Stochastic Analysis, Stochastic Systems, and Applications to Finance

Allanus Tsoi, David Nualart, George Yin Editors

Finding Expectations of Monotone Functions of Binary Random Variables by Simulation, with Applications to Reliability, Finance, and Round Robin Tournaments

Mark Brown

Department of Mathematics
City College, CUNY, New York, NY
Email: cybergarf@aol.com

Erol A. Peköz

School of Management
Boston University
595 Commonwealth Avenue
Boston, MA 02215
Email: pekoz@bu.edu

Sheldon M. Ross

Department of Industrial and Systems Engineering
University of Southern California
Los Angeles, CA 90089
Email: smross@usc.edu

We study the quantity $E[\phi(X_1, \ldots, X_n)]$ when $X_1, \ldots, X_n$ are independent Bernoulli random variables, and ϕ is a nondecreasing function. With $T = \sum_{i=1}^n X_i$, we note that $E[\phi(X_1, \ldots, X_n)|T]$ is a nondecreasing function of T, and show how it can be efficiently estimated by a simulation study that stratifies on T. Our results are applied to static and dynamic reliability systems, the pricing of derivatives related to basket default swaps, and to round robin tournaments.

1. Introduction

Let $X_1, \ldots, X_n$ be independent Bernoulli random variables, and let $T = \sum_{i=1}^n X_i$ be their sum. Also, suppose that ϕ is a nondecreasing function of the vector $\mathbf{X} = (X_1, \ldots, X_n)$. In Section 2 we show how we can efficiently estimate $E[\phi(\mathbf{X})]$ by a simulation approach that stratifies on T. In Section 3, we note that $E[\phi(\mathbf{X})|T]$ is a nondecreasing function of T, and show how this result can sometimes be used to modify our simulation estimates. In Section 4 we apply our results to the classical reliability problem of finding the probability that a system composed of n independent binary components will function. In Section 5 we consider a dynamic version of the preceding model, in which each component works for a random length of time and then fails. Depending on the structure of the system, the component failures eventually cause the system to fail. Supposing that there is a cost incurred if the system fails before some specified time t^*, with the cost depending on the time of failure, we present a stratification approach to estimate the expected cost incurred. In Section 6 we apply our results to the pricing of derivatives related to basket default swaps, where our methods are similar to, but an improvement on, a recently proposed simulation approach (see[2]). In Section 7 we apply our results to estimating win probabilities in round robin tournaments.

2. Estimating $E[\phi(X_1, \ldots, X_n)]$ by Simulation

Suppose that $\mathbf{X} = (X_1, \ldots, X_n)$ is a vector of independent Bernoulli random variables with $E[X_i] = p_i$, and that we are interested in using simulation to estimate $E[\phi(\mathbf{X})]$. Our approach is to stratify on $T = \sum_{i=1}^{n} X_i$. To do so we need to first determine the probability mass function of T, and then show how to simulate $\mathbf{X}$ conditional on $T = i$, for $i = 0, \ldots, n$.

To start, we note that if

$$P_j(i) = P(X_j + \ldots + X_n = i)$$

then, analogous to Example 3.22 of,[6] $P_j(i)$ satisfy

$$P_j(i) = p_j P_{j+1}(i-1) + (1-p_j) P_{j+1}(i) \tag{1}$$

$$P_n(1) = p_n\,, \quad P_n(0) = 1 - p_n\,, \quad P_n(i) = 0\,,\ i \neq 0, 1$$

Starting with the preceding expression for $P_n(i)$, the equations (1) are easily solved by a recursion that first solves for $P_{n-1,}$ then P_{n-2}, and so on.

To generate $(X_1, \ldots, X_n)$ conditional on $T = i$, generate in sequence

- X_1 given $T = i$
- X_2 given $T = i, X_1$
- X_3 given $T = i, X_1, X_2$

and so on. To generate X_j, given both that $T = i$ and the values of

$X_1, \ldots, X_{j-1}$, use that

$$P(X_j = 1|T = i, X_1, \ldots, X_{j-1}) = P(X_j = 1|\sum_{k=j}^{n} X_k = i - \sum_{k=1}^{j-1} X_k)$$
$$= \frac{p_j P_{j+1}(i - 1 - \sum_{k=1}^{j-1} X_k)}{P_j(i - \sum_{k=1}^{j-1} X_k)} \quad (2)$$

Suppose one is planning on doing r simulation runs. One way to employ the stratification approach is to do $rP_1(i)$ runs conditional on $T = i$ for each $i = 0, \ldots, n$. Call the preceding the "standard proportional stratification" approach, Because T and $\phi(\mathbf{X})$ are both increasing functions of the vector of independent random variables $\mathbf{X}$ it follows that they are positively correlated and thus the standard proportional stratification will result in an estimator with smaller variance than the estimator based on r non-stratified runs (see[7]).

However, a more efficient approach than using standard proportional stratification is to first do a preliminary study by simulating $\phi(\mathbf{X})$ conditional on $T = i$ enough times so as to get a rough estimate of

$$s_i^2 \equiv \text{Var}(\phi(\mathbf{X})|T = i), \quad i = 0, \ldots, n$$

Then, if you are planning to do r simulation runs, do $r\frac{s_i P_1(i)}{\sum_k s_k P_1(k)}$ of these runs conditional on $T = i$. (That this results in the minimal variance of the final estimator see.[7]) Letting ϕ_i be the average of the runs done conditional on $T = i$, estimate $E[\phi(\mathbf{X})] = \sum_{i=0}^{n} E[\phi(\mathbf{X})|T = i]P_1(i)$ by $\sum_{i=0}^{n} \phi_i P_1(i)$.

Remark: Whether using the standard or the more efficient procedure, further variance reduction can be obtained by using antithetic variables. That is, in generating $(X_1, \ldots, X_n)$ conditional on $T = i$, first generate random numbers $U_1, \ldots, U_n$, and then generate the value of X_j, given both that $T = i$ and the values of $X_1, \ldots, X_{j-1}$, by letting it equal 1 if U_j is less than the right side of (2.) In the next generation of $(X_1, \ldots, X_n)$ conditional on $T = i$, we can utilize the same set of random numbers, but this time subtracting each from 1. (That is, we use the random numbers $1-U_1, \ldots, 1-U_n$ in the next run.) Because ϕ is a monotone function the reuse of the random number set will lead to a variance reduction when compared with using a new independent set of random numbers (see[7] for a proof).

Example 1: Let $X_i, i = 1, \ldots, 20$, be independent Bernoulli random variables with common mean $P(X_i = 1) = 1/2$, and let

$$\phi(x_1, \ldots, x_{20}) = (\sum_{i=1}^{20} i x_i)^2$$

The standard deviation of the standard Monte Carlo (commonly referred to as the *raw*) simulation estimator of $E[\phi(X_1, \ldots, X_n)]$ based on 10^4 runs is 57, whereas the standard deviation of the stratified estimator based on the same number of runs is 26, a variance reduction by a factor of approximately 4. However, noting that the lower indexed X_i have a much smaller effect on the value of $\phi(X_1, \ldots, X_n)]$ than do the higher ones, it seems reasonable that rather than stratifying on $\sum_{i=1}^{20} X_i$ it would be better to stratify on, say, $\sum_{i=10}^{20} X_i$. A further simulation indicated that stratifying on $\sum_{i=10}^{20} X_i$ reduced the standard deviation (again based on 10^4 runs) down to 8, a variance reduction over the standard Monte Carlo estimator by a factor of about 50.

3. Monotonicity of the Conditional Expectation given T

Not only, as noted in the preceding section, are T and $\phi(X_1, \ldots, X_n)$ positively correlated but, even stronger, is that the conditional distribution of $\phi(\mathbf{X})$ given that $T = k$ is stochastically increasing in k. That is, we have the following result.

Theorem 3.1. *If $X_1, \ldots, X_n$ are independent Bernoulli random variables, and ϕ a nondecreasing function, then $E[\phi(X_1, \ldots, X_n)|T = k]$ is a nondecreasing function of k.*

Theorem 3.1 is a special case of a general result of[3] which states that the preceding is true whenever the X_i are independent and have logconcave densities or mass functions. As the proof of[3] is rather involved, we now present a proof that is not only elementary but also in the spirit of this paper. We being with a lemma.

Lemma 3.1. *The conditional distribution of X_n given T is stochastically increasing in T. That is, $P(X_n = 1|T = k)$ is a nondecreasing function of k.*

Proof: The proof makes use of the log concavity result that $P(T = k)/P(T = k-1)$ is nonincreasing in k. (For a proof, see.[5]) Let $p_n =$

$P(X_n = 1) = 1 - q_n$. Also, let $b_r = P(\sum_{i=1}^{n-1} X_i = r)$. Then, we have

$$\begin{aligned} P(X_n = 1|T = k) &= \frac{p_n b_{k-1}}{P(T = k)} \\ &= \frac{p_n b_{k-1}}{p_n b_{k-1} + q_n b_k} \\ &= \frac{p_n}{p_n + q_n b_k / b_{k-1}} \end{aligned}$$

and the result follows since b_k/b_{k-1} is nonincreasing by log concavity. ■

We now prove the theorem.

Proof of Theorem 3.1: Let $S_r = \sum_{i=1}^{r} X_i$. Also, for fixed $k > 0$, let $Y_1, ..., Y_n$ be distributed as $X_1, ..., X_n$ conditional on $S_n = k$; and let $Z_1, ..., Z_n$ be distributed as $X_1, ..., X_n$ conditional on $S_n = k - 1$. We now show how we can generate random vectors $Y_1, ..., Y_n$ and $Z_1, ..., Z_n$ that are distributed according to the preceding and are such that $Y_i \geq Z_i, i = 1, \ldots, n$. To do so, first note that by the Lemma 3.1

$$P(Y_n = 1) \geq P(Z_n = 1)$$

We generate the random vectors as follows:

KEY STEP Generate a random number U. Then,

$$\text{if } U \leq P(Z_n = 1), \text{ set } Z_n = 1, \text{ else set it equal to } 0$$

$$\text{if } U \leq P(Y_n = 1), \text{ set } Y_n = 1, \text{ else set it equal to } 0$$

There are now 3 cases:

Case 1: $Z_n = Y_n = 1$

Given the scenario of this case, $Y_1, ..., Y_{n-1}$ is distributed as $X_1, .., , X_{n-1}$ conditional on $S_{n-1} = k-1$; and $Z_1, ..., Z_{n-1}$ is distributed as $X_1, .., , X_{n-1}$ conditional on $S_{n-1} = k - 2$. Consequently, by Lemma 3.1

$$P(Y_{n-1} = 1|Y_n = 1) \geq P(Z_{n-1} = 1|Z_n = 1)$$

Thus, we can repeat KEY STEP to generate Y_{n-1} and Z_{n-1} so that $Y_{n-1} \geq Z_{n-1}$.

Case 2: $Z_n = Y_n = 0$

Given the scenario of this case, $Y_1, ..., Y_{n-1}$ is distributed as $X_1, .., , X_{n-1}$

conditional on $S_{n-1} = k$; and $Z_1, ..., Z_{n-1}$ is distributed as $X_1, .., , X_{n-1}$ conditional on $S_{n-1} = k - 1$. Consequently, by Lemma 3.1

$$P(Y_{n-1} = 1 | Y_n = 0) \geq P(Z_{n-1} = 1 | Z_n = 0)$$

Thus, we can repeat KEY STEP to generate Y_{n-1} and Z_{n-1} so that $Y_{n-1} \geq Z_{n-1}$.

Case 3: $Z_n = 0$, $Y_n = 1$

Given the scenario of this case, the random vectors $Y_1, ..., Y_{n-1}$ and $Z_1, ..., Z_{n-1}$ have the same joint distribution. Thus, we can generate them so that $Y_i = Z_i, i = 1, \ldots, n-1$.

The preceding shows how to generate the vectors so that $Y_i \geq Z_i, i = 1, \ldots, n$. By the monotonicity of ϕ this implies that $\phi(Y_1, \ldots, Y_n) \geq \phi(Z_1, \ldots, Z_n)$. Consequently,

$$\begin{aligned} E[\phi(X_1, \ldots, X_n) | T = k] &= E[\phi(Y_1, \ldots, Y_n)] \\ &\geq E[\phi(Z_1, \ldots, Z_n)] \\ &= E[\phi(X_1, \ldots, X_n) | T = k-1] \quad \blacksquare \end{aligned}$$

Remark: While Theorem 1 might seem quite intuitive, it does depend on the X_i being Bernoulli random variables. For instance, suppose that X_1 and X_2 both put all their mass on the values $0, 2, 3$. Then, with

$$\phi(x_1, x_2) = x_1^2 + x_2^2$$

we would have that

$$E[\phi(X_1, X_2) | T = 3] > E[\phi(X_1, X_2) | T = 4] \quad \blacksquare$$

Suppose now that our simulation of the preceding section resulted in the estimate ϕ_i of $E[\phi(\mathbf{X}) | T = i]$ for $i = 0, \ldots, n$. If it results that ϕ_i is not nondecreasing in i, then we can modify these estimates by using the ideas of isotonic regression. Isotonic regression takes preliminary estimates $e_1, \ldots, e_n$ of unknown quantities that are known to be nondecreasing, and obtains final estimates $a_1, \ldots, a_n$ by solving the minimization problem

$$\min_{a_1 \leq \ldots \leq a_n} \sum_{i=1}^{n} (e_i - a_i)^2,$$

which can generally be solved in time linear in n (see[1] for details).

The following corollary will be used in the sequel.

Corollary 3.1. *Suppose that the random vectors $(X_i, W_i), i = 1, \ldots, n$ are independent, where $X_i, i = 1, \ldots, n$ are Bernoulli random variables, and where*

$$W_i | X_i = 1 \geq_{st} W_i | X_i = 0$$

where by the preceding we mean that $P(W_i \geq y | X_i = 1) \geq P(W_i \geq y | X_i = 0)$, for all y. Let $T = \sum_{i=1}^{n} X_i$. Then, for any nondecreasing function h, $E[h(W_1, \ldots, W_n) | T = k]$ is a nondecreasing function of k.

Proof: Let $g(\mathbf{X}) = E[h(\mathbf{W}) | \mathbf{X}]$ where $\mathbf{W} = (W_1, \ldots, W_n)$ and $\mathbf{X} = (X_1, \ldots, X_n)$. Because the random vectors are independent and W_i given $X_i = 1$ is stochastically larger than W_i given $X_i = 0$, it follows that $g(\mathbf{X})$ is a nondecreasing function of $\mathbf{X}$. Hence, by Theorem 3.1, it follows that $E[g(\mathbf{X}) | T = k]$ is a nondecreasing function of k. The result now follows because

$$\begin{aligned} E[h(\mathbf{W}) | T = k] &= E[E[h(\mathbf{W}) | T = k, \mathbf{X}] | T = k] \\ &= E[E[h(\mathbf{W}) | \mathbf{X}] | T = k] \\ &= E[g(\mathbf{X}) | T = k] \quad \blacksquare \end{aligned}$$

4. The Classical Reliability Model

Consider an n component system in which each component is either working or failed, and suppose that there exists a nondecreasing binary function ϕ such $\phi(x_1, \ldots, x_n)$ is 1 if the system works when x_i is the indicator variable for whether component i is working, $i = 1, \ldots, n$. The function ϕ is called the structure function, and to rule out trivialities we assume that $\phi(0, 0, \ldots, 0) = 0$ and $\phi(1, 1, \ldots, 1) = 1$. Now consider the problem of determining $E[\phi(X_1, \ldots, X_n)]$ when the X_i are independent Bernoulli random variables with $p_i = E[X_i]$, $i = 1, \ldots, n$. With $T = \sum_{i=1}^{n} X_i$, we can use the approach of the preceding sections to estimate $E[\phi(X_1, \ldots, X_n)]$ by doing a simulation that stratifies on T.

Call any set of components having the property that the system necessarily works when all of these components are working a *path* set. If no proper subset of a path set is itself a path set, call the path set a *minimal path set*. Let m_1 denote the size of the smallest minimal path set, and let

m_2 be the size of the largest minimal path set. Because the system works whenever $T > m_2$ and does not work when $T < m_1$, we have that

$$E[\phi(X_1, \ldots, X_n)] = \sum_{i=m_1}^{m_2} E[\phi(X_1, \ldots, X_n)|T = i]P(T = i) + P(T > m_2)$$

Consequently, we need only estimate the quantities $E[\phi(X_1, \ldots, X_n)|T = i]$ for $m_1 \leq i \leq m_2$, and this can be done using the approach of Section 2. If r simulation runs are planned then one can either do $\frac{rP(T=i)}{P(m_1 \leq T \leq m_2)}$ runs conditional on $T = i$, $m_1 \leq i \leq m_2$ or, better, do an initial small size simulation to estimate the conditional variances $\mathrm{Var}(\phi(X_1, \ldots, X_n)|T = i)$, $m_1 \leq i \leq m_2$ and then do a larger simulation in which the number of runs done conditional on $T = i$ is proportional to $P(T = i)$ times the square root of the estimate of the $\mathrm{Var}(\phi(X_1, \ldots, X_n)|T = i)$. If the resulting estimates of $E[\phi(X_1, \ldots, X_n)|T = i]$ are not monotone in i, then the estimates can be modified by an isotonic regression.

5. A Dynamic Reliability Model

Again suppose that ϕ is a structure function for an n component system. Suppose that each of the n components is initially working, and that component i works for random time W_i, $i = 1, \ldots, n$. In addition, suppose that $W_1, \ldots, W_n$ are independent, with W_i having distribution function F_i. If L is the amount of time that the system itself works, then

$$L = \max_{i=1,\ldots,s} \min_{j \in M_i} W_j$$

where $M_1, \ldots, M_s$ are the minimal path sets for the structure function ϕ.

Suppose that a cost $C(t)$ is incurred if the lifetime of the system is t, where, for some specified time t^*,

$$C(t) = h((t^* - t)^+)$$

where h is a nondecreasing function having $h(0) = 0$. In other works, there is no cost if system life exceeds t^* and a cost $h(s)$ if the system fails at time $t^* - s$, $s < t^*$. We are interested in using simulation to estimate $E[C(L)]$.

Let X_i equal 1 if component i is still working at time t^* and let it equal 0 otherwise. Then $E[\phi(\mathbf{X})]$ is the probability that the system life exceeds t^*. Let $T = \sum_{i=1}^{n} X_i$. With $p_i = 1 - F_i(t^*)$, $i = 1, \ldots, n$, let $P_j(i), i, j = 1, \ldots, n$ be the solution of (1).

We propose to estimate $E[C(L)]$ by stratifying on T. To begin, let m be the size of the largest minimal path set. Because the system cannot be failed if there are more than m working components, we have

$$E[C(L)] = \sum_{i=0}^{m} E[C(L)|T = i]P_1(i)$$

To simulate $C(L)$ conditional on $T = i$, $i = 0, \ldots, m$, first use the method given in Section 2 to generate $X_1, \ldots, X_n$ conditional on $T = i$. If $\phi(X_1, \ldots, X_n) = 1$ then take 0 as the estimate of $E[C(L)|T = i]$ from that run. If $\phi(X_1, \ldots, X_n) = 0$, then for any j for which $X_j = 0$, generate W_j according to the distribution

$$F_j^*(t) = P(W_j \leq t | W_j \leq t^*) = F_j(t)/F_j(t^*)\,, \quad 0 \leq t \leq t*$$

For j such that $X_j = 1$, set $W_j = t^*$. With L being the lifetime of the system according to the preceding values of W_j, take $C(L)$ as the estimate of $E[C(L)|T = i]$ from that run. Of course the number of runs to do conditional on $T = i$ should be determined by a preliminary small simulation study to estimate the quantities $\mathrm{Var}(C(L)|T = i)$, $i = 0, \ldots, m$.

Because the random vectors (X_i, W_i) satisfy the conditions of Corollary 2 and $C(L)$ is a nonincreasing function of $(W_1, \ldots, W_n)$ it follows from that Corollary that $E[C(L)|T = i]$ is a nonincreasing function of i. Hence, if the resulting estimates of $E[C(L)|T = i]$ are not monotone, then an isotonic regression can be employed to modify them.

A set of components is said to be a *cut* set if the system is necessarily failed when all components in this set are failed. A cut set is said to be a *minimal cut set* if none of its proper subsets are cut sets. A second simulation approach, which can be used when the number of minimal cut sets is not too large, uses an identity for Bernoulli sums. Suppose there are r minimal cut sets, $C_1, \ldots, C_r$, and for the set C_i define an indicator variable Z_i equal to 1 if all the components in C_i fail before time t^* and equal to 0 otherwise. Let $Z = \sum_{i=1}^{r} Z_i$, and let

$$\lambda = E[Z] = \sum_{i=1}^{r} \prod_{j \in C_i} (1 - p_j)$$

Now, it can be shown (see Section 11.3 of[7] for a proof) that for any random variable R

$$E[ZR] = \lambda E[R|Z_I = 1] \tag{3}$$

where I is independent of Z, R and is equally likely to be any of the values $1, \ldots, r$. Now, let

$$R = \begin{cases} 0 & \text{if } Z = 0 \\ \frac{C(L)}{Z} & \text{if } Z > 0 \end{cases}$$

Because $Z = 0$ if and only if $C(L) = 0$, the identity (3) yields that

$$E[C(L)] = \lambda\, E[\frac{C(L)}{Z}|Z_I = 1] \tag{4}$$

Using that

$$P(I = j|Z_I = 1) = \frac{P(Z_j = 1)}{\sum_{i=1}^{r} P(Z_i = 1)} = \frac{a_j}{\sum_{i=1}^{r} a_i}$$

where

$$a_i = \prod_{j \in C_i} (1 - p_j)$$

we can use the preceding to obtain a simulation estimate of $E[C(L)]$ by performing each simulation run as follows:

(1) Generate I such that $P(I = j) = \frac{a_j}{\sum_{i=1}^{r} a_i}$, $j = 1, \ldots, r$. Suppose $I = j$.
(2) For $i \in C_j$, generate W_i according to the distribution

$$F_i^*(t) = P(W_i \leq t | W_i \leq t^*) = F_i(t)/F_i(t^*)\,, \quad 0 \leq t \leq t^*$$

(3) For $i \notin C_j$, generate W_i according to the distribution F_i.
(4) Determine Z_i, $i = 1, \ldots, r$
(5) Determine $Z = \sum_{i=1}^{r} Z_i$
(6) Determine L
(7) Determine $C(L)$
(8) Return the estimator $\frac{\lambda\, C(L)}{Z}$

In contrast to our first estimator, the preceding estimator need not have a smaller variance than the raw simulation estimator. However, it should be very efficient when λ is small.

6. Modeling Basket Default Costs

A model for basket default swaps in which a portfolio consists of n assets, the i^{th} of which defaults at a random time having distribution F_i, was considered in[4] and.[2] It was supposed in these papers that if at least r

assets default by a fixed time t^* then a cost depending on L, the time of the r^{th} default, was incurred. Thus, the model of[4] is a special case of the model of the preceding section, in which the system structure is an $n-r+1$ of n system which works if and only if at least $n-r+1$ of the n components fail. Letting T be the number of assets that do not fail (i.e., do not default) by time t^*, it was suggested in,[2] as an improvement on the method of,[4] that $E[C(L)]$ be estimated by continually simulating the system conditional on the event that $T \leq n-r$, and then use the average of the values obtained for $C(L)$ multiplied by $P(T \leq n-r)$. That is,[2] notes that

$$E[C(L)] = E[C(L)|T \leq n-r]P(T \leq n-r)$$

and then uses simulation to estimate $E[C(L)|T \leq n-r]$. However, because our approach estimates $E[C(L)|T \leq n-r]$ by estimating all the quantities $E[C(L)|T = j]$, $j \leq n-r$, while using the known probabilities $P(T = j)$, it is a stratified version of the estimator of[2] and thus necessarily has a smaller variance. (Intuitively, there will be a lot more variance in the conditional distribution of $C(L)$ given that $T \leq n-r$, than there will be in the conditional distribution of $C(L)$ given that $T = j$, for any $j \leq n-r$.)

In cases where $P(T \geq r)$ is small, and $\binom{n}{r}$ is not too large, the second simulation method of the previous section should also be considered.

Example 2: The paper[2] gave an example in which there are 10 independent exponential random variables $W_i, i = 1, \ldots, 10$, with respective rates $0.03, 0.01, 0.02, 0.01, 0.005, 0.001,$
$0.002, 0.002, 0.017, 0.003,$
and with respective additive costs $0.3, 0.1, 0.2, 0.1, 0.3, 0.1, 0.2, 0.2, 0.1, 0.3$ incurred if $W_i < t, i = 1, \ldots, 10$. The method of[2] performed slightly better than raw simulation. The following indicates the variance of the raw estimator and of our estimator for different values of t.

t	raw estimator variance	variance of our estimator
5	0.021	0.003
10	0.034	0.006
15	0.045	0.008
20	0.052	0.009
25	0.057	0.010
30	0.062	0.011

As can be seen our estimator is far superior to the raw simulation estimator, and thus to the estimator of.[2]

7. A Round Robin Tournament

In a round robin tournament of $n+1$ players, each of the $\binom{n+1}{2}$ pairs play a match. The players who win the greatest number of matches are the winners of the tournament. Suppose that the results from all matches are independent, and that $P(i,j) = 1 - P(j,i)$ is the probability that i beats j in their match. Let I be the indicator of the event that player $n+1$ is the sole tournament winner, and suppose further that we want to use simulation to estimate $E[I] = P(I = 1)$. Letting X_i be the indicator of the event that player $n+1$ beats player i in their match, $i = 1, \ldots, n$, an efficient way to estimate $E[I]$ would be to let $T = \sum_{i=1}^{n} X_i$ and then do the simulation stratified on T. That is, compute $P(T = j), j = 0, 1, \ldots, n$, and use

$$E[I] = \sum_{j=[n/2]+1}^{n-1} E[I|T = j]P(T = j) + P(T = n)$$

To estimate $E[I|T = j]$, we would generate $X_1, \ldots, X_n$ conditional on $T = j$ and then generate the outcomes of the $\binom{n}{2}$ games that do not involve player $n+1$, and then take I as the estimator of $E[I|T = j]$ from that run. Antithetic variables should be effective, so when doing the next run we should use (in the same manner) the same uniforms (subtracted from 1) that were just used. As always it is advised to do a small simulation preparatory study to estimate the conditional variances, so as to set the number of runs done conditional in each strata in the final study.

Let W_i be the event that player i is the sole tournament winner, and suppose now that, rather than just wanting to estimate $P(W_{n+1})$, we want to estimate $P(W_i)$ for all $i = 1, \ldots, n+1$. In this situation, we suggest a post-stratification technique. Start by solving the $n+1$ sets of linear equations so as to determine the quantities

$$P(T_i = j), \quad j = 1, \ldots, n, \quad i = 1, \ldots, n+1$$

where T_i is the number of matches that player i wins. Now perform a fixed number of raw simulations of the tournament. Based on the results, let $N(i,j)$ be the number of simulation runs in which player i wins exactly j matches, and let $W(i,j)$ denote the number of simulation runs in which player i both wins exactly j matches and, in addition, is the sole winner of the tournament. Now, take $W(i,j)/N(i,j)$ as an estimate of $P(W_i|T_i = j)$. Since

$$P(W_i) = \sum_{j=1}^{n} P(W_i|T_i = j)P(T_i = j)$$

this yields our estimate of $P(W_i)$; namely,

$$\sum_{j=1}^{n} \frac{W(i,j)}{N(i,j)} P(T_i = j)$$

Acknowledgments

The research of Mark Brown was supported by the National Security Agency, under Grant H98230-06-01-0149.

References

1. Barlow, R. E., Bartholomew, D. J., Bremner, J. M., and Brunk, H. D., (1972), Statistical Inference under Order Restrictions: Isotonic Regression, Wiley, New York.
2. Chen, Z. and P. Glasserman, (2008), Fast Pricing of Basket Default Swaps, *Operations Research*, **56**, 2, 286-303.
3. Efron, B., (1965) Increasing Properties of Polya Frequency Functions. *Annals of Mathematical Statistics*, **36**, 272-279.
4. Joshi, M. and Kainth, D., (2004), Rapid and Accurate development of Prices and Greeks for nth to Default Credit Swaps in the Li Model, *Quantitative Finance*, Vol. 4, Institute of Physics Publishing, London, UK, 266-275.
5. Keilson, J., and Gerber, H., (1971), Some Results for Discrete Unimodality, *Jour. Amer. Statist. Assc.* **66**, 386-389.
6. Ross, S. M., (2007), Introduction to Probability Models, Nineth ed., *Academic Press*, Burlington, MA.
7. Ross, S. M., (2006), Simulation, fourth ed., *Academic Press*, Burlington, MA.