STATISTICAL ANALYSIS OF DISCRETE--EVENT SIMULATIONS

Michael A. Crane[1] and

Donald L. Iglehart[2].

[1]Control Analysis Corporation

[2]Stanford University

ABSTRACT

This paper surveys past and ongoing work by the authors in statistical analysis of simulations. An earlier developed technique for analyzing simulations of GI/G/S queues and Markov chains is shown to apply to discrete-event simulations which can be modeled as regenerative processes. It is possible to address questions of simulation run duration and of starting and stopping simulations because of the existence of a random grouping of observations which produces independent identically distributed blocks in the course of the simulation. This grouping allows one to obtain confidence intervals for a general function of the steady-state distribution of the process being simulated. The technique is illustrated with a simulation of a retail inventory distribution system.

In some simulations, more than one random grouping may exist which produce independent identically distributed blocks. In this case, it is shown that the simulator may choose any such grouping without fear of obtaining less efficient confidence intervals. That is, simulations of the same time duration will produce confidence intervals which are asymptotically the same in length. Finally, approximation techniques are presented which may be appropriate for simulations in which it is not possible to find a random grouping as above.

1. INTRODUCTION AND SUMMARY

In this paper, we continue the development of a technique for analyzing simulations of stochastic systems in the steady state. From the viewpoint of classical statistics, we address the questions of simulation run duration and of starting and stopping simulations. We are able to do so by avoiding two difficulties which have previously made classical statistics inappropriate for simulation analyses. These are the statistical dependence between successive observations and the inability of the simulator to begin the system in the steady state.

In earlier papers, [3], [4], and [5], a random blocking technique was introduced for analyzing the steady-state distributions of GI/G/s queues and of Markov chains in discrete or continuous time. This technique enabled the simulator to observe independent identically distributed (i.i.d.) blocks in the course of the simulation, thus facilitating statistical analysis. Now in this paper we shall show how this method can be extended to cover discrete-event simulations which can be modeled as regenerative processes. A discrete-event simulation is one in which the state of the system being simulated changes at only a discrete, but possibly random, set of time points called event times. Furthermore, with probability one there are only a finite number of these time points in any finite interval. The frequently used simulation languages GPSS, SIMSCRIPT, and TRANSIM are designed specifically to handle this type of simulation.

Suppose our discrete-event simulation can be represented as a stochastic process $\{\underset{\sim}{X}(t) : t \geq 0\}$ with state space $\underset{\sim}{R}^k$, k-dimensional Euclidean space. Our fundamental regenerative assumption is that there exists an increasing sequence of random times $0 \leq \beta_1 < \beta_2 < \cdots$, called regeneration times, such that at each of these times the process $\{\underset{\sim}{X}(t) : t \geq 0\}$ starts from scratch according to the same probabilistic structure governing it at time β_1. Furthermore, we shall assume that the expected time between successive regeneration times is finite. The existence of these regeneration times and certain mild regularity conditions guarantees that a stationary (limiting) distribution exists and also allows us to decompose the simulation into independent identically distributed blocks, thus enabling us to use the method developed in [3], [4], and [5]. This method results in a confidence interval for a general function of the limiting distribution.

The typical situation in which our regenerative assumption is satisfied is when β_i represents the time of the i^{th} entrance to a fixed state $\underset{\sim}{r}$, say, and upon hitting this state the simulation can proceed without any knowledge of its past history. In other words, at the points β_i the process is Markovian. Examples of such regeneration times are the times a queueing system becomes idle or the times a positive recurrent, irreducible Markov chain hits a fixed state; see [3], [4], and [5] for details.

Historically, COX and SMITH [2], p. 136, suggested using the regenerative structure to analyze queueing systems with Poisson arrivals. This idea was partially developed and extended by KABAK [12]. More recently FISHMAN [8], [9], and [10] has also applied these ideas to queueing models, in research concurrent with [3], [4], [5], and [6].

The organization of this paper is as follows. The probabilistic structure of regenerative processes is studied in Section 2. In Section 3, we discuss discrete-event simulations and their relationship to regenerative processes. In Section 4, we show how the probabilistic results may be applied in obtaining confidence intervals for regenerative processes. The methods are illustrated in Section 5 with a simulation of an inventory distribution system. In Section 6, it is shown that the results are invariant to alternative methods of selecting

*This research was sponsored by Office of Naval Research contract N00014-72-C-0086 [NR-047-106].

regeneration times. Finally, in Section 7, approximation techniques are presented which may be useful in cases where the regenerative approach is not directly applicable.

2. REGENERATIVE PROCESSES

The stochastic processes of concern in this paper are regenerative processes. A regenerative process $\{\underset{\sim}{X}(t) : t \geq 0\}$ with state space R^k, k-dimensional Euclidean space, is loosely speaking a stochastic process which starts from scratch at an increasing sequence of regeneration times $\{\beta_i : i \geq 1\}$. That is, between any two consecutive regeneration times β_i and β_{i+1}, say, the portion $\{\underset{\sim}{X}(t) : \beta_i \leq t < \beta_{i+1}\}$ of the regenerative process is an independent, identically distributed replicate of the portion between any other two consecutive regeneration times. However, the portion of the process between time 0 and β_1, while independent of the rest of the process, is allowed to have a different distribution. For complete details on the construction of these processes consult [6] and [13].

The regenerative property is an extremely powerful tool for obtaining analytical results for the process $\{\underset{\sim}{X}(t) : t \geq 0\}$. Before stating these results, we first introduce some notation and make a few assumptions. Let α_i denote the time between the i^{th} and $(i+1)^{th}$ regeneration times, that is, $\alpha_i = \beta_{i+1} - \beta_i$, $i \geq 1$, and assume $E\{\alpha_i\} < \infty$. Let F denote the common distribution function of the α_i's. We shall say that F is arithmetic with span λ if it assigns probability one to a set $\{0, \lambda, 2\lambda, \ldots\}$ for some $\lambda > 0$. For our simulation applications we shall assume that the process $\{\underset{\sim}{X}(t) : t \geq 0\}$ is piece-wise constant, right-continuous, and makes only a finite number of jumps in each finite time interval. Then if F is not arithmetic, it is known that $\underset{\sim}{X}(t) \Rightarrow \underset{\sim}{X}$ as $t \to \infty$; i.e., there exists a random vector $\underset{\sim}{X}$ such that the $\lim_{t \to \infty} P\{\underset{\sim}{X}(t) \leq \underset{\sim}{x}\} = P\{\underset{\sim}{X} \leq \underset{\sim}{x}\}$ for every $\underset{\sim}{x} \in R^k$ at which $P\{\underset{\sim}{X} \leq \underset{\sim}{x}\}$ is continuous. On the other hand, if F is arithmetic with span λ, then there exists a random vector $\underset{\sim}{X}$ such that $\underset{\sim}{X}(n\lambda) \Rightarrow \underset{\sim}{X}$ as $n \to \infty$.

Now let $f : R^k \to R^1$ be a measurable function and define

$$Y_i = \int_{\beta_i}^{\beta_{i+1}} f[\underset{\sim}{X}(s)]ds \quad .$$

The goal of our simulation is to estimate $E\{f(\underset{\sim}{X})\}$, and a confidence interval for this quantity may be obtained through application of the following two propositions. The first follows from the structure of regenerative processes and the second case is proved in [6].

PROPOSITION 1. The sequence $\{(Y_i,\alpha_i) : i \geq 1\}$ consists of independent and identically distributed random vectors.

PROPOSITION 2. If $E\{|f(\underset{\sim}{X})|\} < \infty$, then

$$E\{f(\underset{\sim}{X})\} = E\{Y_1\}/E\{\alpha_1\} \quad .$$

We shall show in Section 4 how these two facts may be used to obtain a confidence interval for $E\{f(\underset{\sim}{X})\}$.

3. APPLICATIONS

Our earlier papers have dealt with two specific examples of regenerative processes, namely general multi-server queues and Markov chains in discrete and continuous time. In the former case, the regeneration times were the times that a customer arrived to an empty system. In the latter case, the regeneration times were the times that the process returned to an arbitrary fixed state. Both of these simulations were examples of discrete-event simulations. In this section, we show how the concept of regeneration times may be applied to more general discrete-event simulations.

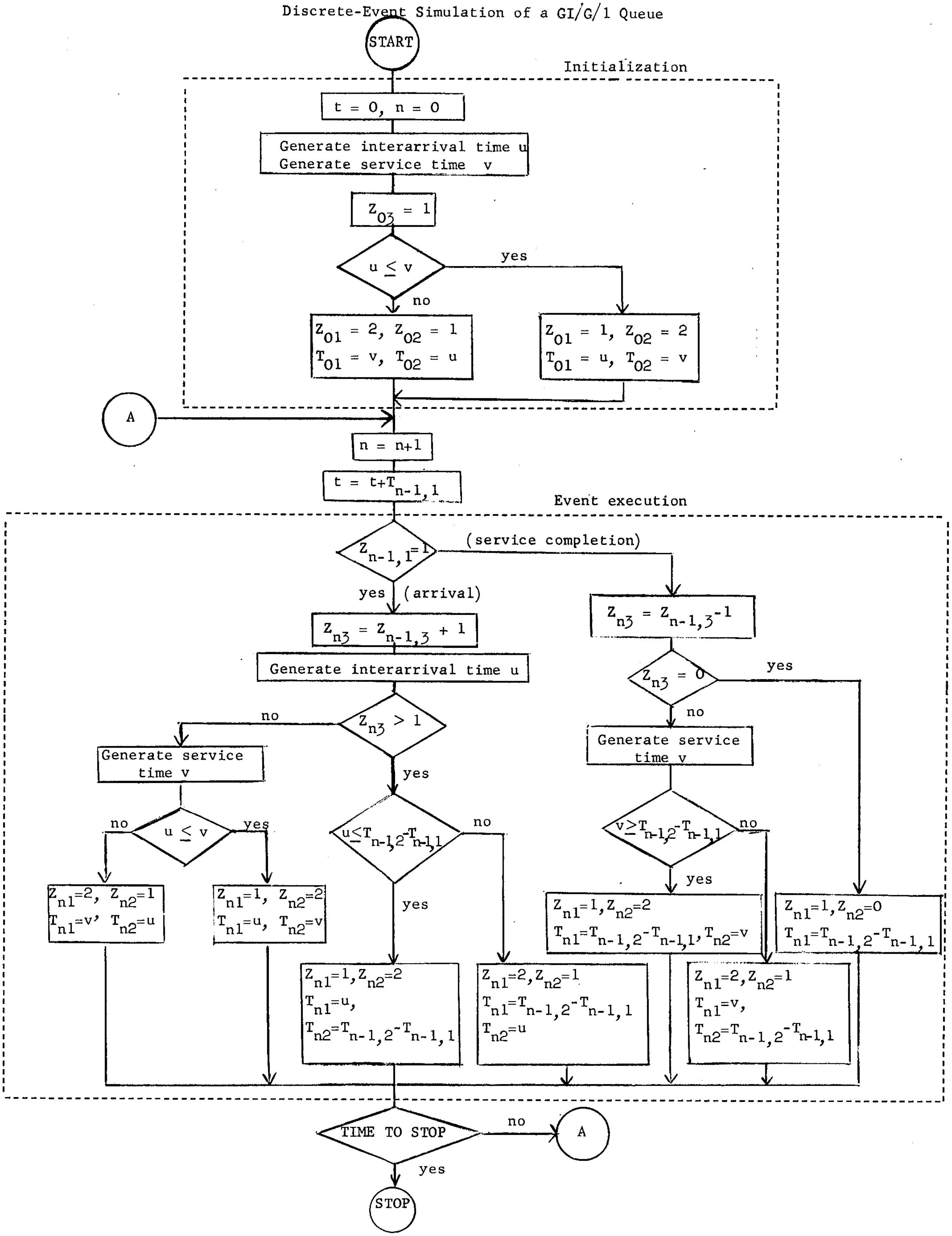
Loosely speaking, a discrete-event simulation is a sample path realization of two vector stochastic processes $\{\underset{\sim}{T}(t) : t \geq 0\}$ and $\{\underset{\sim}{Z}(t) : t \geq 0\}$ which change state at a finite number (in finite time) of event times $0 \leq t_1 \leq t_2 \leq t_3 \leq \cdots$ generated in the course of the simulation. At time t_n, the vector $\underset{\sim}{T}_n = \underset{\sim}{T}(t_n)$ is a chronologically ordered listing of the times to all future events which have been generated through the completion of the n^{th} event, so that the first component of $\underset{\sim}{T}_n$ gives the time duration between the n^{th} and $(n+1)^{th}$ events. (However, the time for the $(n+2)^{th}$ event may not be known until t_{n+1}, for it may in fact be generated at that time.) The vector $\underset{\sim}{Z}_n = \underset{\sim}{Z}(t_n)$ consists of a number of variables describing the simulation model at time t_n, including any variables needed to specify the events to occur at the times given by $\underset{\sim}{T}_n$.

In practice, the processes $\underset{\sim}{T}(t)$ and $\underset{\sim}{Z}(t)$ are generated as follows. At the start of the simulation, $\underset{\sim}{T}(t)$ and $\underset{\sim}{Z}(t)$ are set to initial values $\underset{\sim}{T}_0$ and $\underset{\sim}{Z}_0$ respectively, and the simulation clock is set to $t_0 = 0$. The time of the first event is obtained from the first component of $\underset{\sim}{T}_0$, and the event type is specified by observation of $\underset{\sim}{Z}_0$. The clock is then advanced to time t_1, and the first event is executed. Event execution consists of generating the vectors $\underset{\sim}{T}_1$ and $\underset{\sim}{Z}_1$ based upon observation of $\underset{\sim}{T}_0$, $\underset{\sim}{Z}_0$, and one or more random variables produced at time t_1. The time of the second event is then obtained from $\underset{\sim}{T}_1$ and the process is repeated recursively.

These ideas are best illustrated with a simple example. Figure 1 shows how a simulation of a GI/G/1 queue would be modeled in the above framework. At any given time, there are two possible future events: a customer arrival or a service completion. The first component of $\underset{\sim}{Z}_n = (Z_{n1}, Z_{n2}, Z_{n3})$ describes the first future event, and the second component describes the second future event, if any. These components may take on the values 0, 1, or 2 with 0 denoting no event, 1 denoting

FIGURE 1

Discrete-Event Simulation of a GI/G/1 Queue



an arrival, and 2 denoting a service completion. The third component Z_{n3} denotes the number of customers in the system. Interarrival time and service time random variables are denoted respectively by u and v.

We next discuss the notion of stationarity. Given values for $\tilde{Z}_n$ and $\tilde{T}_n$, we note that $\tilde{Z}_{n+1}$ and $\tilde{T}_{n+1}$ can be determined following the generation of one or more random variables. We may thus speak of state transition probabilities from $(\tilde{Z}_n, \tilde{T}_n)$ to $(\tilde{Z}_{n+1}, \tilde{T}_{n+1})$. The simulation is said to be stationary if these transition probabilities are independent of n. (We note in passing that a periodic simulation could be restructured as a stationary simulation by an appropriate modification of the state space.)

Now consider a stationary simulation and define $\tilde{X}(t) = (\tilde{Z}(t), \tilde{T}(t))$ and $\tilde{X}_n = (\tilde{Z}_n, \tilde{T}_n)$. It follows from the above discussion that $\{\tilde{X}_n : n \geq 0\}$ is a discrete-time Markov chain with general state space and stationary transition probabilities. Furthermore, given $\tilde{X}(t_n) = \tilde{x}$, the time between the n^{th} and $(n+1)^{th}$ transitions of $\tilde{X}(t)$ is known deterministically and depends only on $\tilde{x}$. We thus find a natural relationship between discrete-event simulations and Markov renewal processes with general state space, which we now formalize as follows:

DEFINITION. <u>A stationary discrete-event simulation is a sample path realization of a stationary Markov renewal process $\{\tilde{X}(t) : t \geq 0\}$ with general state space</u>.

The reader will note from our discussion that this definition includes fixed time-increment as well as variable time-increment models.

The application of our methods of Section 2 is now straightforward. If there exists a recurrent state $\tilde{r}$ of the process $\{\tilde{X}(t) : t \geq 0\}$, it may easily be shown that the entry times to $\tilde{r}$ are regeneration times for the process. That is, the process is governed by an identical probabilistic structure following each entrance to $\tilde{r}$. We can thus speak of cycles or tours as being time intervals between successive returns to $\tilde{r}$, and identical random variables defined in successive cycles are then i.i.d. If $\tilde{X}(t) \Rightarrow \tilde{X}$ as $t \to \infty$ and f is a measurable mapping such that $E\{|f(\tilde{X})|\} < \infty$, then we can estimate $E\{f(\tilde{X})\}$ as follows. Let Y_i be the integral of $f[\tilde{X}(t)]$ over the i^{th} cycle and let α_i be the length of the i^{th} cycle. The results of Section 2 imply that the pairs $\{(Y_i, \alpha_i) : i \geq 1\}$ are i.i.d. and that $E\{f(\tilde{X})\} = E\{Y_1\}/E\{\alpha_1\}$. Confidence intervals may thus be obtained by the method reviewed in the following section.

We conclude this section by remarking that the clever simulator may not find it necessary to define a state space as rich as that discussed above in applying our methods to a specific simulation. In many simulations, simpler state space descriptions are sufficient to guarantee a regenerative process. One example is the numerical illustration of Section 5, where we have not found it necessary to include the times to future events in our state space.

4. CONFIDENCE INTERVALS

In this section, we show how Propositions 1 and 2 may be used to obtain confidence intervals for $\nu \equiv E\{f(\tilde{X})\}$. We know from these propositions that $\{(Y_i, \alpha_i) : i \geq 1\}$ are i.i.d. and that $\nu = E\{Y_1\}/E\{\alpha_1\}$. Suppose we observe the n (column) vectors $\tilde{U}_1, \tilde{U}_2, \ldots, \tilde{U}_n$, where $\tilde{U}_i = (Y_i, \alpha_i)$. Let the sample mean of the n observations be denoted by

$$\bar{\tilde{U}} = \begin{pmatrix} \bar{Y} \\ \bar{\alpha} \end{pmatrix} = \frac{1}{n} \sum_{i=1}^{n} \tilde{U}_i$$

and the sample covariance matrix by

$$\tilde{S} = \begin{pmatrix} s_{11} & s_{12} \\ s_{12} & s_{22} \end{pmatrix} = \frac{1}{n-1} \sum_{i=1}^{n} (\tilde{U}_i - \bar{\tilde{U}})\,(\tilde{U}_i - \bar{\tilde{U}})' \quad .$$

Now let $Z_i = Y_i - \nu\alpha_i$, $i = 1, 2, \ldots, n$, and let $\bar{Z} = \frac{1}{n}\sum_{i=1}^{n} Z_i$. Observe that $\{Z_i : i \geq 1\}$ are i.i.d. and that $E\{Z_i\} = 0$, and let $\sigma^2 = E\{Z_i^2\}$. If we let $s^2 = s_{11} - 2\nu s_{12} + \nu^2 s_{22}$, then it may be shown that $s^2 \to \sigma^2$ with probability one as $n \to \infty$. Using this fact together with the continuous mapping theorem and the central limit theorem applied to the Z_i's, and noting that $\bar{Z} = \bar{Y} - \nu\bar{\alpha}$, we see that if $0 < \sigma < \infty$, then

$$\frac{n^{1/2}(\bar{Y} - \nu\bar{\alpha})}{(s_{11} - 2\nu s_{12} + \nu^2 s_{22})^{1/2}} \Rightarrow N(0,1) \qquad (1)$$

as $n \to \infty$, where $N(0,1)$ is a normal random variable with mean zero and variance one. Algebraic manipulation then yields the following approximate $100(1-\gamma)\%$ confidence interval for $\nu = E\{f(\tilde{X})\}$:

$$\frac{[\bar{Y}\bar{\alpha} - ks_{12}] - D^{1/2}}{\bar{\alpha}^2 - ks_{22}} \leq E\{f(\tilde{X})\} \leq \frac{[\bar{Y}\bar{\alpha} - ks_{12}] + D^{1/2}}{\bar{\alpha}^2 - ks_{22}} \qquad (2)$$

where $k = [\Phi^{-1}(1-\gamma/2)]^2/n$, $D = [\bar{Y}\bar{\alpha} - ks_{12}]^2 - [\bar{\alpha}^2 - ks_{22}]\,[\bar{Y}^2 - ks_{11}]$, and Φ is the distribution function for a normal random variable with mean zero and variance one.

We note that the inequalities (2) require that $\bar{\alpha}^2 - ks_{22} > 0$ and $D \geq 0$. In fact, a more precise statement of the result is that (2) holds, together with these conditions, with approximate probability $1-\gamma$ for large n. See [3] for a careful treatment of this point.

The above procedure requires that statistics be gathered over a fixed number of a regeneration times, so that the actual run length of the simulation is not known in advance. Alternatively, it is possible to perform a similar analysis for a simulation of fixed run length T. Let $N(T)$ denote the number of complete regeneration cycles observed by time T; that is,

$$N(T) = n \qquad \text{on } \{\beta_{n+1} \le T < \beta_{n+2}\} \quad .$$

Now redefine $\bar{\underset{\sim}{U}}$ and $\underset{\sim}{S}$ by replacing n with $N(T)$. Based on the central limit theorem for partial sums with a random number of terms, it is easily shown that (1) continues to hold, with n replaced by $N(T)$, as $T \to \infty$. Consequently, for large T, (2) provides an approximate $100(1-\gamma)\%$ confidence interval for $E\{f(\underset{\sim}{X})\}$, again with the substitution of $N(T)$ for n. In other words, the procedure for a fixed run length T is identical to the procedure for a fixed number of regeneration cycles, except that one computes statistics only for those cycles completed by time T.

The above confidence interval has length which is asymptotically proportional to $\Phi^{-1}(1-\gamma/2)/T^{1/2}$ for large T. The simulator might take a small sample to obtain a rough estimate of the proportionality constant. Such an estimate would enable him to determine the ultimate run length and level of confidence, with the appropriate tradeoff between sample cost and precision.

5. AN INVENTORY DISTRIBUTION MODEL

As an illustration of the techniques developed in the preceding sections, we consider the following model of an inventory distribution system for a single item. An inventory warehouse fills orders for two types of customers: retail stores, ordering in large quantities; and direct customers, ordering at most a single unit at a time. Table 1 shows the ordering statistics for each of three retail stores. In addition to the lot size, the table shows p_{sj}, the probability that j weeks transpire between orders of store s. For example, if store 3 orders this week, the probability is 1/4 that it will order again in 3 weeks, 1/2 in 4 weeks, and 1/4 in 5 weeks. Given that it orders, the quantity ordered will be 50.

In addition to the retail stores, there are 160 small customers who order directly from the warehouse. Each of these customers has a probability of 1/4 of ordering in any given week, regardless of the length of time since the last order. The order quantity for each direct customer is one.

The simulation outputs to be observed are: q_{ns}, the number of units ordered by store s in the n^{th} week ($s = 1, 2, 3$); and q_{n4}, the number of orders received from direct customers in the n^{th} week. Let $\underset{\sim}{q}_n = (q_{n1}, q_{n2}, q_{n3}, q_{n4})$ and let β_i denote the i^{th} time after the start of the simulation that $\underset{\sim}{q}_i \in A$, where

$$A = \{\underset{\sim}{x} \in R^4 : x_1 > 0,\ x_2 > 0,\ x_3 > 0\} \quad .$$

Then the times $\{\beta_i : i \ge 1\}$ are regeneration times for the process $\{\underset{\sim}{q}_n : n \ge 0\}$, and may be used to partition the output sequence into i.i.d. blocks as discussed in Section 2. By doing so, we are able to obtain confidence intervals for $E\{f(\underset{\sim}{q})\}$ where $\underset{\sim}{q}$ is the random vector to which $\underset{\sim}{q}_n$ converges in distribution.

The following functions $f : R^4 \to (-\infty, \infty)$, are considered. For $\underset{\sim}{x} \in R^4$, let y = (number of components of x_1, x_2, x_3 which are positive) and define:

$$f_1(\underset{\sim}{x}) = y + x_4$$

$$f_2(\underset{\sim}{x}) = x_1 + x_2 + x_3 + x_4$$

$$f_3(\underset{\sim}{x}) = (y + x_4)^2$$

$$f_4(\underset{\sim}{x}) = (x_1 + x_2 + x_3 + x_4)^2$$

$$f_5(\underset{\sim}{x}) = \begin{cases} 1, & x_1 + x_2 + x_3 + x_4 > 75 \\ 0, & x_1 + x_2 + x_3 + x_4 \le 75 \end{cases}$$

$$f_6(\underset{\sim}{x}) = \begin{cases} 1, & x_1 > 0 \text{ and } x_2 > 0 \\ 0, & \text{otherwise} \end{cases}$$

$$f_7(\underset{\sim}{x}) = \begin{cases} 1, & x_1 = x_2 = x_3 = 0 \\ 0, & \text{otherwise} \end{cases}$$

$$f_8(\underset{\sim}{x}) = x_1 + x_2 + x_3$$

TABLE 1

Retail Store Ordering Statistics

Store s	Lot Size	Ordering Probabilities p_{sj}					
		$j = 1$	$j = 2$	$j = 3$	$j = 4$	$j = 5$	$j = 6$
1	40	0	$\frac{1}{8}$	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{8}$
2	30	0	$\frac{1}{5}$	$\frac{1}{5}$	$\frac{1}{5}$	$\frac{1}{5}$	$\frac{1}{5}$
3	50	0	0	$\frac{1}{4}$	$\frac{1}{2}$	$\frac{1}{4}$	0

TABLE 2

Simulation Results for the Inventory Model* (run length = 20 cycles, level of confidence = 95%)

Parameter	Theoretical value	Point estimate	Confidence interval
$E\{f_1(\underset{\sim}{q})\}$ = E{# of orders}	40.75	40.63	[40.33, 40.94]
$E\{f_2(\underset{\sim}{q})\}$ = E{# of units}	70.00	69.96	[69.37, 70.55]
$E\{f_3(\underset{\sim}{q})\}$ = $E\{(\#\text{ of orders})^2\}$	1691.13	1684.59	[1658.64, 1710.53]
$E\{f_4(\underset{\sim}{q})\}$ = $E\{(\#\text{ of units})^2\}$	5867.50	5854.24	[5713.39, 5995.08]
$E\{f_5(\underset{\sim}{q})\}$ = P{more than 75 units ordered}	.4351	.4475	[.4278, .4673]
$E\{f_6(\underset{\sim}{q})\}$ = P{stores 1 and 2 both order}	.0625	.0690	[.0588, .0792]
$E\{f_7(\underset{\sim}{q})\}$ = P{no stores order}	.4219	.4162	[.3927, .4398]
$E\{f_8(\underset{\sim}{q})\}$ = E{# units ordered by stores}	30.00	30.07	[29.72, 30.43]
$E\{f_9(\underset{\sim}{q})\}$ = E{warehouse operation cost}	300.02	300.11	[299.34, 300.88]

*We are grateful to Mr. Roy Shapiro for his assistance in the simulation programming and the calculation of theoretical values.

TABLE 3

Simulation Results for the Inventory Model (run length = 1280 weeks, level of confidence = 95%)

Parameter	Theoretical value	Point estimate	Confidence Interval
$E\{f_1(\underset{\sim}{q})\}$ = E{# of orders}	40.75	40.70	[40.35, 41.04]
$E\{f_2(\underset{\sim}{q})\}$ = E{# of units}	70.00	70.48	[69.80, 71.15]
$E\{f_3(\underset{\sim}{q})\}$ = $E\{(\#\text{ of orders})^2\}$	1691.13	1685.45	[1658.10, 1712.80]
$E\{f_4(\underset{\sim}{q})\}$ = $E\{(\#\text{ of units})^2\}$	5867.50	5933.27	[5815.21, 6051.32]
$E\{f_5(\underset{\sim}{q})\}$ = P{more than 75 units ordered}	.4351	.4429	[.4207, .4650]
$E\{f_6(\underset{\sim}{q})\}$ = P{stores 1 and 2 both order}	.0625	.0627	[.0480, .0775]
$E\{f_7(\underset{\sim}{q})\}$ = P{no stores order}	.4219	.4132	[.3929, .4336]
$E\{f_8(\underset{\sim}{q})\}$ = E{# units ordered by stores}	30.00	30.54	[29.99, 31.09]
$E\{f_9(\underset{\sim}{q})\}$ = E{warehouse operation cost}	300.02	300.89	[299.93, 301.86]

$$f_9(\underset{\sim}{x}) = 5 + .1(y+x_4) + .1(x_1+x_2+x_3+x_4) + 100(x_1+x_2+x_3+x_4)^{1/4} \quad .$$

These functions are used, respectively, to estimate the expected number of orders in a week, the expected number of units ordered in a week, the second moment of the number of orders, the second moment of the number of units, the probability that more than 75 units are ordered in a week, the probability that both store # 1 and store # 2 place an order in a given week, the probability of no store orders in a week, the expected number of units ordered by stores in a week, and the expected value of a cost function on the warehouse operations. The cost function may be interpreted as follows: a setup cost of 5 per week, a requisition processing cost of .1 per order received, a material handling cost of .1 per unit ordered, and a transportation cost per unit ordered of $100(\#\text{ of units})^{-3/4}$. The transportation cost reflects economies of scale.

Now let $\alpha_i = \beta_{i+1} - \beta_i$, and define

$$Y_i^{(k)} = \sum_{j=\beta_i}^{\beta_{i+1}-1} f_k(\underset{\sim}{q}_j), \quad i \geq 1,\ k = 1,\ldots,9 \ .$$

The sequences $\{(Y_i^{(k)}, \alpha_i) : i \geq 1\}$, $k = 1,\ldots,9$, are i.i.d. and, from Proposition 2,

$$E\{f_k(\underset{\sim}{q})\} = E\{Y_1^{(k)}\}/E\{\alpha_1\} \quad .$$

The statistical methods of Section 4 may thus be used in order to obtain confidence intervals for $E\{f_k(\underset{\sim}{q})\}$.

Tables 2 and 3 show the results of simulation runs for the above model. The level of confidence chosen was 95%. Table 2 shows the results of a run with a fixed number of cycles (20) and Table 3 shows the results with a fixed simulation length (1280 weeks).

6. COMPARING TWO SEQUENCES OF REGENERATION TIMES

In a particular simulation the simulator may have more than one sequence of regeneration times on which to base his confidence interval. Suppose in addition to the sequence $\{\beta_i : i \geq 1\}$ there is a second sequence $\{\beta_i' : i \geq 1\}$ which may be used to decompose $\{\underset{\sim}{X}(t) : t \geq 0\}$ into i.i.d. blocks. This is the case, for example, when $\{\underset{\sim}{X}(t) : t \geq 0\}$ is a positive recurrent, irreducible Markov chain; see [4]. Our goal in this section is to point out that asymptotically as t becomes large the ratio of the length of the confidence interval $I(t)$ based on $\{\beta_i : i \geq 1\}$ to that based on $\{\beta_i' : i \geq 1\}$, call it $I'(t)$, is one with probability one. Thus the simulator can select which ever sequence of regeneration times is most convenient without fear of obtaining an excessively large confidence interval.

Let α_i', Y_i', Z_i' and σ' be defined for the sequence $\{\beta_i' : i \geq 1\}$ as α_i, Y_i, Z_i and σ were defined earlier for $\{\beta_i : i \geq 1\}$. Using a proof similar to that of CHUNG (1960), p. 64, it may be shown that, if $0 < \sigma < \infty$,

$$\frac{\int_0^t f[\underset{\sim}{X}(s)]ds - \nu t}{(\sigma/E^{1/2}\{\alpha_1\})t^{1/2}} \Rightarrow N(0,1)$$

as $t \to \infty$. Similarly, if $0 < \sigma' < \infty$,

$$\frac{\int_0^t f[\underset{\sim}{X}(s)]ds - \nu t}{(\sigma'/E^{1/2}\{\alpha_1'\})t^{1/2}} \Rightarrow N(0,1)$$

as $t \to \infty$. Hence $(\sigma/E^{1/2}\{\alpha_1\}) = (\sigma'/E^{1/2}\{\alpha_1'\})$. Now it may also be shown using the strong law of large numbers and the earlier expression for the confidence interval, that

$$\lim_{t\to\infty} I(t)\sqrt{t} = 2\sigma\Phi^{-1}(1-\gamma/2)/E^{1/2}\{\alpha_1\} \quad \text{with probability one ,}$$

and

$$\lim_{t\to\infty} I'(t)\sqrt{t} = 2\sigma'\Phi^{-1}(1-\gamma/2)/E^{1/2}\{\alpha_1'\} \quad \text{with probability one .}$$

It follows easily that

$$\lim_{t\to\infty} I(t)/I'(t) = \frac{\sigma/E^{1/2}\{\alpha_1\}}{\sigma'/E^{1/2}\{\alpha_1'\}} = 1 \quad \text{with probability one .}$$

Thus we see that with arbitrarily high probability, the lengths of the two confidence intervals will be approximately equal when the length, t, of the simulation run is large. See [11] for more details.

7. APPROXIMATION TECHNIQUES

In order to apply the results of Section 2 in simulating a stochastic model, we must find an increasing sequence of regeneration times $0 \leq \beta_1 < \beta_2 < \cdots$ such that at each of these times the simulation starts from scratch according to the same probabilistic structure governing it at time β_1. Furthermore, it is necessary that the expected time between successive regeneration times be finite and sufficiently small that a number of regeneration times may be observed in the course of the simulation. We have seen in Section 3 that any stationary discrete-event simulation may be modeled as a Markov renewal process $\{\underset{\sim}{X}(t) : t \geq 0\}$. Thus if we let $\beta_i(\underset{\sim}{r})$ denote the i^{th} time that $\underset{\sim}{X}(t)$ enters a state $\underset{\sim}{r}$, then the process will indeed start from scratch at each of the times $\{\beta_i(\underset{\sim}{r}) : i \geq 1\}$. However, it may not be possible to find a state $\underset{\sim}{r}$ such that the expected times between successive returns are finite, or small enough for practical considerations. In this section we present a number of approximation techniques which may be useful in circumventing this problem.

For purposes of illustration, we consider a simulation model of an M/M/1 queue with interarrival times $\{u_n : n \geq 1\}$ and service times $\{v_n : n \geq 0\}$. The process of interest is the sequence of customer waiting times $\{W_n : n \geq 0\}$. As above, let $\beta_i(r)$ denote the ith value of n such that $W_n = r$. It is known that for any fixed $r \geq 0$, the times $\{\beta_i(r) : i \geq 1\}$ are regeneration times for the process. However, since the distributions for u_n and v_n are continuous, the expected times between successive returns to r are finite only in the case $r = 0$. In our earlier paper [3], we used this property of the zero state in order to establish regeneration times appropriate to carry out the statistical analysis. We shall concentrate here, however, on the case where $r > 0$ in order to illustrate techniques which may be used when such a natural "return state" does not exist. In such a case, there are four alternative techniques which may be useful, and these are discussed below. The first three involve applying our methods to a simulation which is modified slightly from the original simulation, and the fourth involves applying approximating methods to the original simulation. Numerical comparisons of these techniques are given in [7].

Partial State-Space Discretization

Recall that the sequence of customer waiting times $\{W_n : n \geq 0\}$ is generated recursively from the interarrival and service times as follows:

$$W_0 = 0$$

$$W_n = \begin{cases} 0 & \text{if } W_{n-1}+v_{n-1}-u_n \leq 0, \\ W_{n-1}+v_{n-1}-u_n & \text{if } W_{n-1}+v_{n-1}-u_n > 0. \end{cases}$$

Suppose now that we consider a modified waiting time process $\{W_n' : n \geq 0\}$ generated as follows:

$$W_0' = 0$$

$$W_n' = \begin{cases} 0 & \text{if } W_{n-1}'+v_{n-1}-u_n \leq 0, \\ W_{n-1}'+v_{n-1}-u_n & \text{if } 0 < W_{n-1}'+v_{n-1}-u_n < r-\epsilon \\ & \text{or } r+\epsilon < W_{n-1}'+v_{n-1}-u_n, \\ r & \text{if } r-\epsilon \leq W_{n-1}'+v_{n-1}-u_n \leq r+\epsilon, \end{cases}$$

where $0 < \epsilon < r$. The modified process differs from the original process only in that each time that the waiting time falls in a "trapping interval" $[r-\epsilon, r+\epsilon]$ around r, the waiting time is set equal to r. This amounts to a discretization of the state space in the neighborhood of r.

Now the entry times $\{\beta_i'(r) : i \geq 1\}$ to r in the modified process are regeneration times for the modified process. Furthermore, since $\epsilon > 0$, the expected time between regeneration times is finite. We may thus apply the methods presented earlier to analyze statistically the modified process. This can then be viewed as an approximate analysis of the original process. The approximation should improve as ϵ gets smaller. At the same time, however, the expected time between regeneration times increases as ϵ decreases.

Stochastic Bounding

One shortcoming of the preceding method is that the simulator is not able to judge the accuracy of the approximation. Thus if confidence intervals are desired for the original process and it is essential that a pre-specified minimum level of confidence be achieved, the technique is not appropriate, since the confidence interval strictly applies only to the approximate process. In this case, an alternative version of the technique may be used for certain processes possessing a monotonicity property.

To illustrate in the case of the M/M/1 queue, we define two modified waiting time processes $\{W_n' : n \geq 0\}$ and $\{W_n'' : n \geq 0\}$:

$$W_0' = 0$$

$$W_n' = \begin{cases} 0 & \text{if } W_{n-1}'+v_{n-1}-u_n \leq 0, \\ W_{n-1}'+v_{n-1}-u_n & \text{if } 0 < W_{n-1}'+v_{n-1}-u_n < r-\epsilon \\ & \text{or } r+\epsilon < W_{n-1}'+v_{n-1}-u_n, \\ r-\epsilon & \text{if } r-\epsilon \leq W_{n-1}'+v_{n-1}-u_n \leq r+\epsilon, \end{cases}$$

and

$$W_0'' = 0$$

$$W_n'' = \begin{cases} 0 & \text{if } W_{n-1}''+v_{n-1}-u_n \leq 0, \\ W_{n-1}''+v_{n-1}-u_n & \text{if } 0 < W_{n-1}''+v_{n-1}-u_n < r-\epsilon \\ & \text{or } r+\epsilon < W_{n-1}''+v_{n-1}-u_n, \\ r+\epsilon & \text{if } r-\epsilon \leq W_{n-1}''+v_{n-1}-u_n \leq r+\epsilon. \end{cases}$$

In both cases, the modified process is again based upon a "trapping interval" about r. However, in the first case, waiting times falling within the trapping interval are set to the lower boundary of the interval, and in the second case, waiting times are set to the upper boundary of the interval.

Now let W, W', and W'' denote the limiting random variables to which the original process and the two modified processes converge in distribution, and let f be a monotonically increasing real-valued function. The queueing model possesses a property of monotonicity which may be stated as follows:

$$E\{f(W')\} \leq E\{f(W)\} \leq E\{f(W'')\} .$$

We may take advantage of this property in order to obtain confidence intervals for $E\{f(W)\}$ with a pre-specified minimum level of confidence.

The entry times $\{\beta_i'(r-\epsilon) : i \geq 1\}$ to $r-\epsilon$ in the first modified process are regeneration times for that process, with finite expected time between regeneration times. We may thus obtain a $100(1-\gamma)\%$ confidence interval for $E\{f(W')\}$ by applying the methods of Sections 2 and 4:

$$L' \leq E\{f(W')\} \leq U' \quad \text{with confidence } 100(1-\gamma)\%.$$

The same method can be used to provide a $100(1-\gamma/2)\%$ lower confidence bound for $E\{f(W')\}$:

$$L' \leq E\{f(W')\} \qquad \text{with confidence } 100(1-\gamma/2)\%.$$

Similarly, we can obtain a $100(1-\gamma/2)\%$ upper confidence bound for $E\{f(W'')\}$ using entry times $\{\beta_i''(r+\epsilon) : i \geq 1\}$ to $r+\epsilon$:

$$E\{f(W'')\} \leq U'' \qquad \text{with confidence } 100(1-\gamma/2)\% .$$

As a consequence of the monotonicity property, these bounds may be combined to obtain the following confidence interval for $E\{f(W)\}$:

$$L' \leq E\{f(W)\} \leq U'' \qquad \text{with at least confidence } 100(1-\gamma)\%.$$

Complete State-Space Discretization

A third alternative is to discretize the basic simulation building blocks with the consequence that the state space is completely discrete. In the case of the queueing simulation, this would involve approximating the distributions for the interarrival and service times by discrete distributions. That is, the interarrival times and service times would be permitted only to take on values $\{\delta, 2\delta, 3\delta, \ldots\}$ for some $\delta > 0$. Consequently, the customer waiting times would be confined to the state space $\{0, \delta, 2\delta, \ldots\}$ and every state would be a recurrent state. Hence, regeneration times could be chosen as the return times to any fixed state, and the statistical methods applied.

Approximate Regeneration Times

Consider now the original waiting time process $\{W_n : n \geq 1\}$ and let $\beta_i(r,\epsilon)$ denote the i^{th} time that the process enters the interval $[r-\epsilon, r+\epsilon]$. The expected time between entries is finite and decreases as ϵ is increased. However, the times $\beta_i(r,\epsilon)$ are not regeneration times for the process. That is, partitioning the process by means of the times $\beta_i(r,\epsilon)$ does not result in independent identically distributed blocks. However, if ϵ is small, this situation will hold in an approximate sense. Random variables observed in successive blocks will have small correlation and will have approximately identical distributions. We are led then to consider a statistical technique which treats the times $\beta_i(r,\epsilon)$ as though they are regeneration times. The simulation may then be analyzed by means of the methods of Section 2 and 4, although the confidence intervals obtained will be only approximate.

BIBLIOGRAPHY

[1] CHUNG, K. L. Markov Chains with Stationary Transition Probabilities. Springer, Berlin, 1960.

[2] COX, D. R. and SMITH, W. L. Queues. Methuen, London, 1961.

[3] CRANE, M. A. and IGLEHART, D. L. Simulating stable stochastic systems, I: General multi-server queues. J. Assoc. Comput. Mach., to appear, 1974.

[4] CRANE, M. A. and IGLEHART, D. L. Simulating stable stochastic systems, II: Markov chains. J. Assoc. Comput. Mach., to appear, 1974.

[5] CRANE, M. A. and IGLEHART, D. L. A new approach to simulating stable stochastic systems. Proceedings of the 1973 Winter Simulation Conference, San Francisco, January 17-19, 1973.

[6] CRANE, M. A. and IGLEHART, D. L. Simulating stable stochastic systems, III: Regenerative processes and discrete-event simulations. Technical Report No. 86-6, Control Analysis Corporation, Palo Alto, California, 1973.

[7] CRANE, M. A. and IGLEHART, D. L. Simulating stable stochastic systems, IV: Approximation techniques. Technical Report No. 86-11, Control Analysis Corporation, Palo Alto, California, 1973.

[8] FISHMAN, G. S. Statistical analysis of multi-server queueing simulations. Technical Report 64, Department of Administrative Sciences, Yale University, 1973.

[9] FISHMAN, G. S. Output analysis for queueing simulations. Technical Report 56, Department of Administrative Sciences, Yale University, 1972a.

[10] FISHMAN, G. S., Estimation in multiserver queueing simulation. Technical Report 58, Department of Administrative Sciences, Yale University, 1972b.

[11] IGLEHART, D. L. Discrete-event simulations with two or more sequences of regeneration times. Technical Report No. 86-9, Control Analysis Corporation, Palo Alto, California 1973.

[12] KABAK, I. W. Stopping rules for queueing simulations. Operations Res., Vol. 16, pp. 431-437, 1968.

[13] MILLER, D. R. Existence of limits in regenerative processes. Ann. Math. Statist., Vol. 43, pp. 1275-1282, 1973.

MICHAEL A. CRANE is Vice President of Control Analysis Corporation, Palo Alto, California. He received a S.B. and S.M. in Electrical Engineering, both from the Massachusetts Institute of Technology. He received a Ph.D. in Operations Research from Stanford University in 1971. He is a member of Sigma Xi, Tau Beta Pi, and Eta Kappa Nu.

DONALD L. IGLEHART is a Professor of Operations Research at Stanford University. Previously he was a member of the Cornell University Department of Operations Research. He holds a B. Engineering Physics from Cornell University and a M.S. and Ph.D. in statistics from Stanford. His current technical interests are in pure and applied probability theory, queueing theory, and the statistical analysis of simulations.