Mathematical Methods in Operations Research 2005
Special issue in honor of *Arie Hordijk*

Extending Simulation uses of Antithetic Variables: Partially Monotone Functions, Random Permutations, and Random Subsets

Sheldon M. Ross

Epstein Department of Industrial and Systems Engineering
University of Southern California
smross@usc.edu

Abstract We show how to effectively use antithetic variables to evaluate the expected value of (a) functions of independent random variables, when the functions are monotonic in only some of their variables, (b) Schur functions of random permutations, and (c) monotone functions of random subsets.

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

Analyzing risk models often involves a determination of the expected values of functions of multiple random variables. As it is often not possible to explicitly compute these values, simulation is often applied. Because there are usually many such quantities, depending on various risk assumptions, it is important that the simulations be done in such a manner as to quickly give accurate estimates of the desired quantities. One approach that has often been applied is to try to use successive simulation runs to obtain identically distributed unbiased run estimators that are not independent but rather are negatively correlated. The value of this is that it results in an unbiased estimator (the average of the estimates from all the runs) that has a smaller variance than would the average of identically distributed run estimators that are also independent. This raises the question of how we can obtain such identically distributed but negatively correlated estimators from successive simulation runs. One known result along this line is concerned with the situation where we want to estimate $\theta = E[h(\mathbf{U_n})]$, where $\mathbf{U_n} = (U_1, \ldots, U_n)$ is an n-vector of independent uniform $(0,1)$ random variables, and h is some specified function. It is known (see [1]) that if h is a monotone function of each of its coordinates (possibly increasing in some and decreasing in others) then the approach, known as the antithetic variables approach, of using a generated vector $\mathbf{U_n}$ to obtain two unbiased and identically distributed estimators $h(\mathbf{U_n})$ and $h(\mathbf{1_n} - \mathbf{U_n})]$, where $\mathbf{1_n}$ is an n-vector of $1's$, yields an estimator with a smaller variance than would be obtained in using the average of two independent estimators distributed as $h(\mathbf{U_n})$. In Section 2 of this paper, we show how this antithetic variable approach can be applied when h is only a monotone function of some of its variables.

It is of importance in homeland security models to be able to assess the consequences of an attack not only on a single target but also on a randomly chosen subset of targets. Thus, to analyze such problems by simulation we need to first generate a random subset. A question of interest is whether one should generate an independent random subset for each simulation run or rather utilize one that depends on past choices. To answer this question, in Sections 3 and 4 we consider using simulation to estimate the expected values of functions of random permutations (section 3) and then apply these results to functions of random subsets (section 4). In these sections we show when utilizing "antithetic" random permutations and random subsets in different simulation runs is better than utilizing independent ones.

2 Expected Values of Partially Monotone Functions

Suppose we want to use simulation to estimate $\theta = E[h(\mathbf{U_n})]$, where $\mathbf{U_n} = (U_1, \ldots, U_n)$ is an n-vector of independent uniform $(0,1)$ random variables, and where h is a function that is monotone in only some of its variables, say in its first r components. In this case, we show that

$$\mathrm{Var}[h(\mathbf{U_n}) + h(\mathbf{1_r} - \mathbf{U_r}, V_{r+1}, \ldots, V_n)] \leq 2\mathrm{Var}[h(\mathbf{U_n})] \tag{1}$$

when $U_1, \ldots, U_r, U_{r+1}, \ldots, U_n, V_{r+1} \ldots, V_n$ are all independent uniform $(0,1)$ random variables. Thus, when h is a monotone function of its first r variables, the random vector $U_1, \ldots, U_r \ldots, U_n$ should be generated and h evaluated at this vector. But then the next estimator should use the random vector $1 - U_1, \ldots, 1 - U_r$ along with an additional $n - r$ independently generated uniform $(0,1)$ random variables $V_{r+1}, \ldots, V_n$.

We prove (1) by proving the equivalent result:

Theorem 1: If h is monotone in its first r variables, then

$$\mathrm{Cov}(h(\mathbf{U_n}),\ h(\mathbf{1_r} - \mathbf{U_r}, V_{r+1}, \ldots, V_n)) \leq 0 \tag{2}$$

Proof of Theorem 1: Suppose h is monotonically increasing in its first r variables. Let $\mathbf{U} = (U_{r+1}, \ldots, U_n)$ and let $\mathbf{V} = (V_{r+1}, \ldots, V_n)$. Because, given $\mathbf{U}, \mathbf{V}$, the random variables $h(\mathbf{U_n})$ and $-h(\mathbf{1_r} - \mathbf{U_r}, V_{r+1}, \ldots, V_n)$ are both monotone increasing functions of $U_1, \ldots, U_r$ it follows that

$$\mathrm{Cov}(h(\mathbf{U_n}),\ -h(\mathbf{1_r} - \mathbf{U_r}, V_{r+1}, \ldots, V_n)|\mathbf{U}, \mathbf{V}) \geq 0$$

implying that

$$E[\mathrm{Cov}(h(\mathbf{U_n}),\ h(\mathbf{1_r} - \mathbf{U_r}, V_{r+1}, \ldots, V_n)|\mathbf{U}, \mathbf{V})] \leq 0 \tag{3}$$

Also, $E[(h(\mathbf{U_n})|\mathbf{U}, \mathbf{V})]$ and $-E[h(\mathbf{1_r} - \mathbf{U_r}, V_{r+1}, \ldots, V_n)|\mathbf{U}, \mathbf{V})]$ are both increasing in $U_1, \ldots, U_r$, implying that

$$\mathrm{Cov}(E[(h(\mathbf{U_n})|\mathbf{U}, \mathbf{V})],\ -E[h(\mathbf{1_r} - \mathbf{U_r}, V_{r+1}, \ldots, V_n)|\mathbf{U}, \mathbf{V})] \geq 0 \tag{4}$$

The result (2) now follows from (3) and (4) upon application of the conditional covariance formula that

$$\mathrm{Cov}(X, Y) = E[\mathrm{Cov}(X, Y|Z)] + \mathrm{Cov}(E[X|Z], E[Y|Z]).$$

Now suppose that h is monotone increasing in some of its first r coordinates and monotone decreasing in the remaining ones. For instance, suppose h is monotone increasing in its first k coordinates and monotone decreasing in its next $r - k$ coordinates, $k \leq r$. Then upon replacing $U_1, \ldots, U_r$ by $U_1, \ldots, U_k, 1 - U_{k+1}, \ldots, 1 - U_r$, the function h is monotone increasing in each of $U_1, \ldots, U_r$ and the argument proceeds as before. QED

3 Random Permutations

Let $I_1, \ldots, I_n$ be equally likely to be any of the $n!$ permutations of $1, \ldots, n$, and suppose we are interested in using simulation to estimate

$$\theta = E[f(v_{I_1}, \ldots, v_{I_n})],$$

for specified values $v_1 < v_2 < \ldots < v_n$, and a specified function f. After generating a random permutation $\mathbf{V} = (v_{I_1}, \ldots, v_{I_n})$, two "antithetic permutations" suggest themselves. Namely,

$$\mathbf{V_1} = (v_{I_n}, \ldots, v_{I_1}) \tag{5}$$

and

$$\mathbf{V_2} = (v_{n+1-I_1}, \ldots, v_{n+1-I_n}) \tag{6}$$

We now show that if h is a Schur convex or concave function (to be defined) then using either $\mathbf{V_1}$ or $\mathbf{V_2}$ in conjunction with $\mathbf{V}$ is better than evaluating h at $\mathbf{V}$ and at another random permutation independent of $\mathbf{V}$.

Let $\mathbf{v_1} = (v_{i_1}, \ldots, v_{i_n})$ and $\mathbf{v_2} = (v_{j_1}, \ldots, v_{j_n})$ both be permutations of $v_1, \ldots, v_n$. Say that $\mathbf{v_1}$ majorizes $\mathbf{v_2}$ if

$$\sum_{r=1}^{k} v_{i_r} \geq \sum_{r=1}^{k} v_{j_r} \quad \text{for all} \quad k = 1, \ldots, n$$

Say that h is a Schur convex (concave) function if $h(\mathbf{v_1}) \geq (\leq) h(\mathbf{v_2})$ whenever $\mathbf{v_1}$ majorizes $\mathbf{v_2}$.

In the following, suppose that $\mathbf{V} = (v_{I_1}, \ldots, v_{I_n})$ is equally likely to be any of the $n!$ permutations of $v_1, \ldots, v_n$, and suppose that $\mathbf{V_1}$ and $\mathbf{V_2}$ are as defined by (5) and (6).

Theorem 2 If g and h are either both Schur convex or both Schur concave functions defined on permutations of $v_1, \ldots, v_n$, then

$$\text{Cov}(g(\mathbf{V}),\, h(\mathbf{V_i})) \leq 0, \quad \mathbf{i = 1, 2}$$

To prove the theorem we will use the following lemma.

Lemma 1. If h is Schur convex, then $E[h(\mathbf{V})|I_1 = i]$ is increasing in i, and $E[h(\mathbf{V_1})|I_1 = i]$ and $E[h(\mathbf{V_2})|I_1 = i]$ are both decreasing in i. If h is Schur concave, then the preceding remains true when the terms "increasing" and "decreasing" are interchanged.

Proof of Lemma 1. Let $i > 1$, and let $\mathbf{P}$ denote the set of all $(n-2)!$ permutations of the values $1, \ldots, i-2, i+1, \ldots, n$. Then, for a Schur convex function h

$$\begin{aligned}
&E[h(\mathbf{V})|I_1 = i] \\
&= \frac{1}{(n-1)!} \sum_{(x_1,\ldots,x_{n-2})\in\mathbf{P}} \sum_{k=1}^{n-1} h(v_i, v_{x_1}, \ldots, v_{x_{k-1}}, v_{i-1}, v_{x_k}, \ldots, v_{x_{n-2}}) \\
&\geq \frac{1}{(n-1)!} \sum_{(x_1,\ldots,x_{n-2})\in\mathbf{P}} \sum_{k=1}^{n-1} h(v_{i-1}, v_{x_1}, \ldots, v_{x_{k-1}}, v_i, v_{x_k}, \ldots, v_{x_{n-2}}) \\
&= E[h(\mathbf{V})|I_1 = i-1]
\end{aligned}$$

where the inequality follows from the Schur convexity of h. Similarly,

$$\begin{aligned}
&E[h(\mathbf{V_1})|I_1 = i] \\
&= \frac{1}{(n-1)!} \sum_{(x_1,\ldots,x_{n-2})\in\mathbf{P}} \sum_{k=1}^{n-1} h(v_{x_1}, \ldots, v_{x_{k-1}}, v_{i-1}, v_{x_k}, \ldots, v_{x_{n-2}}, v_i) \\
&\leq \frac{1}{(n-1)!} \sum_{(x_1,\ldots,x_{n-2})\in\mathbf{P}} \sum_{k=1}^{n-1} h(v_{x_1}, \ldots, v_{x_{k-1}}, v_i, v_{x_k}, \ldots, v_{x_{n-2}}, v_{i-1}) \\
&= E[h(\mathbf{V_1})|I_1 = i-1]
\end{aligned}$$

Also, with $\mathbf{P}'$ denoting the set of all $(n-2)!$ permutations of the values $1, \ldots, n+3-i, n-i, \ldots, n$

$$\begin{aligned}
&E[h(\mathbf{V_2})|I_1 = i] \\
&\quad = \frac{1}{(n-1)!} \times \\
&\qquad \sum_{(x_1,\ldots,x_{n-2})\in\mathbf{P}} \sum_{k=1}^{n-1} h(v_{n+1-i}, v_{x_1}, \ldots, v_{x_{k-1}}, v_{n+2-i}, v_{x_k}, \ldots, v_{x_{n-2}}) \\
&\quad \leq \frac{1}{(n-1)!} \times \\
&\qquad \sum_{(x_1,\ldots,x_{n-2})\in\mathbf{P}} \sum_{k=1}^{n-1} h(v_{n+2-i}, v_{x_1}, \ldots, v_{x_{k-1}}, v_{n+1-i}, v_{x_k}, \ldots, v_{x_{n-2}}) \\
&\quad = E[h(\mathbf{V_2})|I_1 = i-1]
\end{aligned}$$

The proof for a Schur concave function is similar. QED

Proof of Theorem 2. The proof is by induction on n. Suppose g and h are Schur convex functions. As the theorem is true for $n = 1$ (because the covariance is 0 in this case), assume it to be true for $n-1$. Because the functions $g(v_j, x_1, \ldots, x_{n-1})$, $h(x_1, \ldots, x_{n-1}, v_j)$ and $h(v_{n+1-j}, x_1, \ldots, x_{n-1})$ are all Schur convex functions defined on the $(n-1)!$ permutations $(x_1, \ldots, x_{n-1})$ of $v_1, \ldots, v_{j-1}, v_{j+1}, \ldots, v_n$, it follows from the induction hypothesis that for $\mathbf{i} = \mathbf{1}, \mathbf{2}$

$$\text{Cov}(g(\mathbf{V}),\, h(\mathbf{V_i})|I_1 = j) \leq 0$$

implying that

$$E[\text{Cov}(g(\mathbf{V}),\, h(\mathbf{V_i})|I_1)] \leq 0$$

It follows from Lemma 1 that $E[g(\mathbf{V})|I_1]$ is increasing in I_1, whereas both $E[h(\mathbf{V_1})|I_1]$ and $E[h(\mathbf{V_2})|I_1]$ are decreasing in I_1. Consequently, for $\mathbf{i} = \mathbf{1}, \mathbf{2}$

$$\text{Cov}(E[g(\mathbf{V})|I_1],\, E[h(\mathbf{V_i})|I_1]) \leq 0$$

and the result follows from the conditional covariance identity. The proof when f and g are Schur concave is similar. QED

Remark: The theorem is not true without conditions on g. For instance, suppose the function g defined on permutations of $1, \ldots, n$ is large for permutations with 1 or n near the front or the back of the permutation.

Example 1. Suppose that n jobs must be processed sequentially on a single machine. Suppose that the processing times are $t_1 \leq t_2 \leq \cdots \leq t_n$, and that a reward $R(t)$ is earned whenever a job processing is completed at time t. Consequently, if the jobs are processed in the order $i_1, \ldots, i_n$, then the total reward is

$$h(i_1, \ldots, i_n) = \sum_{j=1}^{n} R(t_{i_1} + \ldots + t_{i_j})$$

If $R(t)$ is a monotone function of t then it easily follows that h is a Schur function (concave if h is decreasing, convex if h is increasing). Hence, when using simulation to approximate the expected total return when the processing order is equally likely to be any of the $n!$ orderings, each randomly generated permutation should be used twice, in the manner of Theorem 2.

4 Random Subsets

Suppose one wants to use simulation to determine $\theta = E[g(B)]$ where B is equally likely to be any of the $\binom{n}{k}$ subsets of $S = \{1, 2, \ldots, n\}$ that contain k elements, and g is a function defined on $k-$element subsets of S. Say that the function g is increasing (decreasing) if for all subsets A of size $k-1$, $g(A \cup i)$ is an increasing (decreasing) function of i for $i \notin A$. Now, rather than generating independent subsets of size k to estimate θ, one can also generate first a random $k-$element subset of S, call it R_1; then generate a random $k-$element subset from $S - R_1$, call it R_2; then generate a random $k-$element subset from $S - R_1 - R_2$, call it R_3, and so on. We now show that when g is a monotone function this latter approach results in a better estimate of θ than would be obtained from generating independent subsets.

Theorem 3. With R_j as specified in the preceding,

$$\mathrm{Cov}(g(R_i),\, g(R_j)) \leq 0$$

when g is either an increasing or decreasing function.

Proof: Suppose $n \geq 2k$. Define the function h on permutations of S by

$$h(i_1, \ldots, i_n) = g(i_1, \ldots, i_k)$$

Because h is a Schur function, it follows from Theorem 2 that for a random permutation $I_1, \ldots, I_n$, the covariance between $h(I_1, \ldots, I_n)$ and $h(I_n, \ldots, I_1)$ is nonnegative. But this means that

$$\mathrm{Cov}(g(I_1, \ldots, I_k),\, g(I_{n-k+1}, \ldots, I_n) \leq 0$$

The result now follows because the joint distribution of $g(I_1, \ldots, I_k)$, and $g(I_{n-k+1}, \ldots, I_n)$ is the same as the joint distribution of $g(R_i)$ and $g(R_j)$ whenever R_i and R_j are randomly chosen non-overlapping $k-$ element subsets of S. QED

Remarks:

(a) Suppose n is not an integral multiple of k, say $n = ki + j$, where $0 < j < k$. Then after generating $R_1, \ldots, R_i$ one could either start over, or (better yet) one can generate one additional $k-$element subset of S by using the j elements not in any of $R_1, \ldots, R_i$ along with a random selection of $k - j$ elements of $\cup_{t=1}^{i} R_t$

(b) Like Theorem 2, Theorem 3 is not true without some conditions on g. To see this, suppose that $n = 2k$ and that $g(R)$ is large (say equal to 1) when R contains exactly one of values $1, 2$ and is small (say, equal to 0) otherwise. Then $g(R)$ and $g(R^c)$ would be positively correlated.

5 References

1. Ross, S. M., *Simulation*, third ed., Academic Press, 2002