Proceedings of the 2012 Winter Simulation Conference
C. Laroque, J. Himmelspach, R. Pasupathy, O. Rose, and A.M. Uhrmacher, eds

HYBRID SIMULATION FOR CONDITIONAL ESTIMATORS OVER AN INFINITE INTERVAL

Chia-Li Wang
National Dong Hwa University
Hualien, Taiwan 974, ROC

Ronald W. Wolff
University of California
Berkeley, CA 94720, USA

ABSTRACT

Conditional simulation is an efficient variance-reduction method in simulation. Recently, it was applied to a few slowly convergent simulation problems that yielded substantial reduction of the variance. In these applications, the conditional expectations are known or can be computed exactly. We investigate situations where this is not the case; conditional expectations are computed by numerical integration, and are not exact. We construct hybrid simulations that incorporate numerical integration into stochastic conditional simulations. Two key concerns of hybrid simulation are: the effect of the approximation error on the estimator and its computational efficiency. More critically, the pursue of a robust and efficient estimator becomes a real challenge when the integrator has a heavy tail over an infinite interval. We shall resolve both concerns theoretically and provide numerical experiments on queueing simulation and ruin probability estimation to show both the efficiency and quality of our approach.

1 INTRODUCTION AND SUMMARY

Suppose we wish to estimate $\alpha = E(X)$, the expected value of some random variable X, by simulation. The direct way to do this is to generate independent and identically distributed (i.i.d.) sample values $X_1, \cdots, X_n$ of X, and estimate α as their arithmetic mean, $\sum X_i/n$, with variance $\mathrm{Var}(X)/n$.

Now suppose random variables X and Y have a joint distribution with known conditional expectation $E(X|Y)$. Of course, $\alpha = E[E(X|Y)]$. Instead of generating the X_i, we generate corresponding i.i.d. sample values Y_i of Y. What we call a *conditional estimator* of α is the arithmetic average $\sum E(X|Y_i)/n$ with variance $\mathrm{Var}[E(X|Y)]/n$.

From the conditional variance formula, $\mathrm{Var}(X) \geqq \mathrm{Var}[E(X|Y)]$. This inequality is strict except when X is determined by Y. Compared with the direct estimator, the conditional variance is smaller by factor $\mathrm{Var}(X)/\mathrm{Var}[E(X|Y)]$, which we call the statistical efficiency of the conditional estimator.

Thus conditioning in this manner always reduces variance, but is not necessarily a good idea. Generating the Y_i may be more difficult than generating the X_i; determining the ``known'' $E(X|Y_i)$ may require computation that substantially increases simulation run time. In determining whether to use a conditional estimator, these factors must be evaluated.

Recently, conditional simulation was applied to a few slowly convergent simulation problems that yielded substantial reduction of the variance. In these applications, the conditional expectations are known or can be computed exactly.

978-1-4673-4781-5/12/$31.00 ©2012 IEEE

However, it is often the case that E($X|Y$) is not known exactly and must be approximated. This may impair the versatility of conditional estimators. We are not aware of research results in the literature that address this issue. In this paper, we show how to construct and evaluate what we call hybrid simulations that incorporate numerical integration into stochastic simulations.

A hybrid simulation for estimating E(X) works as follows: Generate i.i.d. samples $\{Y_i\}$ of Y, and by numerical integration, approximate E($X|Y_i$) by $T(Y_i)$ for each i. The hybrid estimator is the arithmetic average of the $T(Y_i)$ values.

For the numerical approximation, we use a relative error criterion to bound the error. However, we do not require that every approximation meets the criterion, but only that their average does.

Two key concerns of hybrid simulation are the effect of the approximation error on the estimator and its computational efficiency. To resolve both concerns theoretically, we first consider numerical integration over a finite interval, and use the composite trapezoid rule because of its simplicity and popularity. The implementation includes the allocations of error allowance and the step sizes for the numerical integration and interpolation approximation. For Monte Carlo and regenerative estimators, we obtain an upper bound on the aggregate error of the estimator, explicit formulae for the optimal steps sizes, and asymptotic variance of each estimator and show its convergence toward a robust estimator.

We then consider when conditional expectation is an improper integral for which numerical integration does not directly apply. More critically, heavy-tailed distributions present a real challenge. We discuss two common ways that resolve difficulties: truncation of the integration interval and change of integration variable. The efficiency of the former approach depends on the truncation point. We find the optimal allocation of the error allowance to the numerical integration and the truncation, explicitly and respectively, for Weibull and regularly-varying tails. We compare efficiency between these approaches. We also explore the advantage of using non-uniform step sizes for estimating rare events. Resulting estimators have compelling computational and statistical advantages.

Applying our method to two conditional estimators that have high statistical efficiency: the simulation of single-server queues and the simulation of rare events, we show that the original (pure simulation) statistical efficiency is kept and the computational efficiency is about 70% for a moderate size of samples. The overall performance is better than expected.

REFERENCES

Asmussen, S. & Kroese, D. P., Improved Algorithms for Rare Event Simulation with Heavy Tails, Adv. Appl. Prob., 38, 545--558, 2006.

Wang, C.-L. & Wolff, R. W., Efficient simulation of queues in heavy traffic. ACM TOMACS, 13, 62--81, 2003.

Wang, C.-L. & Wolff, R. W., New Estimators for Efficient GI/G/1 Simulation, Prob. in the Eng. and Inf. Sci., 19, 219 --239, 2005.