Probability in the Engineering and Informational Sciences, **13**, 1999, 329–332. Printed in the U.S.A.

A PSEUDO COUNTEREXAMPLE TO A SIMULATION METATHEOREM*

SHELDON M. ROSS AND KYLE Y. LIN
Department of Industrial Engineering and Operations Research
University of California
Berkeley, California 94720

We state a simulation metatheorem and then present an example that could be considered a counterexample to the metatheorem.

1. THE METATHEOREM AND COUNTEREXAMPLE

Suppose that we want to simulate the value of a random variable X so as to be able to estimate its mean. Furthermore, suppose that, before obtaining the simulated value of X, we obtain the simulated value of a random vector $\mathbf{Y}$.

METATHEOREM: *If $E[X|\mathbf{Y}]$ is easily calculated then a simulation run should be stopped when the value of $\mathbf{Y}$ is obtained.*

For instance, suppose that X is the delay in queue of the nth customer in a FIFO single server queue in which all service times have distribution G with mean μ. Then, if N is the number of customers in queue and R is the remaining service time of the customer being served (or 0 if there is no customer being served) when customer n arrives, then the metatheorem states that, since

$$E[X|N,R] = R + N\mu,$$

we should stop the simulation after gathering the data $\mathbf{Y} = (N,R)$. That is, we should not bother simulating the service times of the N customers in queue. Now, whereas it is certainly true that $E[X|N,R]$ is an unbiased estimator of $E[X]$ and

$$\mathrm{Var}(E[X|N,R]) \leq \mathrm{Var}(X)$$

*This research was supported by the National Science Foundation grant DMI-9610046 with the University of California.

© 1999 Cambridge University Press 0269-9648/99 $12.50

https://doi.org/10.1017/S0269964899133059 Published online by Cambridge University Press

what if we are planning to use additional variance reduction methods to estimate $E[X]$? Can it not be the case that these additional variance reduction techniques will work more efficiently with the raw simulation estimator X than with the conditional expectation estimator?

Example: Consider a single server queueing system in which customers arrive at deterministic times 1,2,3,..., and where the service distribution is exponential with mean 1. Let X denote the delay in queue of customer 5, and let R and N be as before. A simulation of one million runs gives

$$\mathrm{Var}(X) = 2.1389,$$
$$\mathrm{Var}(R+N) = 1.7104.$$

However, if we let S_j denote the service time of customer j and, as suggested in Ross and Lin [2], use each of S_1, S_2, S_3, and S_4 as control variates, then an additional simulation shows that

$$\mathrm{Var}\left(X + \sum_{j=1}^{4} a_j(S_j - 1)\right) = 0.2534$$
$$\mathrm{Var}\left(R + N + \sum_{j=1}^{4} b_j(S_j - 1)\right) = 0.4205,$$

where a_j and b_j are the optimal constants and are chosen by the simulation data (see Ross [1]).

Thus, it appears that there is benefit in simulating the additional service times of those customers waiting in queue when number 5 arrives, which would contradict the metatheorem. The preceding, however, is not the best way to use control variates along with the conditional estimator. Indeed, if we let I_j be the indicator of the event that service on customer j begins before customer 5 arrives, then it follows from the independence of I_j and S_j that, for any constants c_j,

$$E\left[\sum_{j=1}^{4} c_j I_j(S_j - 1)\right] = 0.$$

Thus, we can use an estimator of the type

$$R + N + \sum_{j=1}^{4} c_j I_j(S_j - 1). \tag{1}$$

Note that estimator (1) satisfies the condition of the metatheorem. It turns out that the best estimator of this type has variance given by

$$\mathrm{Var}\left(R + N + \sum_{j=1}^{4} c_j I_j(S_j - 1)\right) = 0.2374.$$

As this variance is smaller than 0.2534, the variance of the control variate estimator based on X, it appears that the metatheorem still stands. However, since the control

https://doi.org/10.1017/S0269964899133059 Published online by Cambridge University Press

variates $S_j, j = 1,2,3,4$, are independent, one needs only to estimate the covariances between X and each S_j to find the optimal values a_j. On the other hand, since the random variables $W_j \equiv I_j(S_j - 1)$, $j = 1,2,3,4$, are dependent, to find the optimal constants c_j it is necessary not only to estimate all of the covariances between $R + N$ and these W_j, but also the covariances between all pairs W_i and W_j; the relatively small variance reduction that is obtained by using estimator (1) over the control variate estimator based on X may not be worth the needed additional computational effort.

A Final Remark: Because the example assumes that the service distribution was exponential, it is not necessary to simulate the value of R. One can stop the simulation when customer 5 arrives and use the (uncontrolled) estimator

$$E[X|N] = N + I,$$

where I is the indicator of the event that there is someone in service when customer 5 arrives. (A simulation indicated that the variance of this estimator is 1.0835.) To improve this estimator, let T_j denote the total amount of time that customer j has already spent in service at the moment that customer 5 arrives, and let

$$T_j^* = \begin{cases} T_j, & \text{if customer } j \text{ leaves before customer 5 arrives} \\ T_j + 1, & \text{if customer } j \text{ is in service when customer 5 arrives} \\ 1, & \text{if customer } j \text{ is waiting in queue when customer 5 arrives} \end{cases}$$

for $j = 1,2,3,4$. That is, if customer j is being served when customer 5 arrives, then T_j^* is equal to the amount of time that j has already spent in service plus 1 (its conditional expected remaining service time). Now, by a well-known argument, it follows that, for any constants d_j, $\{Z_n, 1 \le n \le 4\}$ is a martingale with mean 0 when

$$\begin{aligned} Z_n &\equiv \sum_{j=1}^{n} (d_j T_j^* - d_j E[T_j^* | T_1^*, \ldots, T_{j-1}^*]) \\ &= \sum_{j=1}^{n} d_j (T_j^* - 1), \end{aligned}$$

where the preceding equality follows from the lack of memory property of exponential random variables (which implies that the conditional remaining service time is always equal to the mean service time). Now, if we let M denote the number of customers that have started service by the time customer 5 arrives, then M is easily seen to be a stopping time for this martingale. Hence, by the martingale stopping theorem

$$0 = E[Z_M] = E\left[\sum_{j=1}^{M} d_j (T_j^* - 1)\right].$$

https://doi.org/10.1017/S0269964899133059 Published online by Cambridge University Press

Hence, if we let J_j be the indicator variable for the event that $j \leq M$, then we can utilize an estimator of the type

$$N + I + \sum_{j=1}^{4} d_j J_j (T_j^* - 1).$$

The optimal constants can be estimated from the simulation, and it turns out that the variance of the resulting estimator is 0.1694.

References

1. Ross, S.M. (1997). *Simulation,* 2nd ed. Boston: Academic Press.
2. Ross, S.M. & Lin, K.Y. (in preparation). Applying variance reduction ideas in queueing simulations.

https://doi.org/10.1017/S0269964899133059 Published online by Cambridge University Press