. Bootstrap methodology assumes that the statistic of interest is computed deterministically from the sample (or resample). That is, if two (re)samples are exactly the same, the computed value of the statistic will be exactly the same. For our simulation examples, the finiteness of the run length adds a source of variability (which in conventional output analysis is the only source of variability). That is, given the same empirical input distributions, two replications will produce different values for W_i (unless common random numbers are used).

In the examples that we consider in this paper, the variability due to the finite empirical samples is significantly larger than the variability induced by the finite simulation run length, and so the assumptions for the bootstrap percentile intervals are not grossly violated. There is a need to develop a more general confidence interval construction method when the variability from the finite run length is large relative to the variability from the finite empirical sample sizes. This issue is examined in the parametric bootstrap setting by Cheng (1994).

4.3 Uniformly Randomized EDF or Bayesian Bootstrap Resampling

Another approach to modeling distribution uncertainty for continuous random variables is by randomizing the EDF in a different way. This approach is based on the well-known fact that if F is the cumulative distribution function of the random variable, X, then $F(X)$ will have the distribution of a Uniform(0, 1) random variable, U. For an ordered real world sample $X_{(1)},X_{(2)},...X_{(n)}$, the joint distribution of $F(X_{(1)}),...,F(X_{(n)})$ corresponds to the n order statistics of from a uniform distribution. The marginal distribution for $F(X_{(m)})$ is beta(m, n-m+1) (Law and Kelton 2000).

Although we do not know F, we can observe an ordered real world sample of horizontal coordinates, $\{X_{(i)}: i = 1, 2, \ldots, n\}$, and generate an ordered sample of vertical coordinates, $\{U_{(i)}: i = 1, 2, \ldots, n\}$, for n random points $(X_{(i)}, F(X_{(i)})$ on F. The uniform variates can be generated by sorting a set of n uniform variates, or by a direct method that does not require sorting (Schucany 1972, Schmeiser 1978). This assigns random probabilities to each of atoms in the empirical distribution function. A randomized distribution function to drive a simulation run can be obtained by smoothing these points. We do that by using the U_m as the values for $\hat{F}^{(p)}(x_{(m)})$ in (2-b).

A standard procedure for choosing the atoms, $\hat{F}^{(p)}(x_{(m)})$, in an EDF is to use a good estimator for the *expected value* of $F(X_{(m)})$. However, here we are not ultimately interested in obtaining a good estimate of F, but rather in simulating our ignorance about F. The randomized EDF sampling method used here is an application of the Bayesian bootstrap approach (Rubin 1981).

In a simulation experiment, randomized EDF sampling is implemented as follows:

1. Let $i = 1$. For each of the random quantities (indexed by q) modeled with an empirical input distribution, order the observed data as $\{ x_{q(1)}, x_{q(2)}, ..., x_{q(n)}\}$. For each q, generate a sample of n ordered values $u^i_{q(1)}, u^i_{q(2)}, ..., u^i_{q(n)}$ from a Uniform(0,1) distribution.
2. Set $\hat{F}^{(p)}_q(x_{q(1)}) = u^i_{q(1)}, \ldots, \hat{F}^{(p)}_q(x_{q(n)}) = u^i_{q(n)}$, and construct a smoothed empirical distribution. (Here we used 2-b.)
3. Conduct a simulation run using these new input distributions.
4. Repeat steps 1 - 3 r times, for $i = 2, 3, ..., r$. Construct bootstrap confidence intervals using the outputs of the r replications.

The next section applies these three methods to the construction of confidence intervals for some simple capacitated single-server queues.

5 CONFIDENCE INTERVALS FOR E(W) FOR SOME CAPACITATED QUEUES

We examined the performance of these three resampling methods in computing confidence intervals for E(W) for some simple capacitated single server queues. Four real systems were modeled: two M/M/1/10 systems with traffic intensity either 0.7 or 0.9, and two systems with exponential interarrival rates 0.7 or 0.9, and Uniform (0.5, 1.5) service times. 90% confidence intervals for E(W) were computed using simulation models with empirical interarrival and service time distributions based on three sample sizes: 50 observations each of interarrival and service times, 100 each, and 500 each.

Figure 3 shows the actual coverage of 90% confidence intervals for the M/M/1/10 systems. The notation for each column indicates the sample size and the interarrival rate. For example, the label 100: 0.7 indicates the results when 100 samples were available and the interarrival rate was 0.7. The poor coverage of the conventional no resampling approach is apparent, as was suggested by the discussion accompanying Figure 1. The direct resampling method intervals have low coverage for sample sizes of 50 (split into ten groups of five) and 100 (split into ten groups of ten), but are still much better than the conventional analysis, even when the individual replications are based on empirical distributions using only five samples each. Both bootstrap methods perform well, the uniformly randomized

EDF performing slightly better than the conventional bootstrap in terms of coverage.

Figure 4 shows similar results for the two systems with uniformly distributed service times. Presumably these systems should exhibit less variation in time-in-system, and so the expected time in the system should be easier to estimate.

The figure shows coverages are better generally. The bootstrap and uniformly randomized EDF methods exhibit comparable coverage for all three sample sizes and for both interarrival rates. Coverage for the direct resampling approach is inferior to the bootstrap coverages for the two smaller sample sizes. The figure also shows that the conventional approach still produces confidence intervals with coverage that is much less than 90%.

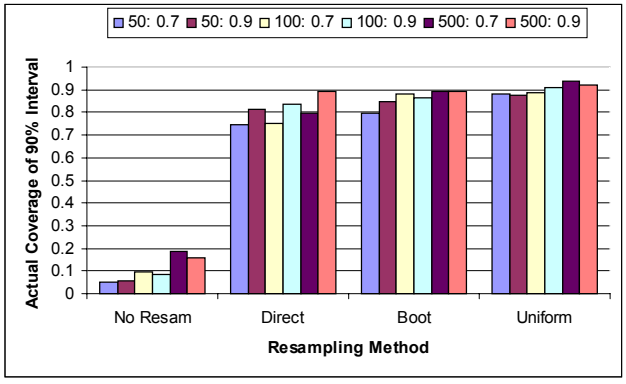

Figure 3: Coverage for Two Exponential/Exponential Examples

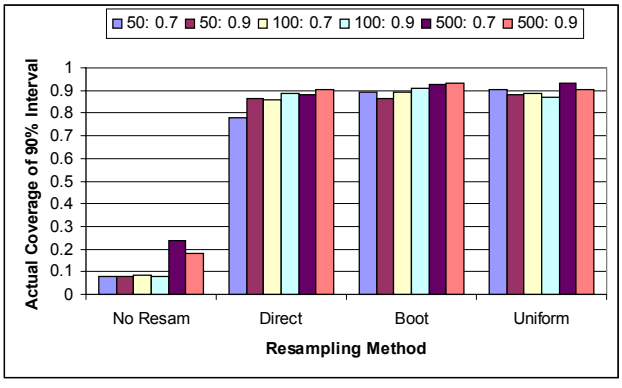

Figure 4: Coverage for Two Exponential/Uniform Examples

6 SUMMARY

The example warns us that current simulation practice of not resampling the EDF can be very risky. In other investigations, we have found that parametric distributions selected using distribution-fitting software also performed poorly when the finiteness of the empirical sample was not taken into account. This is not just a problem with using simulation. Not knowing the exact values of input parameters can cause serious errors for some analytical queueing models as well. Simulation textbooks and much of the research literature treats input distribution estimation and output analysis as distinct topics. When a great deal of simulation data is generated, and input uncertainty is ignored, the consequences can be disastrous.

In light of these findings, what should a practitioner do? Direct, bootstrap, or uniformly randomized EDF resampling methods can be applied in any simulation of a real system where data can be collected. Of course, it is not sensible to generalize to all situations based on our limited analysis of some simple systems. However, it does appear that if real world data collection is expensive it may be good practice to split any large sample of real system data into 10 or so subsamples of data and compute *t*-based confidence intervals using the direct resampling method. This recommendation is consistent with Schmeiser's classic paper that indicates that the marginal value of more than about 20 replications is small (Schmeiser 1982). If sufficient real-system data is not available for at least 10 subsamples, then bootstrap or randomized EDF sampling along with percentile intervals may be able to account for some of the input uncertainty.

ACKNOWLEDGMENTS

Russell Cheng, Steve Chick, Shane Henderson, Jack Kleijnen, Barry Nelson, Bruce Schmeiser, and James Wilson have provided encouragement and insight, and have identified references that helped to improve our understanding of bootstrap methods. This research was supported in part by NSF grants DDM-9308492 and DMI-9322712, by a TQM grant from IBM Corporation, and by research contract 97-FJ-490 with the Semiconductor Research Corporation.

REFERENCES

Babu, G. J., and C. R. Rao. 1993. Bootstrap methodology. In *Handbook of Statistics* 9, 627-659. The Netherlands: Elsevier Science Publishers B.V.

Barton, R. R., and L. W. Schruben. 1993. Uniform and bootstrap resampling of empirical distributions. In *Proceedings of the 1993 Winter Simulation Conference* ed. G.W. Evans, M. Mollaghasemi, W. E. Biles, and E. C. Russell, 503-508. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers.

Cheng, R. C. H. 1994. Selecting input models. In *Proceedings of the 1994 Winter Simulation Conference,* ed. J.D. Tew, S. Manivannan, D.A. Sadowski and A.F. Seila, 184-191. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers.

Cheng, R. C. H., and O. D. Jones. 2000. Analysis of simulation factorial experiments by EDF resample statistics. In *Proceedings of the 2000 Winter Simulation*

Conference, ed. J. A. Joines, R. R. Barton, K. Kang, and P. A. Fishwick, 697-703. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers.

Efron, B. 1982. *The jackknife, the bootstrap and other resampling plans.* Philadelphia: SIAM.

Efron, B., and R. Tibshirani. 1986. Bootstrap methods for standard errors, confidence intervals, and other measures of statistical accuracy. *Statistical Science* 1: 54-77.

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

Kleijnen, J. P. C., R. C. H. Cheng, and B. Bettonvil. 2000. Validation of trace-driven simulation models: more on bootstrap tests. In *Proceedings of the 2000 Winter Simulation Conference*, ed. J. A. Joines, R. R. Barton, K. Kang, and P. A. Fishwick, 882-892. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers.

Law, A. M., and W. D. Kelton. 2000. *Simulation modeling and analysis*, 3rd ed. New York: McGraw-Hill.

Rubin, D. B. 1981. The Bayesian bootstrap. *Annals of Statistics* 9: 130-134.

Schmeiser, B. W. 1978. The generation of order statistics in computer simulation: a survey. In *Proceedings of the 1978 Winter Simulation Conference*, ed. H. J. Highland, N. R. Nielsen, and L. G. Hull, 137-140. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers.

Schmeiser, B. W. 1982. Batch size effects in the analysis of simulation output. *Operations Research* 30: 556-568.

Schucany, W. R. 1972. Order statistics in simulation. *Journal of Statistical Computation and Simulation* 1: 281-286.

Shanker, A., and W. D. Kelton. 1991. Empirical input distributions: an alternative to standard input distributions in simulation modeling, In *Proceedings of the 1991 Winter Simulation Conference*, ed. B. L. Nelson, G. M. Clark, and W. D. Kelton, 975-985. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers.

Wagner, M. A. F., and J. R. Wilson. 1995. Graphical interactive simulation input modeling with bivariate Bézier distributions. *ACM Transactions on Modeling and Computer Simulation* 5: 163-189.

AUTHOR BIOGRAPHIES

RUSSELL R. BARTON is a Professor in the Harold and Inge Marcus Department of Industrial and Manufacturing Engineering at the Pennsylvania State University, University Park. He holds a B.S. degree in Electrical Engineering from Princeton and M.S. and Ph.D. degrees in Operations Research from Cornell University. He worked for three years as a consultant and eight years at RCA Laboratories before entering academia. He is currently the Secretary/Treasurer for the INFORMS College on Simulation. His research interests include applications of statistical and simulation methods to design and manufacturing. His email address is <rbarton@psu.edu>.

LEE W. SCHRUBEN is Professor and Chairman in the Department of Industrial Engineering and Operations Research at the University of California, Berkeley. He holds a B.S. degree in Engineering from Cornell, an M.S. degree in Statistics from the University of North Carolina, and a Ph.D. in Administrative Sciences from Yale. His research is on discrete event simulation modeling and analysis methodologies in a variety of applications including manufacturing, agriculture, health care, banking, and broadcasting. Prior to joining the Berkeley faculty in 1999, Professor Schruben was in the School of Operations Research and Industrial Engineering at Cornell University where he was the Andrew S. Schultz Jr. Professor of Industrial Engineering. He has been the Distinguished Visiting Professor at SEMATECH and the Principal Investigator for Cornell's Center for Semiconductor Manufacturing Research. His email address is <schruben@ieor.berkeley.edu>.