Proceedings of the 1989 Winter Simulation Conference
E.A. MacNair, K.J. Musselman, P. Heidelberger (eds.)

DEVELOPING ANALYTIC MODELS BASED ON SIMULATION RESULTS

A. Alan B. Pritsker
Pritsker Corporation
P.O. Box 2413
West Lafayette, Indiana 47906

ABSTRACT

This paper suggests that a simulation model provides results that lead to the development of analytic models. The analytic models can be used to uncover model relationships. Illustrations are given of simulation models that provide direction for analytic model building.

1. INTRODUCTION

Models of systems should be built for a purpose. If the purpose of modeling is to gain an understanding of system operation then a combined use of simulation and analytic modeling can lead to greater insights and improved system designs.

The paper demonstrates by example how the outputs of simulation models are used to direct the development of analytic models. No claim is made that the observations from the simulation results could not have been made without experimenting with the simulation models.

Models developed for analysis by simulation include the realism necessary to understand the system for which problem solving is sought. The realism included in simulation models provides for greater understanding and a foundation upon which other models can be developed. *Simulation models should be built first and then continually used throughout a problem solving project.*

For the examples included in this paper, each problem is stated with exponentially distributed random variables. This was done to simplify the development of analytic models and to focus attention on the structural characteristics of the simulation model. The assumption of exponential random variables permitted a concentration on the structural aspects of the model. In Section 4, a distribution free analysis is performed to give one example where exponential assumptions are not necessary.

Throughout the examples, only the first iteration of the evolution from simulation model to analytic model is provided. It has been my experience that many iterations take place where transitions are made from a simulation model to an analytic model to a simulation model. The hypothesis that needs further development is that combined analytic-simulation modeling is an appropriate way to obtain a full understanding of system operations.

Further research on the modeling process is required. We have barely begun to scratch the surface of this complex activity. A paradigm needs to be developed that relates the purpose for modeling to the types of results that can be obtained from modeling. In this regard, measures of the benefits and effectiveness of a modeling effort need to be explored. In the following sections, only ideas for developing analytic methods based on simulation results are presented. This is a small step toward the combined modeling activities needed to achieve the benefits of combined analytic and simulation modeling.

2. MODEL OF WORK STATIONS IN SERIES [Pritsker, 1986]

The maintenance facility of a large manufacturer performs two operations. These operations must be performed in series; operation 2 always follows operation 1. The units that are maintained are bulky, and space is available for only eight units including the units being worked on. A proposed design leaves space for two units between the work stations, and space for four units before work station 1. A SLAM II model and its parameters are present in Figure 1. Current company policy is to subcontract the maintenance of a unit if it cannot gain access to the in-house facility.

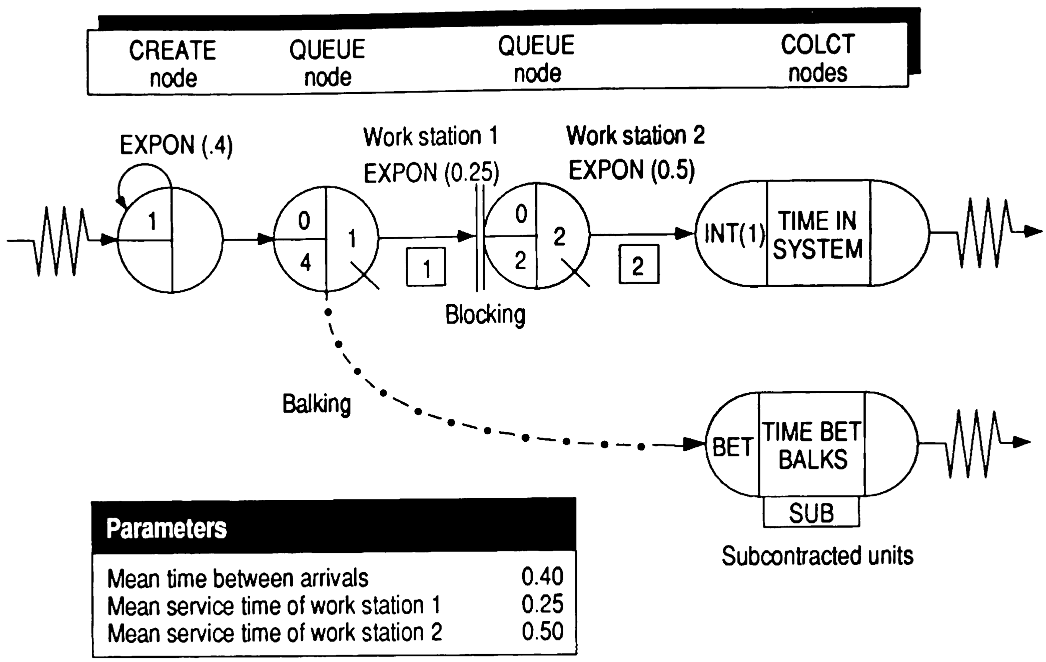

Figure 1 SLAM II network model of maintenance facility.

The serial work station configuration presented in this example results in a significant amount of subcontracting as the mean processing time for work station 2 of 0.5 is greater than the mean time between arrivals which is 0.4. Thus, the utilization factor, ρ, is 1.25. Since there is a finite queue before each work station, steady state results will be obtained, but a significant amount of balking should be expected.

A simulation was run for 300 time units. With a mean time between arrivals of 0.4 time units, we expect 750 arrivals to occur in a 300 time unit simulation. Simulation results [Pritsker, 1986] show that 586 units were processed and 180 units balked. Thus, there were 766 arrivals in the 300 time units. Since the time between arrivals is assumed to be exponentially distributed, the number of arrivals in 300 time units is Poisson distributed with a mean of 750 and a variance of 750. This yields a standard deviation of over 27 and the sample value from the one run is well within one standard deviation of the expected value. For work station 2, there were 586 units processed. From the simulation results, the utilization of work station 2 was 0.9421 indicating that work station 2 worked for 282.63 time units. At the end of the simulation run, work station 2 was processing a unit. Dividing the working time by 587 yields an average service time for work station 2 of 0.481 which is also within one standard deviation of the value of the mean of 0.5 time units that was input.

On first analysis of the maintenance facility, it appears that the designer should have realized that having only two spaces for units before work station 2, and 4 spaces before work station 1, is a poor design since the service time at work station 2 is larger. Based on this, a first alternative would be to put more buffer space before work station 2 in order to decrease the amount of subcontracting. Upon reflection, however, it is the processing time of work station 2 not the amount of buffer assigned to work station 2 that is the problem. To reduce the balking, the processing time at work station 2 needs to be reduced with the addition of another server or more efficient processing techniques. This can be seen by the following simplified analysis of the simulation model.

All randomness was taken out of the model by replacing the distributions by their mean value. Figure 2 shows the status of the two work stations and space 2 of the queue of work station 2. From the status of work station 1, we see that it is either blocked or busy. When exponentially distributed times are used, a similar result is obtained but the length of time work station 1 is blocked or busy is a random variable (see Figure 3).

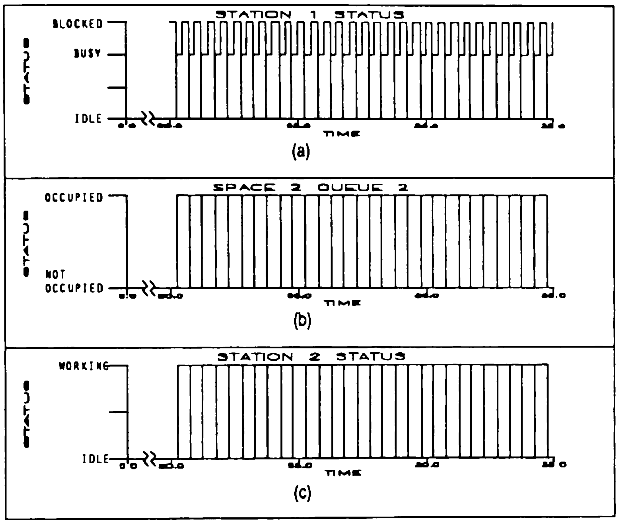

Figure 2 Work stations in series: deterministic simulation.

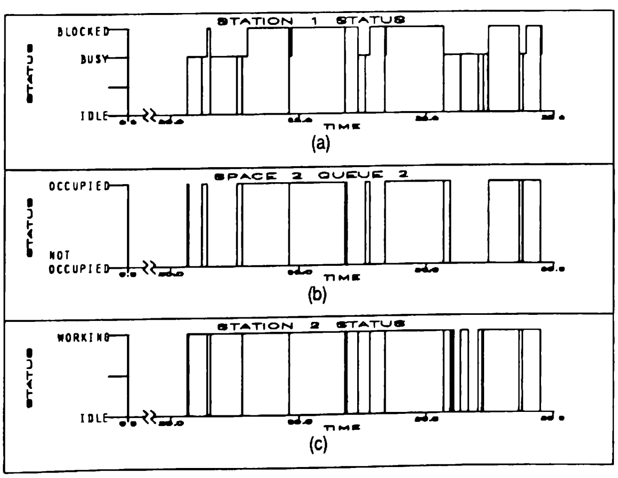

Figure 3 Work stations in series: exponential simulation.

Quantitatively, the average time spent in work station 1 in the stochastic case was 0.4941. Thus, the service time for work station 2 dominates the service time for service station 1. From the above analysis, we can view the model as having a single server with a mean service time of 0.5 and a queue capacity of 7 units (4 for first queue, 1 in the first station, and 2 for second queue). We can now build an analytic model of the situation. From queueing theory [Saaty, 1961], the probability of having 8 units in the system (7 in queue and 1 in service) is given by the following formula:

$$\pi_8 = \frac{(1-\rho)\rho^8}{1-\rho^9}.$$

Since the probability of a unit arriving at any point in time is equal for exponentially distributed arrivals, the probability of a unit balking is the same as the probability of 8 units in the system. Thus, for the analytic model of the abstracted system, the probability of a unit balking is obtained by evaluating the above formula with ρ equal to 1.25 which gives a balking probability of 0.231. Thus, we expect 23% of the arriving units to balk which is approximately the number observed during the simulation run (180/766 = 0.235).

The points of this discussion are:

1. The simulation model with deterministic times provided understanding of the situation;
2. The simulation model was used to communicate system operation;
3. The simulation results pointed to an alternative model;
4. The alternative model was a standard analytic model.

3. DRIVE-IN BANK WITH JOCKEYING [Pritsker, 1986]

A drive-in bank has two windows, each manned by a teller and each with a separate drive-in lane. The drive-lanes are adjacent. From previous observations, it has been determined that the time interval between customer arrivals during rush hours is exponentially distributed with a mean time between arrivals of 0.25 time units (λ = 4). The service time is exponentially distributed for each teller with mean service time of 0.4 (μ = 2.5). It has also been shown that customers have a preference for lane 1 if neither teller is busy or if the waiting lines are equal. At all other times, a customer chooses the shortest line. After entering the system, a customer does not leave until served. However, the last customer in a lane may change lanes if there is a difference of two customers between the two lanes. Because of parking space limitations only 9 cars can wait in each lane. These cars, plus the car of the customer being serviced by each teller, allow a maximum of 20 cars in the system. If the system is full when a customer arrives, the customer balks and is lost to the system. The initial conditions are that both drive-in tellers are busy and there are two customers waiting in each queue.

The drive-in bank example is amenable to modeling with network concepts with the exception of the jockeying of cars between lanes when the lanes differ by two cars or more. Since jockeying can occur only when a teller completes service on a customer, an EVENT node for processing the jockeying of cars between lanes is included in the model following the end-of-service for each entity in the system. The network model for this example is depicted in Figure 4.

In watching an animation of the running of this model, it becomes obvious that the number of cars in teller 1's subsystem differs by at most 1 from the number of cars in teller 2's subsystem. This is also shown in the plot given in Figure 5. Thus, the inclusion of jockeying reduces the number of states required to describe the number of customers in the model.

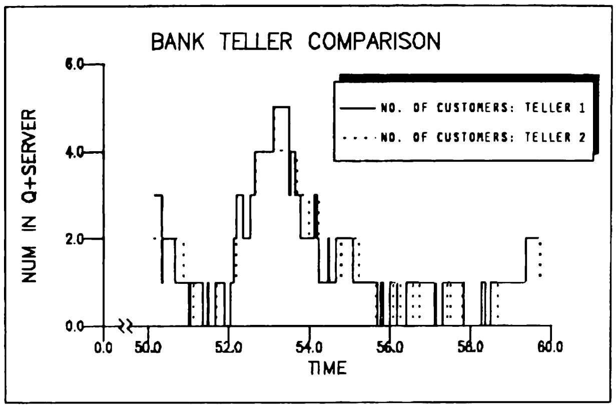

Figure 5 Number of customers for each teller.

With a simplified set of states and all random variables being exponential, a set of differential difference equations can be developed to characterize the probability, $P_{ij}(t)$, of i customers in teller

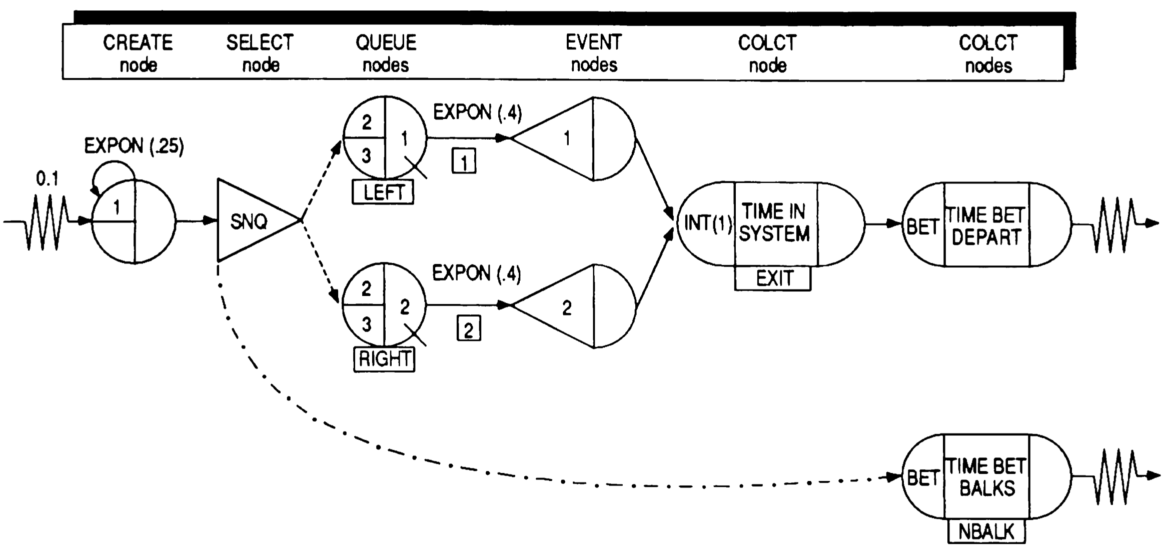

Figure 4 SLAM II network model of drive-in bank.
Source [Pritsker, 1986]

1's subsystem and j customers in teller 2's subsystem at time t. These equations are given below.

$$\frac{dP_{00}(t)}{dt} = -\lambda P_{00}(t) + \mu P_{01}(t) + \mu P_{10}(t)$$

$$\frac{dP_{01}(t)}{dt} = -(\lambda + \mu) P_{01}(t) + \mu P_{11}(t)$$

$$\frac{dP_{10}(t)}{dt} = \lambda P_{00}(t) - (\lambda + \mu) P_{10}(t) + \mu P_{11}(t)$$

For $j = 1, 2, \ldots 9$, we have

$$\frac{dP_{j+1,j}(t)}{dt} = \lambda P_{jj}(t) - (\lambda + 2\mu) P_{j+1,j}(t) + \mu P_{j+1,j+1}(t)$$

$$\frac{dP_{j+1,j}(t)}{dt} = \lambda P_{jj}(t) - (\lambda + 2\mu) P_{j+1,j}(t) + \mu P_{j+1,j+1}(t)$$

$$\frac{dP_{j,j+1}(t)}{dt} = -(\lambda + 2\mu) P_{j,j+1}(t) + \mu P_{j+1,j+1}(t)$$

$$\frac{dP_{jj}(t)}{dt} = -(\lambda + 2\mu) P_{jj}(t) + \lambda P_{j,j-1}(t) + \lambda P_{j-1,j}(t) + 2\mu P_{j+1,j}(t) + 2\mu P_{j,j+1}(t)$$

and

$$P_{10,10}(t) = -2\mu P_{10,10}(t) + \lambda P_{9,10}(t) + \lambda P_{10,9}(t).$$

The continuous capabilities of SLAM II were used within SLAMSYSTEM to evaluate the equations to obtain the plots of $P_{22}(t)$ and $EV(t)$, shown in Figure 6. Note that the initial conditions for the problem have three customers in teller 1's subsystem and three customers in teller 2's subsystem so that $P_{33}(0)=1$ and all other probabilities at time 0 are zero. From these equations, the steady state probabilities and mean number of customers in the drive-in bank teller system were computed as given given in Table 1. Alternatively, the steady state results can be obtained using a Markovian argument [Hazen and Pritsker, 1981 or Glynn, 1984].

The distribution of service times for this example was changed from normal to exponential in

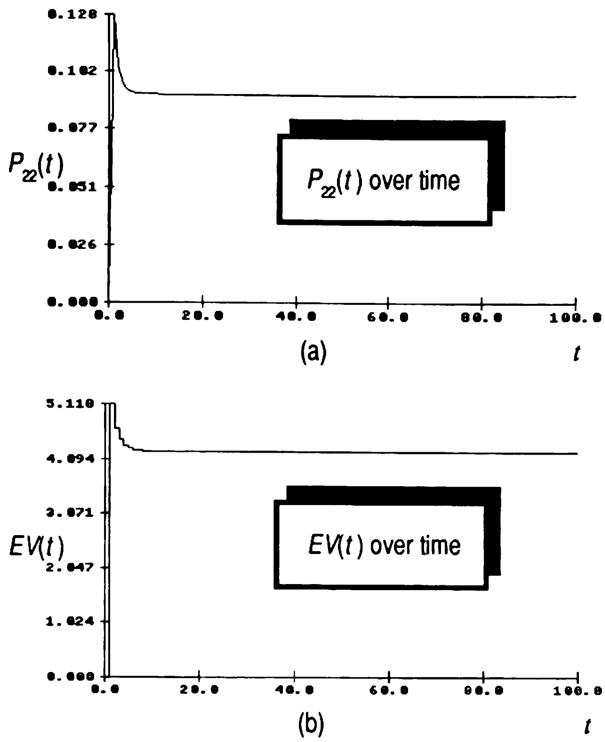

Figure 6 Dynamic behavior of $P_{22}(t)$ and the expected number of customers.

Table 1 Probabilities and Expected Number in System

EXPECTED VALUE IS 4.23161316

NUMBER IN SYSTEM	PROBABILITY
0	0.11226179
1	0.17961882
2	0.14369518
3	0.11495592
4	0.09196500
5	0.07357170
6	0.05885768
7	0.04708581
8	0.03766898
9	0.03013487
10	0.02410818
11	0.01928628
12	0.01542926
13	0.01234321
14	0.00987473
15	0.00789965
16	0.00631983
17	0.00505578
18	0.00404468
19	0.00323571
20	0.00258858

order to do the analytic modeling. A question that needs to be answered is the validity of the analytic models when the form of arrival and service time distributions are changed from normal to exponential. Another question is what would be the impact of changing the random variables to constant values? Further, are there performance measures which are invariant when this certainty equivalence approach is taken and is useful information generated when considering special cases such as $\lambda > 2\mu$ and $\lambda < 2\mu$?

4. CIRCULAR CONVEYOR

Consider the case of five servers stationed along a circular conveyor belt shown in Figure 7.

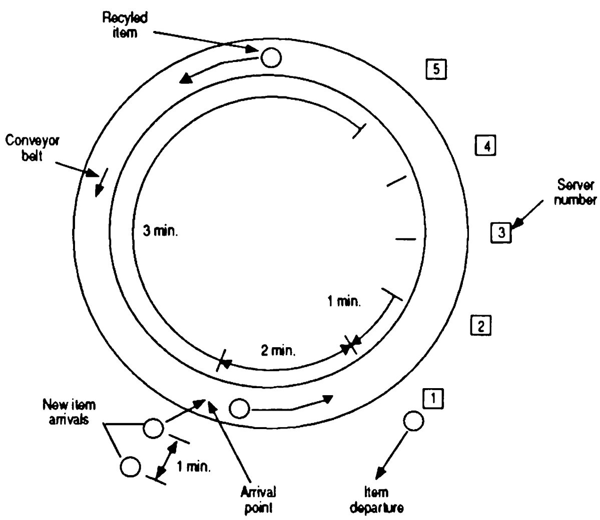

Figure 7 Schematic diagram of a circular conveyor.

Assume that items to be processed by the servers arrive at the conveyor belt with an interarrival time whose mean is 1 minute. After being placed on the conveyor belt, it takes 2 minutes for a new arrival to reach the first server. Service time for each server averages 3 minutes. No storage space for items is provided before any of the servers. If the first server is idle, the item is processed by that server. If the first server is busy when the item arrives, the item continues down the conveyor belt until it arrives at the second server. The delay time between servers is 1 minute. If an item encounters a situation in which all servers are busy, it is recycled to the first server with a time delay of 5 minutes. At the completion of service for an item, the item is removed from the system.

A simulation model was developed for the circular conveyor belt situation. A SLAM II model is shown in Figure 8 with exponentially distributed times. Other distributions for the arrival times and service times were selected and simulation runs performed. From the simulation runs, several observations were made. The length of the delay times on the conveyor belt from the loading point to the first server, between servers and the recirculation delay did not affect the steady state utilization probabilities of the servers. It did, however, affect the amount of time spent in the system by an item. It was also seen that the system was designed to avoid overflows and the recirculation of the items.

Based on this information, an analytic model was built for determining the number of busy servers assuming no recirculation of items and no storage before each server. Under these conditions, the expected number of items being served in an m-server model, L_m, is equal to the sum of the probabilities that a server is busy, that is,

$$L_m = \sum_{j=1}^{m} q_j$$

where q_j denotes the probability that the jth server

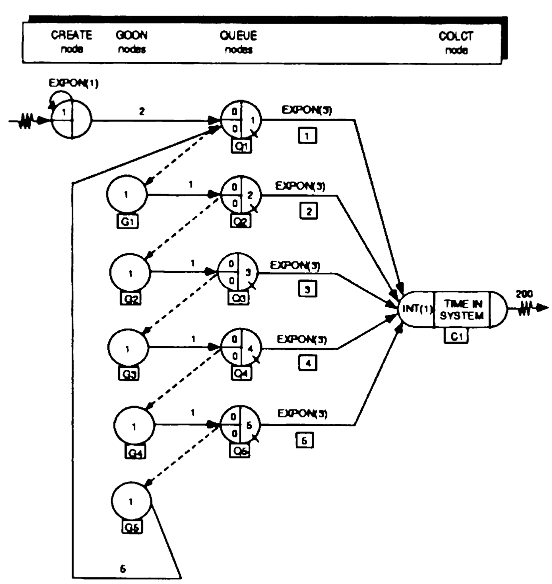

Figure 8 SLAM II network model.

is busy. Under steady-state conditions, the mean duration of the activity of the jth server during an interval of length T is Tq_j and the average number of items served by the jth server during T is μTq_j where μ is the service rate for all servers. Thus, the total number served during T is

$$\sum_{j=1}^{m} \mu T q_j = \mu T \sum_{j=1}^{m} q_j = \mu T L_m$$

The total number of arrivals which are serviced during T is $\lambda T(1 - p_m)$ where λ is the arrival rate, and $(1 - p_m)$ is the steady-state probability of an arrival being served (not lost). Since the total number of arrivals served equals the total number served, we have

$$\lambda T(1 - p_m) = \mu T L_m$$

and $$L_m = \rho(1 - p_m),$$

where $$\rho = \lambda/\mu.$$

In the above development no assumptions regarding the arrival distribution or the service distribution were made and the equation for L_m is a general result. It was assumed that there is no delay between servers. As indicated earlier, simulation results showed that delays between servers do not affect the steady-state probabilities associated with a conveyor system [Pritsker, 1966].

For the ordered selection of the servers, the expected busy time of server j does not depend on the servers after j. Therefore, it is possible to compute the probability of server j being busy starting in the $m = 1$ case and then using all previously computed probabilities to compute the probability that server i, $i > j$ is busy when $m = i$. From

$$L_m = \sum_{j=1}^{m} q_j$$

we have the recurrence relationship

$$q_m = L_m - \sum_{k=1}^{m-1} q_k = L_m - L_{m-1}, \; m = 1, 2, \ldots$$

with $$L_0 = 0.$$

Note that the probability that a channel is busy can be obtained directly from L_m which in turn is dependent only on p_m and ρ. Results from queueing theory for ordered-entry systems can be used to compute p_m.

This example is interesting in that a distribution free analysis is performed. In addition, for those quantities for which distributions were important, it was possible to use known results for deterministic and general distribution types to obtain further insight into the problem. These results were then related to the important variables in the design of the circular conveyor system.

5. CONCLUSIONS

This paper advocates that simulation models be built first to lead the direction of analytic model developments. By example, it demonstrates that there is an evolutionary path from simulation models to analytic models. The combined use of simulation and analytic modeling can lead to a greater understanding of systems.

The paper also calls for a paradigm to relate the purpose for modeling and the results desired from models. Additional research efforts are needed to enhance our understanding of the simulation and analytic problem solving process as is a greater elaboration on the benefits of modeling in problem solving.

ACKNOWLEDGEMENT

This material is based on work supported by the National Science Foundation under Grant No. DMS-8717799. The government has certain rights in this material. I would like to thank Terry Markiewicz and Mary Grant who provided assistance in obtaining the computer outputs for this paper, and Kathy Shoemaker and Miriam Walters for the typing, layout and final preparation of the paper.

REFERENCES

Glynn, P. W., "Some Asymptotic Formulas of Markov Chains with Applications to Simulation," *Journal of Statistical Computation and Simulation*, Vol. 19, 1984, pp. 97-112.

Hazen, G. and A. A. B. Pritsker, "Formulas for the

Variance of the Sample Mean in Finite State Markov Processes," *Journal of Statistical Computation and Simulation*, Vol. 12, 1981, pp. 25-40.
Pritsker, A. A. B., "Application of Multichannel Queueing Results to the Analysis of Conveyor Systems," *Industrial Engineering*, Vol. 17, 1966, pp. 14-21.
Pritsker, A. A. B., *Introduction to Simulation and SLAM II*, 3rd ed., Wiley and Systems Publishing, New York and W. Lafayette, IN, 1986.
Pritsker, A. A. B., C. E. Sigal, and R. D. J. Hammesfahr, *SLAM II Network Models for Decision Support*, Prentice Hall, Inglewood Cliffs, NJ, 1989.
Pritsker, A. A. B., "Why Simulation Works," *Proceedings, Winter Simulation Conference*, 1989.
Saaty, T. L., *Elements of Queueing Theory*, McGraw-Hill, New York, 1961.
SLAMSYSTEM User's Guide, Pritsker Corporation, West Lafayette, IN, 1988.

AUTHOR'S BIOGRAPHY

A. ALAN B. PRITSKER is Chairman of the Board of Pritsker Corporation. He graduated from Columbia University with a BSEE and MSIE. He obtained a Ph.D. from The Ohio State University in 1961. From 1956 through 1962, Dr. Pritsker worked for Battelle Memorial Institute in Columbus, Ohio. From 1962 through 1981, he was Professor of Industrial Engineering at Arizona State University, Virginia Tech, and Purdue University. In 1973, he co-founded Pritsker Corporation.

Dr. Pritsker has written more than 100 technical papers and eight books. He is a member of the National Academy of Engineering. He is a Fellow of AIIE and recipient of AIIE's Distinguished Research Award and Innovative Achievement Award. Dr. Pritsker served as a member of the Board of Directors of the Winter Simulation Conference from 1970 to 1974 and 1980 to 1987, and as WSC Board Chairman in 1984 and 1985.

A. Alan B. Pritsker
Pritsker Corporation
P.O. Box 2413
West Lafayette, IN 47906
(317) 463-5557