DRAFT

Report #06-001

A New Simulation Approach to Estimating Expected Values of Functions of Bernoulli Random Variables under Certain Types of Dependencies

Ross, S.M.

CREATE REPORT
Under Office of Naval Research
Grant No. N00014-05-1-0630
August 2, 2006

HOMELAND SECURITY CENTER

Center for Risk and Economic Analysis of Terrorism Events
University of Southern California
Los Angeles, CA

This research was supported by the United States Department of Homeland Security through the Center for Risk and Economic Analysis of Terrorism Events (CREATE) under grant number N00014-05-1-0630. However, any opinions, findings, and conclusions or recommendations in this document are those of the authors and do not necessarily reflect views of the United States Department of Homeland Security.

A New Simulation Approach to Estimating Expected Values of Functions of Bernoulli Random Variables under Certain Types of Dependencies

Sheldon M. Ross
Department of Industrial and Systems Engineering
University of Southern California
Los Angeles, CA 90089
smross@usc.edu

Abstract

Consider an n component system, where each component either works or is failed. With X_i equal to 1 if component i works and 0 otherwise, we present a new simulation approach, based on an innovative use of stratified sampling, for estimating $E[h(X_1, \ldots, X_n)]$, when h is a monotone function and the vector $X_1, \ldots, X_n$ is exchangeable. We then show how to extend our approach to the case where there is a random environmental parameter Θ such that, conditional on $\Theta = \theta$, the components act independently, with component i working with probability θp_i.

Acknowledgement. This research was supported by the United States Department of Homeland Security through the Center for Risk and Economic Analysis of Terrorism Events (CREATE), grant number EMW-2004-GR-0112. However, any opinions, findings, and conclusions or recommendations in this document are those of the author and do not necessarily reflect views of the U.S. Department of Homeland Security.

1 Introduction

Consider an n component system, where each component either works or is failed. Let X_i equal 1 if component i works and 0 otherwise. For a nondecreasing function $h(x_1, \ldots, x_n)$, defined on n-dimensional binary vectors, we are interested in using simulation to efficiently estimate $\alpha = E[H]$, where $H = h(X_1, \ldots, X_n)$.

In the special case where h is a binary function, we can interpret it as a structure function for a component system (see [1]) and so $E[H]$ would be the probability that the system functions. When the X_i are independent (which we are not assuming), $E[H]$ is the reliability function.

In Section 2 we present a new approach, based on an improved use of standard stratified sampling, for using simulation to estimate $E[H]$ when the vector $X_1, \ldots, X_n$ is exchangeable. In Section 3 we consider the case where there is a random environmental parameter Θ such that, conditional on $\Theta = \theta$, the X_i are independent with $E[X_i] = \theta p_i$. In subsection 3.1 we consider the case where the X_i are independent, and show how to use this to improve the method of Section 3. In Section 4 we present other simulation approaches for estimating system reliability that have appeared in the literature.

2 The Exchangeable Bernoulli Case

Suppose that $X_1, \ldots, X_n$ is an exchangeable random vector. Let $S = \sum_{i=1}^{n} X_i$, and further suppose that the probability mass function

$$P_j = P\{S = j\}, \quad j = 1, \ldots, n$$

is computable.

Before continuing let us mention some examples in which the preceding would hold.

Example

1. The simplest model resulting in the X_i being exchangeable and the distribution of S computable is when the components are independent, with each component working with probability p. In this case,

$$P_j = \binom{n}{j} p^j (1-p)^{n-j}$$

2. A generalization of the preceding case, is when the components are conditionally independent and identically distributed (iid). That is, there is an

environmental parameter Θ, having distribution G, such that if $\Theta = \theta$ then each component independently works with probability $p(\theta)$. In this case

$$P_j = \binom{n}{j} \int p(\theta)^j (1 - p(\theta))^{n-j}\, dG(\theta)$$

The preceding probabilities, being one dimensional integrals, are easily numerically computed.

3. Another possibility is the cascading failure model, introduced and studied in [8] and [2]. In this model, each of n components has an initial load, with these loads being the values of n independent uniform $(0, 1)$ random variables. Following a disturbance, an additional load d is added to each of these n loads. If the resulting load of a component exceeds 1, that component fails. Each failure then adds a new load of fixed amount to those still unfailed components, possibly causing some of them to fail (if their total load now exceeds 1) and thus adding new loads to the so-far unfailed components, and so on. The probability mass function of the total number of components that eventually fail is determined in [8] and [2].

To use simulation to estimate $\alpha = E[H] = E[h(X_1, \ldots, X_n)$, note that

$$\alpha = \sum_{j=0}^{n} E[H|S = j]P_j$$

Our method will, on each simulation run, estimate all of the quantities $E[H|S = j]$, $j = 0, \ldots, n$. This is done by letting each simulation run be as follows.

1. Generate $R_1, \ldots, R_n$, a random permutation of $1, 2, \ldots, n$.
2. Set $j = 0, \quad x_i = 0, i = 1, \ldots, n$
3. Let $H_j = h(x_1, \ldots, x_n)$
4. Let $j = j + 1$, $x_{R_j} = 1$
5. If $j < n$, return to Step 3
6. $H_n = h(1, \ldots, 1)$
7. Return the estimate $\hat{\alpha} = \sum_{j=0}^{n} H_j P_j$

By the exchangeable assumption, given that $S = j$ each of the $\binom{n}{j}$ sets of j components is equally likely to be the set of ones that work. Hence, because $\{R_1, \ldots, R_j\}$ is equally likely to be any set of j components, it follows that H_j is an unbiased estimator of $E[H|S = j]$, implying that $\hat{\alpha}$ is an unbiased estimator of $E[H]$.

We now show that $\hat{\alpha}$ has a smaller variance than does the raw simulation estimator $h(X_1, \ldots, X_n)$.

Proposition $\mathrm{Var}(\hat{\alpha}) \leq \mathrm{Var}(h(X_1, \ldots, X_n))$

Proof We prove this result by showing that $\hat{\alpha}$ can be written as a conditional expectation of the raw simulation estimator. To show this, suppose that $X_1, \ldots, X_n$ is generated by first generating the value of S, then generating a random permutation $\mathbf{R} = (R_1, \ldots, R_n)$, and setting

$$X_i = \begin{cases} 1, & \text{if } i \in \{R_1, \ldots, R_S\} \\ 0, & \text{if otherwise} \end{cases}$$

The result follows because $\hat{\alpha} = E[h(X_1, \ldots, X_n)|\mathbf{R}]$. ■

In the special case where h is a binary nondecreasing function, (that is, when h is a structure function), the estimator has a particularly appealing form. With

$$\begin{aligned} N &= \min\{k : h(x_1, \ldots, x_n) = 1 \text{ when } x_{R_1} = \cdots x_{R_k} = 1,\ x_j = 0,\ j \neq R_1, \ldots, R_k\} \\ &= \min\{k : (R_1, \ldots, R_k) \text{ is a minimal path set}\} \end{aligned}$$

the estimator can be written

$$\hat{\alpha} = \sum_{j \geq N} P_j$$

Remarks We expect our simulation approach to be extremely efficient for the following reasons.

1. When h is a binary structure function, the approach only requires enough knowledge of h to be able to determine when the set of working components ensures that the system is working. (That is, we need only be able to identify when the working components constitute a minimal path set, as opposed to needing a complete list of all such sets.)

2. Each simulation run should be quite fast. The values of P_j can be quickly computed before the simulation begins.

3. In the spirit of antithetic variables, the random permutation $R_1, \ldots, R_n$ can be used twice. First as given, and secondly in its reverse order $R_n, \ldots, R_1$. (For some sufficient conditions as to when using a random permutation both in its forward and reverse direction is better than using two independent random permutations, see [14].)

4. Most important is that because our approach simultaneously estimates all the conditional (stratified) expectations $E[H|S = j]$ in a single run, we expect its variance to be quite small (though because the estimators of $E[H|S = j]$ are positively correlated, the variance is not as small as it would be if these estimators were independent.) ∎

To illustrate our last remark, consider the following example:

Example 1: The bridge system is a five component binary system with minimal path sets $\{1,4\}, \{2,5\}, \{1,3,5\}, \{2,3,4\}$. It is easy to check that

$$P\{N = 2\} = 1/5,\ P\{N = 3\} = 3/5,\ P\{N = 4\} = 1/5$$

Thus, the simulation estimator is

$$\text{Est} = \begin{cases} \sum_{j=2}^{5} P_j, & \text{with prob. } 1/5 \\ \sum_{j=3}^{5} P_j, & \text{with prob. } 3/5 \\ \sum_{j=4}^{5} P_j, & \text{with prob. } 1/5 \end{cases}$$

Now suppose that conditional on $\Theta = \theta$, the components are independent and each works with probability $p(\theta) = \theta$, where θ has distribution G. The following table gives the variance of the estimator when G is the uniform distribution on $(0, 1)$, and the variances when G is the distribution of a deterministic random variable identically equal to p, for various values of p. The column labelled $\alpha(1-\alpha)$ gives the variance of the raw simulation estimator.

G	α	$\alpha(1-\alpha)$	Var(Est)
uniform(0, 1)	.5	.25	.011
deterministic .5	.5	.25	.039
deterministic .7	.8016	.1590	.0213
deterministic .9	.9785	.0211	.0009
deterministic .95	.9948	.0052	.00008

It is interesting that our estimator's performance is superior when Θ is uniform on $(0, 1)$ than it is when Θ is identically $1/2$, even though there is more randomness in the former case. However, because our approach does not simulate Θ, added randomness on its distribution apparently does not always result in additional variance of the estimator. ■

3 The Conditionally Independent Bernoulli Case having $p_i(\theta) = \theta p_i$

Suppose now that there is a random environmental parameter Θ, taking on values between 0 and 1, and such that if $\Theta = \theta$ then the components act independently, with component i working with probability $p_i(\theta) = \theta p_i$. Let G be the distribution function of Θ.

To simulate the reliability of the system, let $p = \max p_i$, and do the simulation in two stages, where the second stage uses our previous method to estimate the reliability of a system whose component probabilities all equal Θp. In stage 1, generate random numbers $U_1, \ldots, U_n$, and set

$$J_i = \begin{cases} 1, & \text{if } U_i < p_i/p \\ 0, & \text{otherwise} \end{cases}$$

If $J_i = 0$, set $X_i = 0$. Now, consider the system conditional on these values of X_i, under the supposition that all the other components (i.e., those components i for which $J_i = 1$) work with probability Θp, and use our earlier method to estimate the reliability of the resulting system. When most of the p_i are close to p this method should work very well. (One way to think about the preceding two stage procedure is to suppose that in order for components to work they must pass two stages: i passes stage 1 with probability p_i/p and, if successful in stage 1, passes stage 2 with probability Θp.)

3.1 The Independent Case

The special case of independent components results when $P\{\Theta = 1\} = 1$. In this case we can improve on the preceding approach by choosing p to be close to most of the p_i. For instance, p could be the arithmetic mean or the geometric mean of the p_i. Let L be the set of component indices i such that $p_i \leq p$ and let H be the remaining component indices. Again do the simulation in two stages, with the second stage using our earlier method to estimate the reliability of a system whose components all work with probability p. In stage 1, generate random numbers

$U_1, \ldots, U_n$, and set

$$J_i = \begin{cases} 0, & \text{if } i \in L \text{ and } U_i > p_i/p \\ 0, & \text{if } i \in H \text{ and } U_i > (1-p_i)/(1-p) \\ 1, & \text{otherwise} \end{cases}$$

If $J_i = 0$ then set $X_i = 0$ if $i \in L$, or set $X_i = 1$ if $i \in H$. Now, consider the system conditional on these values of X_i, under the supposition that all the other components (i.e., those components i for which $J_i = 1$) have reliability p. Now use the method when all $p_i = p$ to estimate the reliability of the resulting system. When most of the p_i are close to p this method should work very well. (Note that components in L work if they pass two stages: $i \in L$ passes stage 1 with probability p_i/p and, if successful in stage 1, passes stage 2 with probability p. Components in H work if they pass either of two stages: $i \in H$ fails stage 1 with probability $(1-p_i)/(1-p)$ and, if it fails, passes stage 2 with probability p. $J_i = 1$ when component i needs to pass stage 2 to work.)

The following example having independent components indicates the promise of this method.

Example 2: Consider a 100 component system that works if at least m of these components work. Suppose that $p_i = .4 + .002i$, $i = 1, \ldots, 100$. The following table gives, for different values of m, the reliability α, the variance of the raw simulation estimator, and the variance of the estimator based on the preceding approach using $p = .5$.

m	α	$\alpha(1-\alpha)$	Var(Est)
30	0.999987	1.3×10^{-5}	5.15×10^{-10}
35	0.999227	7.72×10^{-4}	9.20×10^{-7}
40	0.98396	.0158	.00017
45	0.8701	.1130	.00412
50	0.5476	.2477	.0140
55	0.1889	.1532	.0067
60	0.0291	.0283	.00045

4 Other Simulation Estimators of System Reliability in the Binary Case

Other approaches for using simulation to estimate system reliability are the random hazard estimator (see [7]), importance sampling (see [6], [11], [3]), quick estimation of rare events (see [9]), an estimator based on a conditional expectation identity (see [12]), and evolution simulation procedures (see [4], [5], [10]), all of which appear to require independent components. The random hazard approach requires finding and simulating the components of a minimal cut set (a nonreducible set of components whose failure ensures that the system is failed), then finding and simulating the components of a minimal cut set for the system conditional on the results of the first simulation, and so on. Importance sampling depends on the choice of sampling probabilities and sometimes has unforeseen perils (see Example 3 in the following). Our proposed simulation procedure when adapted to the case of independent components is related to evolution simulation procedures which, although originally developed to estimate network reliability, can be used in the more general reliability systems considered. They work as follows. Suppose the components are independent, with i working with probability p_i. To estimate α, imagine that each component is initially failed and that i remains failed until the random time T_i, from which time it works forever. Suppose that $T_1, T_2, \ldots, T_n$ are independent exponential random variables with rates λ_i, where $e^{-\lambda_i} = 1 - p_i, i = 1, \ldots, n$, and note that p_i represents the probability that component i is working at time 1. Now simulate the T_i, and let I_j be the index of the j^{th} smallest, $j = 1, \ldots, n$. (So $T_{I_1} < T_{I_2} < \ldots < T_{I_n}$.) With M defined as before as the minimal value of k for which $I_1, \ldots, I_k$ constitutes a minimal path set, the raw simulation estimator of $\alpha = P\{\text{system is working at time 1}\}$ is $I(\{T_{I_M} < 1\})$. An improved estimator is

$$E[I(\{T_{I_M} < 1\})|M, I_1, \ldots, I_M] = P\{T_{I_M} < 1|M, I_1, \ldots, I_M\}$$

Using the fact that the distribution of the minimum of independent exponential random variables is independent of the rank ordering of these exponentials, it was noted in the preceding references that the conditional distribution of T_{I_M} given $M, I_1, \ldots, I_M$ is that of a hypoexponential random variable (that is, a convolution of independent exponential random variables) and its distribution function can be computed (see [13]). However, the computation needed is much more involved than what is needed in our new approach.

Importance sampling is a widely used technique for estimating α in the case of independent components. However, as the following example indicates there is often a hidden danger when using importance sampling as a variance reduction technique.

Example 3. Suppose $n = 20$ components are arranged in a linear order, and suppose that the system is failed if there is a block of 3 consecutive components that are all failed. Suppose further that each component is independently failed with probability $q = .1$. A seemingly reasonable importance sampling approach is to simulate the components with failure probability .5. Suppose this simulation results in the data $\mathbf{y} = (y_1, \ldots, y_{20})$, where $y_i = 1$ if component i is failed, and $y_i = 0$ if component i is working. Then, with $s(\mathbf{y}) = \sum_{i=1}^{20} y_i$, the importance sampling simulation estimator of the probability that the system is failed based on this run is

$$I(\mathbf{y})\frac{(.1)^{s(\mathbf{y})}(.9)^{20-s(\mathbf{y})}}{(.5)^{20}} \tag{1}$$

where $I(\mathbf{y})$ is equal to 1 if the vector $\mathbf{y}$ contains a run of at least 3 consecutive values of 1. Now, consider a vector $\mathbf{y}$ with $s(\mathbf{y}) = 3$ and suppose that the 3 values of 1 occur consecutively. For such a data outcome, the value of this estimator is $(.1)^3(.9)^{17}2^{20} = 174.87$. Since there are 18 such outcomes, the probability one will result is $18(.5)^{20} = .000017$. The contribution of these outcomes to the variance of the estimator, which is roughly their contribution to the expected value of its square, is $(174.87)^2(.000017) \approx .52$. Thus, the variance of the importance sampling estimator will be greater than .5, which is much larger even than the variance of the raw simulation estimator. Also, it is quite possible that such a simulation outcome as considered will never occur (since its probability of occurrence on any one run is .000017), and so the sample variance will be much smaller than the actual variance, resulting in incorrect confidences placed on the interval estimates. (Indeed, if the type of outcomes being considered never occur, then the conditional expected value of the estimate will be less than the probability of system failure by approximately $(174.87)(.000017) \approx .003$.)

References

[1] Barlow, R., and F. Proschan, STATISTICAL THEORY OF RELIABILITY AND LIFE TESTING: PROBABILITY MODELS, Holt-Rhinehart- Winston, (1975)

[2] Dobson, I., Carreras, B.A., and D.E. Newman, "A Loading Dependent Model of Probabilistic Cascading Failure, *Probability in the Engineering and Informational Sciences*, **19**, 15-32, (2005)

[3] El Khadiri, M., and H. Cancela, "An Improvement to the Total Hazard Method for System Reliability Simulation", *Probability in the Engineering and Informational Sciences*, **10**, 2, (1996)

[4] Elperin, T., Gertsbakh, I., and Lomonsov, M., 'Estimation of Network Reliability using Graph Evolution Models", *IEEE Transactions in Reliability*, **40**, 5, 572-581, (1991)

[5] Elperin, T., Gertsbakh, I., and Lomonsov, M., "An Evolution Model for Monte Carlo Estimation of Equilibrium Network Renewal Parameters", *Probability in the Engineering and Informational Sciences*, **6**, 457-469, (1992)

[6] Fishman, G., A Monte Carlo Sampling Plan for Estimating Network Reliability, *Operations Research*, **34**, 581-594, (1986)

[7] Jun, C. H., and S. M. Ross, "System Reliability by Simulation: Random Hazards versus Importance Sampling", *Probability in the Engineering and Informational Sciences*, **6**, 119-126, (1992)

[8] Lefevre, C., "On the Outcome of a Cascading Failure Model, *Probability in the Engineering and Informational Sciences*, **20**, 413-427, (2006)

[9] Lieber, D., Rubinstein, R., and D. Elmakis, "Quick Estimation of Rare Events in Stochastic Networks," *IEEE Transactions on Reliability*, **46**, 2, 254-265, (1996)

[10] Lomonsov, M., "On Monte Carlo Estimates in Network Reliability", *Probability in the Engineering and Informational Sciences*, **8**, 245-264, (1994)

[11] Mazumdar, N., "Importance Sampling in Reliability Estimation", in R. Barlow, J. Fussell, and N. Singpurwalla eds., RELIABILITY AND FAULT TREE ANALYSIS, Philadelphia, SIAM, 153-165, (1975)

[12] Ross, S. M., "A New Simulation Estimator of System Reliability," *Journal of Applied Mathematics and Stochastic Analysis*, **7**, 331-336, (1994)

[13] Ross, S. M., INTRODUCTION TO PROBABILITY MODELS, ninth edition, Academic Press, 2007

[14] Ross, S. M., SIMULATION, fourth ed., Academic Press, 2006

[15] Ross. S.M., "Variance Reduction in Simulation via Random Hazards," *Probability in the Engineering and Informational Sciences*, **4**, 299-310, (1990)