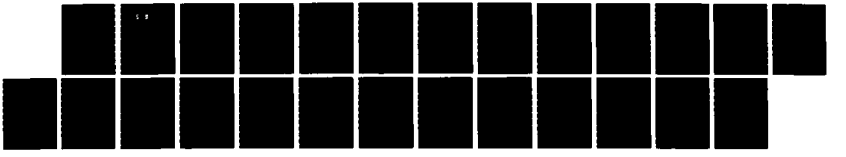
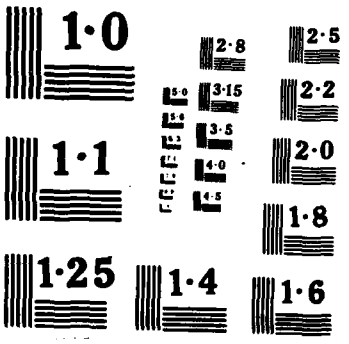
AD-A170 265 THE TOTAL HAZARD CONSTRUCTION ANTITHETIC VARIATES AND SIMULATION OF STOCH.. (U) ARIZONA UNIV TUCSON DEPT OF MATHEMATICS M SHAKED ET AL. AUG 85 AFOSR-TR-86-0401 1/1

UNCLASSIFIED AFOSR-84-0205 F/G 12/1 NL




1·0
1·1
1·25
1·4
1·6
1·8
2·0
2·2
2·5
2·8
3·15
3·5
4·0
4·5
5·0

(2)

Unclassi
SECURITY CLA

AD-A170 265

DTIC
SELECTE
JUL 28 1986
S D

:UMENTATION PAGE

1a. REPORT SECURITY CLASSIFICATION: Unclassified	1b. RESTRICTIVE MARKINGS
2a. SECURITY CLASSIFICATION A[illegible]	3. DISTRIBUTION/AVAILABILITY OF REPORT: Approved for Public Release; Distribution Unlimited
2b. DECLASSIFICATION/DOWNGRA[illegible]	
4. PERFORMING ORGANIZATION [illegible] NUMBER(S)	5. MONITORING ORGANIZATION REPORT NUMBER(S): AFOSR.TR. 86-0401
6a. NAME OF PERFORMING ORGANIZATION: University of Arizona	6b. OFFICE SYMBOL (If applicable):
7a. NAME OF MONITORING ORGANIZATION: AFOSR/NM	
6c. ADDRESS (City, State and ZIP Code): Tucson, AZ 85721	7b. ADDRESS (City, State and ZIP Code): Bldg. 410 Bolling AFB, DC 20332-6448
8a. NAME OF FUNDING/SPONSORING ORGANIZATION: AFOSR	8b. OFFICE SYMBOL (If applicable): NM
9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER: Grant AFOSR-84-0205	
8c. ADDRESS (City, State and ZIP Code): Bldg. 410 Bolling AFB, DC	10. SOURCE OF FUNDING NOS.

PROGRAM ELEMENT NO.	PROJECT NO.	TASK NO.	WORK UNIT NO.
6.1102F	2304	A5	

11. TITLE (Include Security Classification): The total hazard construction and simulation

12. PERSONAL AUTHOR(S): Moshe Shaked and J. George Shanthikumar

13a. TYPE OF REPORT	13b. TIME COVERED	14. DATE OF REPORT (Yr., Mo., Day)	15. PAGE COUNT
Technical Report	FROM _____ TO _____	August, 1985	21

16. SUPPLEMENTARY NOTATION

17 COSATI CODES		
FIELD	GROUP	SUB GR

18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number): Antithtic variates, multivariate dependence, Monte Carlo methods, simulation, generation of dependent variables, total hazard construction, coherent life functions, reliability thry.

19. ABSTRACT (Continue on reverse if necessary and identify by block number)

This paper combines recent developments in the area of generation of dependent random variables with the advantages of the use of common and antithetic random numbers. This combination yields new efficient methods for estimating complicated stochastic quantities by simulation. Some theoretical and practical aspects of use of antithetic and common random numbers for variance reduction while using the total hazard construction are given. A proof of their optimality in estimating the expected value of the response sum or the response difference of functions of vector arguments with dependent components is presented. Some numerical examples illustrate the theory.

20. DISTRIBUTION/AVAILABILITY OF ABSTRACT	21. ABSTRACT SECURITY CLASSIFICATION
UNCLASSIFIED/UNLIMITED ☒ SAME AS RPT ☐ DTIC USERS ☐	Unclassified

22a. NAME OF RESPONSIBLE INDIVIDUAL	22b. TELEPHONE NUMBER (Include Area Code)	22c. OFFICE SYMBOL
Brian W. Woodruff, MAj.		AFOSR/NM

DD FORM 1473, 83 APR EDITION OF 1 JAN 73 IS OBSOLETE. Unclassified

86 214

DTIC FILE COPY

The Total Hazard Construction, Antithetic Variates and Simulation of Stochastic Systems

Moshe Shaked
Department of Mathematics
University of Arizona

J. George Shanthikumar
School of Business Administration
University of California, Berkeley

August 1985

Supported by the Air Force Office of Scientific Research. U.S.A.F., under Grant AFOSR-84-0205. Reproduction in whole or in part is permitted for any purpose of the United States Government.

Abstract

This paper combines recent developments in the area of generation of dependent random variables with the advantages of the use of common and antithetic random numbers. This combination yields new efficient methods for estimating complicated stochastic quantities by simulation. Some theoretical and practical aspects of use of antithetic and common random numbers for variance reduction while using the total hazard construction are given. A proof of their optimality in estimating the expected value of the response sum or the response difference of functions of vector arguments with dependent components is presented. Some numerical examples illustrate the theory.

Key words and phrases. Antithetic variates, multivariate dependence, Monte Carlo methods, simulation, generation of dependent variables, total hazard construction, coherent life functions, reliability theory.

1. Introduction

Methods of variance reduction in simulation analysis are useful for the purpose of improving accuracy of estimates of output parameters of complex stochastic systems. The development of such methods is important because they yield estimators which are superior in accuracy to the crude Monte Carlo estimator.

One of the most successful variance reduction techniques is the use of common and antithetic random numbers (see the bibliography of Rubinstein, Samorodnitsky and Shaked (1985)).

This paper presents new ways of application of common and antithetic random variables. It combines recent developments in the area of generation of dependent random variables with the advantages of the use of common and antithetic random variates. This combination yields new efficient methods for estimating complicated probabilistic quantities by simulation (see Section 4).

More explicitly, in the present paper these methods are used in order to minimize the variance of $g(\underline{X}) - h(\underline{Y})$ where g and h are real measurable functions which are monotonic in the same (or the opposite) direction and the joint cdf (=cumulative distribution function) of $\underline{X}$ and $\underline{Y}$ is restricted to belonging to some set Θ of cdf's. To see an instance of such a set Θ, suppose, for example, that $\underline{X} = (X_1, ..., X_n)$ and $\underline{Y} = (Y_1, ..., Y_n)$ are both n-dimensional random vectors with respective cdf's $F_{\underline{X}}$ and $F_{\underline{Y}}$. Then Θ can be taken to be the set of all cdf's on R^{2n} with marginals $F_{\underline{X}}$ and $F_{\underline{Y}}$ or (see, e.g., Rubinstein, Samorodnitsky and Shaked (1985)) the set of all cdf's on R^{2n} with marginals $F_{\underline{X}}$ and $F_{\underline{Y}}$ such that dependence is permitted only between like components of $\underline{X}$ and $\underline{Y}$ (components with the same indices).

Usually $\underline{X}$ and $\underline{Y}$ are generated by constructing them from independent uniform random variables. In this respect, three alternative constructions have been used. They are (i) *standard construction* (e.g., Arjas and Lehtonen (1978)), (ii) *dynamic construction* (e.g., Shaked and Shanthikumar (1985b)), and (iii) *total hazard construction* (e.g., Norros (1984), Shaked and Shanthikumar (1985a)). Each of these constructions determine vector functions $\underline{T}_{\underline{X}}$ and $\underline{T}_{\underline{Y}}$ depending, respectively, on the cdf's of $\underline{X}$ and $\underline{Y}$ such that $\underline{X} \stackrel{d}{=} \underline{T}_{\underline{X}}(\underline{U})$ and $\underline{Y} \stackrel{d}{=} \underline{T}_{\underline{Y}}(\underline{V})$ where $\underline{U}$ and $\underline{V}$ are vectors of independent uniform [0,1] random variables (here $\stackrel{d}{=}$ denotes equality in law).

When $\underline{X}$ and $\underline{Y}$ are both n-dimensional (as will be assumed throughout the paper), then the standard and the total hazard constructions (to be described below in detail) require $\underline{U}$ and $\underline{V}$ to be n-dimensional. The dynamic construction requires $\underline{U}$ and $\underline{V}$ to be of (the same) dimension $n' > n$.

We say that VCRN (vector of common random numbers) is used if $\underline{V} = \underline{U}$. We say that VARN (vector of antithetic random numbers) is used if $\underline{V} = \underline{1} - \underline{U}$ where $\underline{1} \equiv (1, \ldots, 1)'$.

Let g and h be real measurable functions and suppose we are interested in reducing or minimizing $Var(g(\underline{X}) - h(\underline{Y}))$ subject to $F_{\underline{X}\underline{Y}} \in \Theta$ where Θ is some set of $2n$-dimensional cdf's. Assume we construct $\underline{X}$ and $\underline{Y}$ by $\underline{T}_{\underline{X}}(\underline{U})$ and $\underline{T}_{\underline{Y}}(\underline{V})$ where $\underline{U}$ and $\underline{V}$ are not necessarily independent vectors of independent uniform [0,1] random variables. Then

$$(1.1) \qquad Var(g(\underline{X}) - h(\underline{Y})) = Var(g^*(\underline{U}) - h^*(\underline{V}))$$

where

$$(1.2) \qquad g^*(\underline{u}) = g(\underline{T}_{\underline{X}}(\underline{u})),$$

$$(1.3) \qquad h^*(\underline{v}) = h(\underline{T}_{\underline{Y}}(\underline{v})).$$

Representation (1.1) together with the following result (Lemma 1.1 below) of Rubinstein and Samorodnitsky (1982) [see also Rubinstein, Samorodnitsky and Shaked (1985)] will be used to find optimal variance reduction techniques in Sections 2 and 3. First we need to define the following:

Let Φ be the set of $2n$-dimensional cdf's corresponding to random vectors $(\underline{U}, \underline{V})$ of uniform [0,1] random variables with n-dimensional marginals $F_{\underline{U}}(\underline{u}) = \prod_{i=1}^{n} u_i$ and $F_{\underline{V}}(\underline{v}) = \prod_{i=1}^{n} v_i, 0 \le u_i \le 1,\ 0 \le v_i \le 1,\ i = 1, \ldots, n$, with dependence permitted only between like components of $\underline{u}$ and $\underline{v}$.

Lemma 1.1. Let $\tilde{g}$ and $\tilde{h}$ be real measureable functions on $[0,1]^n$. (a) If $\tilde{g}$ and $\tilde{h}$ are monotonic in the same direction with respect to the i-th component, $i = 1, \ldots, n$, then

$$\min_{F_{\underline{U}\underline{V}} \in \Phi} Var(\tilde{g}(\underline{U}) - \tilde{h}(\underline{V})) = Var(\tilde{g}(\underline{U}) - \tilde{h}(\underline{U})). \tag{1.4}$$

(b) If $\tilde{g}$ and $\tilde{h}$ are monotonic in the opposite direction with respect to the i-th component, $i = 1, \ldots, n$, then

$$\min_{F_{\underline{U}\underline{V}} \in \Phi} Var(\tilde{g}(\underline{U}) - \tilde{h}(\underline{V})) = Var(\tilde{g}(\underline{U}) - \tilde{h}(\underline{1} - \underline{U})). \tag{1.5}$$

That is, use of VCRN is optimal for problem (1.4) and use of VARN is optimal for problem (1.5).

In the following sections we will use (1.1) and Lemma 1.1 as follows. For a given construction $\underline{T}_{\underline{X}}(\underline{U})$ and $\underline{T}_{\underline{Y}}(\underline{V})$ we will find conditions under which g^* of (1.2) and h^* of (1.3) are monotone in the same [respectively, opposite] direction and then from Lemma 1.1 it will follow that use of VCRN [respectively, VARN] is optimal.

In Section 2 we deal with variance reduction while using the standard construction. We briefly describe the main results of Rubinstein, Samorodnitsky and Shaked (1985) and their underlying ideas, and in doing this we motivate the developments of the following section. Section 3 deals with variance reduction while using the total hazard construction. We follow the lines of Section 2, thus obtaining new results which are the main contributions of this paper. Finally, in Section 4, we present some simulation results involving dependent stochastic variables.

2. The standard construction and variance reduction

First the standard construction is described in detail. The following notation is used. For any random vector $\underline{Z} = (Z_1, ..., Z_n)$, let

$$F_{Z_i \mid z_1,...,z_{i-1}}(z_i) = P\{Z_i \le z_i \mid Z_1 = z_1, ..., Z_{i-1} = z_{i-1}\}$$

and let

$$F^{-1}_{Z_i \mid z_1,...,z_{i-1}}(u_i) = \inf\{z_i : F_{Z_i \mid z_1,...,z_{i-1}}(z_i) \ge u_i\}, i = 1, ..., n.$$

According to the standard construction, $\underline{X} = (X_1, ..., X_n)$ and $\underline{Y} = (Y_1, ..., Y_n)$ are generated as follows:

(2.1) $$X_1 = F^{-1}_{X_1}(U_1) \equiv \tilde{a}_1(U_1),$$

(2.2) $$X_i = F^{-1}_{X_i \mid \tilde{a}_1(U_1),...,\tilde{a}_{i-1}(U_1,...,U_{i-1})}(U_i) \equiv \tilde{a}_i(U_1, ..., U_i), i = 2, ..., n,$$

(2.3) $$Y_1 = F^{-1}_{Y_1}(V_1) \equiv \tilde{b}_1(V_1),$$

(2.4) $$Y_i = F^{-1}_{Y_i \mid \tilde{b}_1(V_1),\ldots,\tilde{b}_{i-1}(V_1,\ldots,V_{i-1})}(V_i) \equiv \tilde{b}_i(V_1,\ldots,V_i), i = 2,\ldots,n,$$

where $\underline{U}$ and $\underline{V}$ are vectors of independent uniform [0,1] random variables. If we define $a_i(\underline{U}) \equiv \tilde{a}_i(U_1,\ldots,U_i)$ and $b_i(\underline{V}) \equiv \tilde{b}_i(V_1,\ldots,V_i)$ then (2.1) - (2.4) can be rewritten as

(2.5) $$X_i = a_i(\underline{U}), i = 1,\ldots,n,$$

(2.6) $$Y_i = b_i(\underline{V}), i = 1,\ldots,n.$$

Thus, the transformation $T_{\underline{X}}$ and $T_{\underline{Y}}$, mentioned in Section 1, which corresponds to the standard construction is described in (2.5)-(2.6).

A random vector $\underline{Z} = (Z_1,\ldots,Z_n)$ is said to be CIS (conditionally increasing in sequence) if $P(Z_i > z_i \mid Z_1 = z_1,\ldots,Z_{i-1} = z_{i-1})$ is nondecreasing in $z_1,\ldots,z_{i-1}$ for every $z_i, i = 2,\ldots,n$. Rubinstein, Samorodnitsky and Shaked (1985) proved:

Lemma 2.1. If $\underline{X}$ and $\underline{Y}$ are CIS then $a_i(\underline{u})$ and $b_i(\underline{v})$ [of (2.5)-(2.6)] are nondecreasing functions of $\underline{u}$ and $\underline{v}, i = 1,\ldots,n$.

That is, if $\underline{X}$ and $\underline{Y}$ are CIS then the transformations $T_{\underline{X}}$ and $T_{\underline{Y}}$ defined by (2.5)-(2.6) are monotone nondecreasing.

Let Ξ be the set of all cdf's on R^{2n} with marginals $F_{\underline{X}}$ and $F_{\underline{Y}}$ such that dependence is permitted only between like components of $\underline{U}$ and $\underline{V}$ of (2.5) and (2.6). Consider the problem

(2.7) minimize $Var(g(\underline{X}) - h(\underline{Y}))$ subject to $F_{\underline{X}\underline{Y}} \in \Xi$.

From Lemmas 1.1 and 2.1 one easily obtains (see, e.g., Rubinstein, Samorodnitsky and Shaked (1985)):

Theorem 2.1. If $\underline{X}$ and $\underline{Y}$ are CIS and g and h are monotone in the same (the opposite) direction then use of VARN (VCRN) yields an optimal solution for problem (2.7).

Note that the cdf $F_{\underline{X}\underline{Y}}$, which corresponds to the choice of independent $\underline{U}$ and $\underline{V}$ in (2.5)-(2.6), belongs to Ξ. Thus, by Theorem 2.1, the use of VCRN (VARN) is preferable to the use of independent $\underline{U}$ and $\underline{V}$. That is, use of VCRN (VARN) is superior to the crude Monte Carlo method. Furthermore, use of VCRN or VARN requires generation of only *half* as many uniform [0,1] random variables.

3. The total hazard construction and variance reduction.

The total hazard construction can be used to generate absolutely continuous nonnegative random vectors (see, e.g., Norros (1984) and Shaked and Shanthikumar (1985a)). It is particularly useful when the numerical inversion of the conditional distributions, described in (2.1)-(2.4), is too involved and time consuming, and on the other side, inversion of the multivariate hazard functions, to be defined below, is simple. An example of such an instance is given in Section 4.

Let $\underline{X} = (X_1, ..., X_n)$ be an absolutely continuous nonnegative random vector to be generated. We will use the following notation.

For $J = \{j_1, ..., j_k\} \subset \{1, ..., n\}$ let $\underline{x}_J$ denote $(x_{j_1}, ..., x_{j_k})$. If $\bar{J} = \{i_1, ..., i_{n-k}\}$ then $\underline{x}_{\bar{J}}$ denotes $(x_{i_1}, ..., x_{i_{n-k}})$. Let $\underline{1} = (1, ..., 1)'$. The length of $\underline{1}$ will vary from one formula to another, but it will be always possible to determine it from the expression in which $\underline{1}$ appears.

For $J \subset \{1, ..., n\}$ and $i \in \bar{J}$ let $\lambda_i(x \mid \underline{X}_J = \underline{x}_J, \underline{X}_{\bar{J}} > x\underline{1})$ denote the conditional hazard rate of X_i at time x given that $\underline{X}_J = \underline{x}_J$ and that $\underline{X}_{\bar{J}} > x\underline{1}$, where $x \geq \bigvee_{j \in J} x_j \equiv \max\{x_j : j \in J\}$. If $J = \phi$ then $\bigvee_{j \in J} x_j \equiv 0$. Formally, for $i \in \bar{J}$,

(3.1) $\lambda_i(x \mid \underline{X}_J = \underline{x}_J, \underline{X}_{\bar{J}} > x\underline{1})$

$$\equiv \lim_{\Delta x \downarrow 0} \frac{1}{\Delta x} P\left\{x < X_i \le x + \Delta x \mid \underline{X}_J = \underline{x}_J, \underline{X}_{\bar{J}} > x\underline{1}\right\}, \; x \ge \bigvee_{j \in J} x_j$$

(J may be empty.) The absolute continuity of $\underline{X}$ ensures that limit exists. To save space we sometimes supress the condition $\underline{X}_{\bar{J}} > x\underline{1}$ and just write $\lambda_i(x \mid \underline{X}_J = \underline{x}_J, \bullet)$ but the reader should keep in mind that "$\bullet$" means $\underline{X}_{\bar{J}} > x\underline{1}$ with x being the same as the first argument of λ_i. Note that $\lambda_i(x \mid \underline{X}_J = \underline{x}_J, \bullet)$ is well defined for all $x \ge \bigvee_{j \in J} x_j$.

For $i \in \bar{J}$ the total hazard accumulated by X_i during the time interval $\left[\bigvee_{j \in J} x_j, \bigvee_{j \in J} x_j + x\right], x \ge 0$, is defined by

$$\Lambda_i(x \mid \underline{X}_J = \underline{x}_J) \equiv \int_{\bigvee_{j \in J} x_j}^{\bigvee_{j \in J} x_j + x} \lambda_i(u \mid \underline{X}_J = \underline{x}_J, \bullet) du, \quad x \ge 0, i \in \bar{J} \; .$$

When $J = \phi$, $\Lambda_i(x \mid \underline{X}_J = \underline{x}_J)$ will be simply denoted by $\Lambda_i(x)$.

We will introduce now a notation for the total hazard accumulated by X_i by time x. Fix $x > 0$ and suppose that it is given that $X_{j_1}, \ldots, X_{j_{k-1}} (k > 1)$ failed at times $x_{j_1}, \ldots, x_{j_{k-1}}$, respectively ($x_{j_1} < \ldots < x_{j_{k-1}} \le x$) and that all other X_j's are alive at time x. For $i \notin \{j_1, \ldots, j_{k-1}\}$ denote

$$\Psi_{i \mid j_1, \ldots j_{k-1}}(x \mid x_{j_1}, \ldots, x_{j_{k-1}}) \equiv \Lambda_i(x_{j_1})$$

$$+ \sum_{\ell=2}^{k-1} \Lambda_i(x_{j_\ell} - x_{j_{\ell-1}} \mid X_{j_1} = x_{j_1}, \ldots, X_{j_{\ell-1}} = x_{j_{\ell-1}})$$

$$+ \Lambda_i(x - x_{j_{k-1}} \mid X_{j_1} = x_{j_1}, \ldots, X_{j_{k-1}} = x_{j_{k-1}}) \; .$$

Also denote (corresponding to the case $k=1$)

$$\Psi_i(x) = \Lambda_i(x), x \geq 0.$$

The *total hazard accumulated by X_i by the time it failed*, given that X_i was the k-th to fail and that $X_{j_1}, ..., X_{j_{k-1}}$ failed before X_i, is $\Psi_{i|j_1,...j_{k-1}}(X_i \mid X_{j_1}, ..., X_{j_{k-1}})$.

Define the inverse functions

$$\Lambda_i^{-1}(t) = \inf\{x \geq 0 : \Lambda_i(x) \geq t\}, \quad i = 1, ..., n, t \geq 0,$$

and for nonempty $J \subset \{1, ..., n\}$, $\underline{x_J} > 0\underline{1}$ and $i \in \bar{J}$,

$$\Lambda_{i|J}^{-1}(t \mid \underline{x_J}) = inf\{x \geq 0 : \Lambda_i(x \mid \underline{X_J} = \underline{x_J}) \geq t\}, t \geq 0.$$

Let $U_1, ..., U_n$ be independent uniform [0,1] random variables. They generate independent standard (i.e., mean one) exponential random variables as follows:

(3.3) $$E_i = -\log(1 - U_i), \quad i = 1, ..., n.$$

The total hazard construction consists of transforming $E_1, ..., E_n$ into $X_1, ..., X_n$ as follows:

Step 1. Let j_1 be the (random) index (which, by absolute continuity, is unique with probability one) such that $\Lambda_{j_1}^{-1}(E_{j_1}) = \min\{\Lambda_i^{-1}(E_i) : i = 1, ..., n\}$ and define

(3.4) $$X_{j_1} = \Lambda_{j_1}^{-1}(E_{j_1}).$$

Step k (k=2,...,n). Given that steps 1,2,...,k-1 resulted in $X_{j_1},...,X_{j_{k-1}}$ let $J=\{j_1,...,j_{k-1}\}$. Let j_k be the (random) index (which, by absolute continuity, is unique with probability one) such that

$$\Lambda_{j_k|J}^{-1}\left[E_{j_k}-\Psi_{j_k|j_1,...j_{k-1}}(X_{j_{k-1}}|X_{j_1},...,X_{j_{k-1}})\,|\,X_{j_1},...,X_{j_{k-1}}\right]$$

$$=\min_{i\in J}\left\{\Lambda_{i|J}^{-1}\left[E_i-\Psi_{i|j_1,...j_{k-1}}(X_{j_{k-1}}|X_{j_1},...,X_{j_{k-1}})\,|\,X_{j_1},...,X_{j_{k-1}}\right]\right\}\ .$$

It is easy to verify, by induction, that the arguments of Λ_i^{-1} and $\Lambda_{j_k}^{-1}$, in the above expression, are nonnegative. Having chosen the (random) index j_k as described above, define (here $J=\{j_1,...,j_{k-1}\}$)

(3.5) $$X_{j_k}=X_{j_{k-1}}$$

$$+\Lambda_{j_k|J}^{-1}\left[E_{j_k}-\Psi_{j_k|j_1,...j_{k-1}}(X_{j_{k-1}}|X_{j_1},...,X_{j_{k-1}})\,|\,X_{j_1},...,X_{j_{k-1}}\right],\ k=2,...,n\ .$$

Notice that (3.4)-(3.5) describe explicitly how to obtain $\underline{X}$ from $\underline{E}=(E_1,...,E_n)$. However (3.4)-(3.5) together with (3.3) define a transformation of $\underline{U}$ into $\underline{X}$. This is the transformation $\underline{T}_{\underline{X}}$, mentioned in Section 1, which corresponds to the total hazard construction. Shaked and Shanthikumar (1985a) proved that indeed (3.4)-(3.5) yields a random vector with the desired cdf. Analogously, starting from independent uniform [0,1] random variables $V_1,...,V_n$ one can generate a nonnegative absolutely continuous random vector $(Y_1,...,Y_n)$ with a desired cdf. That is, the transformation $\underline{T}_{\underline{Y}}$, mentioned in Section 1, which corresponds to the total hazard construction, is described in (3.3)-(3.5) with obvious modification of notation.

In order to give conditions under which $\underline{T}_{\underline{X}}$ and $\underline{T}_{\underline{Y}}$ are nondecreasing, we need the following definition. For $2\le k\le n$ we use the notation

$$c_k(t_k, x_1, ..., x_{k-1})$$

$$\equiv x_{k-1} + \Lambda^{-1}_{k|1,...,k-1}\left[t_k - \Psi_{k|1,...,k-1}(x_{k-1} \mid x_1, ..., x_{k-1}) \mid x_1, ..., x_{k-1}\right]$$

which describes, according to (3.5), how X_k is determined, given that $j_1 = 1, ..., j_{k-1} = k-1, j_k = k$ and that $E_k = t_k, X_1 = x_1, ..., X_{k-1} = x_{k-1}$.

Definition 3.1. The absolutely continuous nonnegative random vector $\underline{X} = (X_1, ..., X_n)$ is said to have CDTH (*conditionally decreasing total hazard*) if

$$\Psi_{i|1,...,i-1}\left[x_i \mid x_1, ..., x_k, c_{k+1}(t_{k+1}, x_1, ..., x_k)\right.,$$

$$\left. c_{k+2}(t_{k+2}, x_1, ..., x_k, c_{k+1}), ..., c_{i-1}(t_{i-1}, x_1, ..., x_k, c_{k+1}, ..., c_{i-2})\right]$$

(the arguments of some of the c_m's are omitted) nonincreases in $x_k \in \{x_k : x_k \geq x_{k-1}, c_{k+1} \geq x_k;\ c_{\ell+1} \geq c_\ell, \ell = k+1, ..., i-2\}$ for all $1 \leq k+1 \leq i \leq n$, $t_{k+1} \geq 0, ..., t_{i-1} \geq 0$, $0 \leq x_1 \leq ... \leq x_{k-1}$, and if the above condition holds for all permutations of the indices 1,2,...,n.

Definition 3.1 is essentially the same as the definition of *supportive system* of Norros (1984). An easy to check condition which implies the CDTH property is given in (3.7) below.

From the proof of Theorem 4.4 of Shaked and Shanthikumar (1985a) the following result follows:

Lemma 3.1. If the absolutely continuous nonnegative random vectors $\underline{X} = (X_1, ..., X_n)$ and $\underline{Y} = (Y_1, ..., Y_n)$ have CDTH then the transformations $T_{\underline{X}}$ and $T_{\underline{Y}}$ defined by (3.3)-(3.5) are monotone nondecreasing.

Let Ξ' be the set of all cdf's on R_+^{2n} with marginals $F_{\underline{X}}$ and $F_{\underline{Y}}$ such that dependence is permitted only between like components of $\underline{U}$ and $\underline{V}$ of the transformations $T_{\underline{X}}(\underline{U})$ and $T_{\underline{Y}}(\underline{V})$ defined by (3.3)-(3.5). Consider the problem

(3.6) minimize $Var(g(\underline{X}) - h(\underline{Y}))$ subject to $F_{\underline{X}\underline{Y}} \in \Xi'$.

From Lemmas 1.1 and 3.1 one easily obtains

Theorem 3.1. If $\underline{X}$ and $\underline{Y}$ have CDTH and g and h are monotone in the same (the opposite) direction then use of VCRN (VARN) yields an optimal solution for problem (3.6).

Note that the cdf $F_{\underline{X}\underline{Y}}$, which corresponds to the choice of independent $\underline{U}$ and $\underline{V}$ in the transformation defined by (3.3)-(3.5), belongs to Ξ'. Thus, by Theorem 3.1, the use of VCRN (or VARN) is preferable to the use of crude Monte Carlo method (i.e., independent $\underline{U}$ and $\underline{V}$).

Shaked and Shanthikumar (1985a) showed that if, for a random vector $\underline{X}$, the multivariate conditional hazard rates (defined in (3.1)) nondecrease as functions of the number of past failures and nonincrease as functions of the failure times of these past failures, then $\underline{X}$ has CDTH. That is, $\underline{X}$ has CDTH if more and earlier failures cause a higher risk for the surviving components than fewer and later failures.

More explicitly Shaked and Shanthikumar (1985a) showed that if for disjoint sets $I, J \subset \{1, \ldots, n\}$ and fixed $\underline{x}_I, \hat{\underline{x}}_I, \underline{x}_J$ (such that $\underline{x}_I \leq \hat{\underline{x}}_I$) and $k \in \overline{IUJ}$ (I or J may be empty),

$$(3.7) \quad \lambda_k((\underset{i \in I}{V} \hat{x}_i) V (\underset{j \in J}{V} x_j) + u \mid \underline{X}_I = \underline{x}_I, \underline{X}_J = \underline{x}_J, \bullet)$$

$$\geq \lambda_k((\underset{i \in I}{V} \hat{x}_i) V (\underset{j \in J}{V} x_j) + u \mid \underline{X}_I = \hat{\underline{x}}_I, \bullet), u \geq 0,$$

then $(X_1, ..., X_n)$ has CDTH. Thus, from Theorem 3.1 we get

Corollary 3.1. If the multivariate hazard rates of $\underline{X}$ satisfy (3.7) and the multivariate hazard rates of $\underline{Y}$ satisfy the same condition with proper notational modifications, and if g and h are monotone in the same (the opposite) direction then use of VCRN (VARN) yields an optimal solution for problem (3.6).

Norros (1984) has obtained another condition which implies that the transformation defined by (3.4)-(3.5) is monotone nondecreasing. Thus, if Norros' condition holds for $\underline{X}$ and $\underline{Y}$ and if g and h are monotone in the same (the opposite) direction then use of VCRN (VARN) yields an optimal solution for problem (3.6).

4. Some applications and examples

In this section we illustrate the use of the results of Section 3 for simulation purposes.

Suppose we want to estimate $\theta^{(1)} = E[g(\underline{X})]$ where $\underline{X}$ has CDTH and g is a real valued measurable monotonic function. An unbiased estimator for $\theta^{(1)}$ is the crude Monte Carlo estimator which can be written as

$$(4.1) \qquad \bar{\theta}^{(1)} = \frac{1}{k}\sum_{m=1}^{k} g(\underline{X}^{(m)}).$$

Here $\underline{X}^{(1)}, \ldots, \underline{X}^{(k)}$ is a sample from the cdf $F_{\underline{X}}$ which is generated by the total hazard construction (3.3)-(3.5).

Another unbiased estimator for $\theta^{(1)}$ is

$$(4.2) \qquad \bar{\theta}_a^{(1)} = \frac{1}{k}\sum_{m=1}^{\frac{k}{2}} \left[g(\underline{X}^{(m)}) + g(\underline{X}_a^{(m)})\right] \quad \text{(k is even)}$$

where $\underline{X}^{(m)}$ is generated by (3.3)-(3.5) using the vector $\underline{U}^{(m)}$ of independent uniform [0,1] random variables, and $\underline{X}_a^{(m)}$ is generated by (3.3)-(3.5) using $\underline{1} - \underline{U}^{(m)}$. We call $\bar{\theta}_a^{(1)}$ the *antithetic estimator* of $\theta^{(1)}$.

Since $g(\underline{X}) + g(\underline{X}_a)$ is a special case of $g(\underline{X}) - h(\underline{Y})$ [with $h(\underline{Y})$ replaced by $-g(\underline{X}_a)$] we obtain from Theorem 3.1 that $Var\left[\bar{\theta}_a^{(1)}\right] \le Var\left[\bar{\theta}^{(1)}\right]$. That is, the antithetic estimator is more accurate than the crude Monte Carlo estimator $\bar{\theta}^{(1)}$.

Assume now that we seek to estimate $\theta^{(2)} = E[g(\underline{X}) - h(\underline{Y})]$, $\underline{X}, \underline{Y} \in R_+^n$, i.e., the expected value of response difference of a pair of functions $g(\underline{X})$ and $h(\underline{Y})$. Assume that both $\underline{X}$ and $\underline{Y}$ have CDTH and that both g and h are monotone functions. The crude Monte Carlo estimator of $\theta^{(2)}$ is

$$\bar{\theta}^{(2)} = \frac{1}{k}\sum_{m=1}^{k}\left[g(\underline{X}^{(m)}) - h(\underline{Y}^{(m)})\right] \tag{4.3}$$

where $\underline{X}^{(m)}$ and $\underline{Y}^{(m)}$ are generated independently by the transformation described by (3.3)-(3.5).

The antithetic estimator for $\theta^{(2)}$ is

$$\bar{\theta}_a^{(2)} = \frac{1}{k}\sum_{m=1}^{k}\left[g(\underline{X}^{(m)}) - h(\underline{Y}_a^{(m)})\right] \tag{4.4}$$

where $\underline{X}$ and $\underline{Y}$ are generated using VCRN, i.e., $\underline{X}^{(m)}$ is generated by (3.3)-(3.5) using $\underline{U}^{(m)}$ and $\underline{Y}_a^{(m)}$ is generated similarly also using $\underline{U}^{(m)}$.

It follows from Theorem 3.1 that $Var\left[\bar{\theta}_a^{(2)}\right] \le Var\left[\bar{\theta}^{(2)}\right]$ i.e., the antithetic estimator is more accurate than the crude Monte Carlo estimator.

The monotonicity assumption in Theorem 3.1 actually holds for a broad class of complex stochastic systems (see, e.g., Rubinstein, Samorodnitsky and Shaked (1985)). Here we take

$$g(x_1, ..., x_5) = \max(\min(x_1, x_2),\ \min(x_1, x_3, x_5), \min(x_4, x_5)),$$

$$h(y_1,...,y_5) = \max(\min(y_1,y_2), \min(y_1,y_3,y_5), \min(y_4,y_5)).$$

These represent the lives of coherent systems described by a Wheatstone Bridge.

Suppose $\underline{X} = (X_1,...,X_5)$ has hazard rates which increase with the number of failed components. For example, for $\alpha > 0, \xi > 0$, let

$$\lambda_i(x \mid \underline{X}_J = \underline{x}_J, \cdot) = \alpha(|J| + \xi), i \in \bar{J}, J \subset \{1,...,5\}, \tag{4.5}$$

where $|J|$ denotes the number of elements of J. Then the λ_i's satisfy (3.7) and thus $\underline{X}$ has CDTH.

Let $\underline{Y} = (Y_1,...,Y_5)$ have hazard rates of the form (4.5), that is, for $\beta > 0, \eta > 0$, let (using obvious notation)

$$\mu_i(y \mid \underline{Y}_J = \underline{y}_J, \cdot) = \beta(|J| + \eta), i \in \bar{J}, J \subset \{1,...,5\} . \tag{4.6}$$

Then also $\underline{Y}$ has CDTH.

For the choice of the cdf which corresponds to (4.5) it is easy to see that

$$\Lambda_i(x) = \alpha\xi x, \quad x \geq 0, i = 1,...,5,$$

$$\Lambda_i(x \mid \underline{X}_J = \underline{x}_J) = \alpha(|J| + \xi)x, \quad x \geq 0, i \in \bar{J}, J \subset \{1,...,5\} .$$

Thus,

$$\Lambda_i^{-1}(t) = [\alpha\xi]^{-1} t, \quad t \geq 0, i = 1,...,5,$$

$$\Lambda^{-1}_{i|J}(t \mid \underline{x}_J) = [\alpha(|J| + \xi)]^{-1} t, \quad t \geq 0, i \in \bar{J}, J \subset \{1, ..., 5\} \ .$$

These were used in (3.4)-(3.5) to generate $\underline{X}^{(i)}$. The vectors $\underline{Y}^{(i)}$ were generated similarly.

The parameters $\theta^{(1)} = E[g(\underline{X})]$ and $\theta^{(2)} = E[g(\underline{X}) - h(\underline{Y})]$ were estimated for various choices of α, β, ξ and η using, respectively, $\bar{\theta}^{(1)}, \bar{\theta}^{(1)}_a, \bar{\theta}^{(2)}$ and $\bar{\theta}^{(2)}_a$. The estimators $\bar{\theta}$ and $\bar{\theta}_a$ were calculated according to (4.1) and (4.2) while estimating $\theta^{(1)}$ and according to (4.3) and (4.4) while estimating $\theta^{(2)}$.

A sensible performance measure is $(t_{\bar{\theta}} \mathrm{Var}(\bar{\theta}) / t_{\bar{\theta}_a} \mathrm{Var}(\bar{\theta}_a))$ where $t_{\bar{\theta}}$ and $t_{\bar{\theta}_a}$ are the CPU times for the crude Monte Carlo and the antithetic estimators respectively. Comparing (4.1) and (4.2) [or (4.3) and (4.4)] we see that $t_{\bar{\theta}_a} \leq t_{\bar{\theta}}$ because the antithetic estimator $\bar{\theta}_a$ needs only half as many random numbers as it is needed by its counterpart $\bar{\theta}$. In the following we shall neglect this advantage of $\bar{\theta}_a$ and take the performance measure to be

(4.7) $\quad \varepsilon = Var(\bar{\theta}) \ / \ Var(\bar{\theta}_a).$

Table 1

The efficiency ε of the antithetic estimator

$\theta_a^{(1)}$ as a function of α and ξ

ξ \ α	1	2	3	4	5
1	2.05	2.83	2.63	2.59	2.27
2	1.90	2.45	1.88	2.31	2.40
3	2.21	2.41	2.16	1.67	1.58
4	2.10	1.83	1.93	1.65	2.05
5	1.81	1.71	1.79	1.53	2.18

Table 1 gives the values of ε while estimating $\theta^{(1)}$ for various choices of α and ξ. Here, in order to compute $\bar{\theta}^{(1)}$ we used $k = 1000$ replications of $\underline{X}$. For the computation of $\bar{\theta}_a^{(1)}$ we used $k/2 = 500$ replications of the pair $(\underline{X}, \underline{X}_a)$. Since

$$Var(\bar{\theta}^{(1)}) = Var(\frac{1}{k}\sum_{m=1}^{\frac{k}{2}} g(\underline{X}^{(2m-1)}) + g(\underline{X}^{(2m)})),$$

from (4.7) one sees that $\varepsilon = Var(g(\underline{X}^{(1)}) + g(\underline{X}^{(2)})) \,/\, Var(g(\underline{X}^{(1)}) + g(\underline{X}_a^{(1)}))$. Hence ε is estimated by $\hat{\varepsilon} = s^2 / s_a^2$, where s^2 is the sample variance of the 500 observations $(g(\underline{X}^{(2m-1)}) + g(\underline{X}^{(2m)}))$, $m = 1, 2, ..., 500$ and s_a^2 is the sample variance of the 500 observations $(g(\underline{X}^{(m)}) + g(\underline{X}_a^{(m)}))$, $m = 1, 2, ..., 500$.

Table 2

The efficiency of ε of the antithetic estimator $\theta_a^{(2)}$ as a function of α, ξ, β and η

(η, ξ) \ (β, α)	(2,1)	(3,2)	(4,3)	(5,4)	(6,5)
(1,1)	4.71	14.73	22.80	38.63	69.94
(2,2)	4.91	15.08	21.00	37.28	62.86
(3,3)	4.98	14.88	23.93	39.53	62.47
(4,4)	5.13	11.51	25.00	43.01	66.10
(5,5)	4.89	11.00	19.83	36.20	58.87

Table 2 gives the values of ε while estimating $\theta^{(2)}$ with various choices of α, ξ, β and η. For the computation of $\bar{\theta}^{(2)}$ we used $k = 500$ replications of $(\underline{X}, \underline{Y})$, where $\underline{X}$ and $\underline{Y}$ are independent. For the computation of $\bar{\theta}_a^{(2)}$ we used $k = 500$ replications of $(\underline{X}, \underline{Y}_a)$. Similar to the earlier case ε is estimated by $\hat{\varepsilon} = s^2 / s_a^2$, where s^2 is the sample variance of the 500 observations $g(\underline{X}^{(m)}) - h(\underline{Y}^{(m)}), m = 1, 2, ..., 500$ and s_a^2 is the sample variance of the 500 observations $g(\underline{X}^{(m)}) - h(\underline{Y}_a^{(m)}), m = 1, 2, .., 500$.

These numerical results clearly indicate the benefits of using antithetic and common random numbers for variance reduction.

Bibliography

1. Arjas, E. and Lehtonen, T. (1978), Approximating many server queues by means of single server queues. *Math. Opr. Res.*, 3, 205-223.

2. Norros, I. (1984), A compensator representation of multivariate life length distributions, with applications, submitted for publication.

3. Rubinstein, Y.R. and Samorodnitsky, G. (1982). Antithetic and common random numbers in simulation of stochastic systems. Mimeograph Series No. 308, Technion, Haifa.

4. Rubinstein, Y.R., Samorodnitsky, G. and Shaked, M. (1985), Antithetic variates, multivariate dependence and simulation of stochastic systems, *Management Science*, 31, 66-77.

5. Shaked, M. and Shanthikumar, J.G. (1985a), The multivariate hazard construction, Technical report, Department of Mathematics, University of Arizona.

6. Shaked, M. and Shanthikumar, J.G. (1985b), Dynamic construction of random vectors, manuscript in preparation.

END

DTIC

8-86