MANAGEMENT SCIENCE
Vol. 21, No. 10, June 1975
Printed in U.S.A.

ANTITHETIC VARIATES, COMMON RANDOM NUMBERS AND OPTIMAL COMPUTER TIME ALLOCATION IN SIMULATION*

JACK P. C. KLEIJNEN† §

Katholieke Hogeschool, Tilburg, Netherlands

Two simple variance reduction techniques are discussed, viz. antithetic variates and common random numbers. Their joint application creates undesirable negative correlations between the responses of two simulated systems. Therefore three alternatives are considered: antithetics only, common random numbers only, antithetic and common random numbers combined. No alternative is always best as is shown by analytical results for extremely simple systems and simulation results for simple queuing systems. Therefore a procedure is derived that starts with some pilot runs for both systems and estimates which alternative minimizes the variance; at the same time this procedure allocates the limited amount of computer time to the two systems in an optimal way. Results of the application of the procedure to several queuing systems are presented. Because of certain disadvantages of the procedure we may decide to select alternative 1 (antithetics only) a priori. Then the procedure can still be used for the optimal computer time allocation.

1. Introduction

Variance reduction techniques (or briefly VRT) may decrease the variance of the estimated response in simulation experiments through replacement of the crude or "straight on" sampling procedure by a revised procedure. In the literature we can find many VRT; see e.g. [5]. Unfortunately most of these techniques have been devised for the Monte Carlo solution of problems in mathematics and physics (e.g. estimation of integrals, eigenvalues etc.). The management scientist, however, is interested in the simulation of queuing systems, inventory systems etc. The adjustments, required to make the VRT applicable to simulation, can be found in e.g. [7]. The resulting VRT are quite complicated and have hardly been applied in practice. Two techniques, however, remain very simple, viz. antithetic variates and common random numbers (or correlated sampling).

2. Antithetic Variates and Common Random Numbers

In the *antithetic variates* technique one simulation run is generated in the "normal" way from the random numbers $r_1, r_2 \ldots$ but a companion run is generated "antithetically" from the complements of these random numbers, i.e. from $(1 - r_1)$, $(1 - r_2), \ldots$.[1] The purpose of this approach is the creation of *negative correlation* between the responses of the two partner runs. Such correlation decreases the variance

* Processed by Professor Charles H. Kriebel, Departmental Editor for Information Systems and Associate Editor Mark B. Garman; received July 25, 1972, revised July 19, 1973. This paper has been with the author 3 months for revision.

† On leave at Katholieke Hogeschool, Tilburg, Netherlands during December 1973–December 1974.

§ The Fortran programming was done by H. Tilborghs (Katholieke Hogeschool Tilburg) and D. Graham (Duke University). We are indebted to the referees for their comments on two earlier drafts of this paper. The basic idea of our paper was first presented at the European Meeting of IMS, TIMS, ES and IASPS, Amsterdam, 2–7 September, 1968.

[1] The antithetic values $1 - r$ may be directly generated by replacing the starting value, say z_0, by its complement $m - z_0$ in the multiplicative congruential generator $z_i = az_{i-1} \pmod{m}$; see [6].

of the average output of the two runs since

$$\text{var}\{(x_1 + x_2)/2\} = \{\text{var}(x_1) + \text{var}(x_2) + 2\,\text{cov}(x_1, x_2)\}/4 \tag{2.1}$$

where x_1 and x_2 are the output of runs 1 and 2 respectively. If the runs were generated in the usual way (i.e. run 2 were using a sequence of random numbers independent of run 1) then $\text{cov}(x_1, x_2)$ would be zero. Whether antithetics indeed create negative correlation in a complicated simulation cannot be proved. Our intuition tells us that negative correlation may be expected; experiments with various simulated systems of moderate complexity show that such correlation is indeed created (some results will be shown in Tables 4 and 8).

Common random numbers can be utilized when we simulate two (or more) systems and want to compare their mean responses. Using the same sequence of random numbers means that the systems are compared "under the same circumstances" or, statistically speaking, their responses are supposed to show positive correlation. Such correlation is desirable since

$$\text{var}(x - y) = \text{var}(x) + \text{var}(y) - 2\,\text{cov}(x, y) \tag{2.2}$$

where x and y are the response of systems 1 and 2 respectively. This VRT is actually the only technique widely used in practice. Some variance reductions obtained by this technique will be presented later on.

3. The Conflict between Antithetic and Common Random Numbers

Several authors have suggested combining both VRT; compare [2, p. 198], [3, p. 23]. We shall show, however, that joint application of the two techniques does not necessarily give best results. Yet, at first sight such a combination may look quite reasonable. For the difference between the mean responses of two systems is estimated by

$$\bar{d} = \bar{x} - \bar{y} = \textstyle\sum_{i=1}^{M} x_i/M - \sum_{j=1}^{N} y_j/N. \tag{3.1}$$

Hence

$$\begin{aligned}\text{var}(\bar{d}) &= \text{var}(\bar{x}) + \text{var}(\bar{y}) - 2\,\text{cov}(\bar{x}, \bar{y}) \\ &= M^{-2} \textstyle\sum_{i\neq g}^{M} \sum_{i\neq g}^{M} \text{cov}(x_i, x_g) + M^{-1}\sigma_1^2 + N^{-2} \sum_{j\neq h}^{N} \sum_{j\neq h}^{N} \text{cov}(y_j, y_h) \\ &\quad + N^{-1}\sigma_2^2 - 2M^{-1}N^{-1} \textstyle\sum_{i=1}^{M} \sum_{j=1}^{N} \text{cov}(x_i, y_j),\end{aligned} \tag{3.2}$$

where σ_1^2 (σ_2^2) is the variance of a run with system 1 (system 2). The covariances in (3.2) are determined by our choice of the random number streams. Antithetic variates reduce the variance of the average response of a particular system, i.e. $\text{var}(\bar{x})$ and $\text{var}(\bar{y})$ are decreased; common random numbers are supposed to create a positive covariance between $\bar{x}$ and $\bar{y}$. But let us consider the joint application of both techniques in more detail, using Table 1. The columns (2) and (4) of this table show that systems 1 and 2 are simulated with antithetics and we suppose that this technique indeed creates the desirable negative correlation between x_1 and x_2, x_3 and x_4, etc. and between y_1 and y_2, y_3 and y_4, etc. Looking at a particular row we see that the two systems use common random numbers. Therefore we suppose that there is positive correlation between x_1 and y_1, x_2 and y_2, etc. However Table 1 also shows *negative* correlation between x_1 and y_2, x_2 and y_1, x_3 and y_4, etc. *These negative cross-correlations are undesirable* as (3.2) shows. Therefore we shall next consider three obvious alternatives.

TABLE 1

Joint application of antithetic variates and common random numbers.

Run	System 1		System 2	
	Random numbers★	Response	Random numbers★	Response
(1)	(2)	(3)	(4)	(5)
1	$\mathbf{R}_1$	x_1	$\mathbf{R}_1$	y_1
2	$I - \mathbf{R}_1$	x_2	$I - \mathbf{R}_1$	y_2
3	$\mathbf{R}_2$	x_3	$\mathbf{R}_2$	y_3
4	$I - \mathbf{R}_2$	x_4	$I - \mathbf{R}_2$	y_4
⋮	⋮	⋮	⋮	⋮

★ **R**: vector of random numbers. I: vector of one's.

4. Three Alternative Methods of Variance Reduction

In this section we shall discuss three alternative methods for the generation of correlated runs and the corresponding variances, $\text{var}(\bar{d})$.

(A) *Antithetic variates only*: With this method the random numbers in the columns (2) and (4) of Table 1 are no longer identical but become different.

(B) *Common random numbers only*: Then x_1 and y_1 are generated from $\mathbf{R}_1$, x_2 and y_2 from $\mathbf{R}_2$, etc.[2]

(C) *Joint application of antithetic variates and common random numbers*: This alternative was shown in Table 1.

The derivation of $\text{var}(\bar{d})$ for each alternative is based on (3.2) above. Let c_1 denote the negative covariance created between two responses x of system 1 when using antithetics; c_2 the negative covariance between two responses y of system 2; c_3 the positive cross-covariance between the responses of systems 1 and 2 generated from the same random numbers, c_4 the undesirable negative cross-covariance between the responses of systems 1 and 2. Note that if the number of runs for a system is odd, then the last run cannot be generated antithetically, and if $M \neq N$ then we cannot match all runs of system 1 with those of system 2.

We illustrate the derivation of $\text{var}(\bar{d})$ by considering the situation where only antithetics are applied and $M = N =$ even. The first summation term in (3.2) then reduces to

$$(4.1)\quad M^{-2}\{2\,\text{cov}(x_1, x_2) + 2\,\text{cov}(x_3, x_4) + \cdots + 2\,\text{cov}(x_{M-1}, x_M)\} = M^{-2}\left\{2\,\frac{M}{2}\,c_1\right\} = M^{-1}c_1 .$$

In the same way we find that the third term in (3.2) reduces to $N^{-1}c_2$. Since no correlation exists between runs of different systems the last term in (3.2) vanishes. Hence

$$(4.2)\quad \text{var}(\bar{d}) = M^{-1}c_1 + M^{-1}\sigma_1^2 + N^{-1}c_2 + N^{-1}\sigma_2^2 \quad (\text{Method A, } M = N \text{ even}).$$

The derivation of $\text{var}(\bar{d})$ for the other situations is analogous. Because (3.2) always

[2] If $N \neq M$ then we may combine common random numbers with antithetics in the last, noncommon runs of the system that is simulated most often, for these last runs do not create negative cross-correlations; see [6].

contains the two terms $(M^{-1}\sigma_1^2 + N^{-1}\sigma_2^2)$ these terms are not shown in Table 2. To save space Table 2 shows $\text{var}(\bar{d})$ only for method (C). For method (A) we obtain $\text{var}(\bar{d})$ by putting $c_3 = c_4 = 0$ (i.e. delete the last term in Table 2); for method (B) put $c_1 = c_2 = c_4 = 0$ (so we obtain $-2c_3/M$). Table 2 does not show $M < N$ but the results for this case are easily obtained by interpreting c_1 as the covariance among the responses (y) of the system that is run most often, etc.

5. Comparisons among Alternatives

From Table 2 it follows that in order to determine which of the three alternatives (A), (B), (C) gives the lowest variance, we need to know the *relative magnitudes of the covariances* c_1 through c_4. For instance, if $c_3 > | c_4 |$ then method (C) is better than (A); method (A) is better than (B) if $c_3 < | c_1 + c_2 |/2$ (and $M = N =$ even); method (B) is better than (C) if $| c_4 | > | c_1 + c_2 |/2$ (and $M = N =$ even), etc. We would like to know whether these inequalities hold in *general* in the simulation of systems.

We first consider some very simple "systems" and by counterexample we prove that no inequality holds for *all* systems. In the "systems" under consideration, the response x depends on a single random number r and is a monotonic increasing function of r, e.g. $x = r^2$. Then the antithetic run is $x^\star = (1 - r)^2$ so that

$$
\begin{aligned}
c_1 = \text{cov}(x, x^\star) &= E[r^2(1 - r)^2] - E(r^2)E[(1 - r)^2] \\
&= \int_0^1 (r^2 - 2r^3 + r^4)\, dr - \int_0^1 r^2\, dr \int_0^1 (1 - 2r + r^2)\, dr \qquad (5.1) \\
&= 1/30 - (1/3)(1/3) = -7/90.
\end{aligned}
$$

Other simple "systems" are specified in the columns (1) and (2) of Table 3; the resulting covariances have been calculated analogous to (5.1) and are given in the columns (3) through (6); the columns (7) through (9) show the comparisons among covariances suggested above. These comparisons prove that no inequality holds for all systems, i.e. *none of the methods* (A), (B) *or* (C) *is best in all circumstances.*

The counterexample in Table 3 actually proves our point, viz. neither method is always best. For illustratory purposes we give *experimental* results for some more *complicated* systems. As Table 4 shows we simulated single-server queuing systems (variant a: the two systems have exponential interarrival and service distributions; variant b: system 2 has constant service times) and four-servers-in-sequence systems (systems 1 and 2 have different parameters for their exponential service time dis-

TABLE 2

The variance of the estimated difference between the responses of systems 1 and 2 (to each entry the common term $(M^{-1}\sigma_1^2 + N^{-1}\sigma_2^2)$ *should be added).*

Case.	C: joint application of antithetic variates and common random numbers.
I $M = N =$ even	$M^{-1}c_1 + N^{-1}c_2 - 2M^{-1}(c_3 + c_4)$
II $M = N =$ odd	$(M^{-1} - M^{-2})c_1 + (N^{-1} - N^{-2})c_2 - 2M^{-1}\{c_3 + (1 - M^{-1})c_4\}$
III M(even) $> N$(even)	$M^{-1}c_1 + N^{-1}c_2 - 2M^{-1}(c_3 + c_4)$
IV M(even $> N$(odd)	$M^{-1}c_1 + (N^{-1} - N^{-2})c_2 - 2M^{-1}(c_3 + c_4)$
V M(odd) $> N$(even)	$(M^{-1} - M^{-2})c_1 + N^{-1}c_2 - 2M^{-1}(c_3 + c_4)$
VI M(odd) $> N$(odd)	$(M^{-1} - M^{-2})c_1 + (N^{-1} - N^{-2})c_2 - 2M^{-1}(c_3 + c_4)$

TABLE 3

The relative magnitudes of the covariances in some simple systems.

Systems		Covariances				Comparisons		
$x = x(r)$	$y = y(r)$	c_1	c_2	c_3	c_4	$\frac{\lvert c_1 + c_2 \rvert}{2c_3}$	$\frac{\lvert c_1 + c_2 \rvert}{2 \lvert c_4 \rvert}$	$\frac{c_3}{\lvert c_4 \rvert}$
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
r^2	$r^2/2$	$-7/90$	$-7/360$	$8/180$	$-7/180$	>1	>1	>1
$2r + 5$	$r^2 + r$	$-1/3$	$-59/180$	$1/3$	$-1/3$	<1	<1	
$2r^2$	$\sqrt{r}$			$8/63$	$-132/905$			<1

tributions).[3] For the single-server systems we took $M = N = 50$ and for the four-servers systems $M = N = 20$. To these systems we also applied the crude method ("method D") where all runs are independent. Since the resulting $\hat{\text{var}}(\bar{d})$ are stochastic variables the values and rankings of Table 4 are subjected to sampling errors. An exact statistical analysis would require multiple comparison procedures to determine simultaneous confidence bands for the comparisons among the estimated variances, or—better—multiple ranking procedures to determine a reliable ranking of the variances; see [6]. A crude test of the hypothesis that method (C) is not the best technique, compares $\hat{\text{var}}(\bar{d})$ for (C) and (B) in the four-servers case (least favorable comparison for C). Since $\hat{\text{var}}(\bar{d})$ is based on only ten independent observations for (C) (observation 1 is $d_1 = (x_1 + x_2)/2 - (y_1 + y_2)/2$) we have

$$F_{9,19} = \frac{\hat{\text{var}}(\bar{d} \mid C)}{\hat{\text{var}}(\bar{d} \mid B)} = \frac{32}{19} = 1.68 \tag{5.2}$$

which is significant at the 25% point ($F_{9,19}^{0.25} = 1.41$) but not at the 10% point of the F-statistic ($F_{9,19}^{0.10} = 1.98$). So the point estimates of Table 4 can yield only preliminary conclusions which, however, do not seem to contradict the results from Table 3, where for simple systems it was shown analytically that (C) is not necessarily best. Our preliminary ranking is (B), (C), (A), (D) for the four-servers and (C), (A), (B), (D) for the single-server systems. Andréasson found the ranking (C), (B), (A), (D) for

TABLE 4

The estimated variance of the estimated difference between the responses of two systems.

Systems	Methods			
	A (antithetics)	B (common)	C (joint)	D (independent)
1. Single-server				
a. Exponential	0.057	0.11	0.053	0.12
b. Exp./Constant	0.090	0.18	0.081	0.33
2. Four-stations	0.022	0.0019	0.0032	0.037

[3] More specific: Case 1a: system 1 has parameters $\lambda_1 = 1.5$ and $\mu_1 = 2$, system 2 has $\lambda_2 = 0.5$ and $\mu_2 = 2.5$. Case 1b: $\lambda_1 = 1.5$ and $\mu_1 = 2$, $\lambda_2 = 1.5$ and service times 0.5. Case 2: the two systems are plans I and II in Table 1 of [8]. We did not simulate systems as complex as met in practice, since such systems would require much modeling and running time and yet serve only illustration purposes.

his simulated multichannel systems (no statistically tested ranking; see [1, p. 19]). Tocher (private communication) reported that (C) performed poorer than (B).

6. Optimum Alternative and Computer Time Allocation

Since no alternative is best for *all* systems, we may try to estimate which alternative gives best results for the two systems we actually want to simulate. Therefore we may generate some pilot runs for these two systems and estimate the variances σ_1^2 and σ_2^2 and the covariances c_1 through c_4 to select the alternative that will be applied in the remaining runs. (In §7 we shall consider the estimation procedure for the variances and covariances in detail.) If available computer time were unlimited then we could take $M = N = \infty$ so that $\text{var}(\bar{d}) = 0$ for all three alternatives and the choice among them would be indifferent. Therefore limited total computer time is taken into account. After the pilot phase in which N_p runs of both systems 1 and 2 are generated, we determine M and N such that the remaining $(M - N_p)$ and $(N - N_p)$ runs consume the total computer time T. Hence if t_1 and t_2 denote the computer time per run of systems 1 and 2 we should satisfy (6.1) and (6.2).

$$t_1 M + t_2 N = T, \tag{6.1}$$

$$M \geqq N_p, \quad N \geqq N_p. \tag{6.2}$$

We want to select M and N such that $\text{var}(\bar{d})$ is minimized under the restrictions (6.1) and (6.2). Unfortunately Table 2 showed that the formula for $\text{var}(\bar{d})$ varies with M and N, and with (A), (B), (C). For the sake of simplicity we introduce an approximate formula for $\text{var}(\bar{d})$ dropping terms in M^{-2} and N^{-2} in Table 2. This results in Table 5.

From Table 5 we see that $\text{var}(\bar{d})$ can be approximated by

$$\text{var}(\bar{d}) = a_1/M + a_2/N \tag{6.3}$$

the values of the coefficients a_1 and a_2 varying with the methods (A), (B), (C) and with the cases $M \geqq N$, $M \leqq N$.

The optimum values of M and N (denoted by M_0 and N_0) depend on the signs of the coefficients a_1 and a_2 :

(i) If one of these coefficients is *negative* (two negative coefficients would yield a negative variance) then

$$M_0 = N_0 = T/(t_1 + t_2). \tag{6.4}$$

For, suppose a_1 is negative. Then N should be taken as large as possible. However, it can be shown that negative a_1 implies $M \geqq N$. Hence (6.4) must hold.

TABLE 5

Approximation for $\text{var}(\bar{d})$ *(based on dropping terms in* M^{-2} *and* N^{-2}*)*

Case	Method	$\text{var}(\bar{d})$
$M \geqq N$	A	$(\sigma_1^2 + c_1)M^{-1} + (\sigma_2^2 + c_2)N^{-1}$
	B	$(\sigma_1^2 - 2c_3)M^{-1} + \sigma_2^2 N^{-1}$
	C	$(\sigma_1^2 + c_1 - 2c_3 - 2c_4)M^{-1} + (\sigma_2^2 + c_2)N^{-1}$
$M \leqq N$	A	$(\sigma_1^2 + c_1)M^{-1} + (\sigma_2^2 + c_2)N^{-1}$
	B	$\sigma_1^2 M^{-1} + (\sigma_2^2 - 2c_3)N^{-1}$
	C	$(\sigma_1^2 + c_1)M^{-1} + (\sigma_2^2 + c_2 - 2c_3 - 2c_4)N^{-1}$

TABLE 6

Optimal values of M *and* N *if* a_1 *and* a_2 *are positive* ($M^\star$ *and* $N^\star$ *specified in* (6.5) *and* (6.6)).

Case	Side-condition	M_0	N_0
$M \geqq N$	$N_p \leqq N^\star \leqq T/(t_1 + t_2)$	$M^\star$	$N^\star$
	$N^\star < N_p$	$(T - t_2 N_p)/t_1$	N_p
	$N^\star > T/(t_1 + t_2)$	$T/(t_1 + t_2)$	$T/(t_1 + t_2)$
$M \leqq N$	$N_p \leqq M^\star \leqq T/(t_1 + t_2)$	$M^\star$	$N^\star$
	$M^\star < N_p$	N_p	$(T - t_1 N_p)/t_2$
	$M^\star > T/(t_1 + t_2)$	$T/(t_1 + t_2)$	$T/(t_1 + t_2)$

(ii) If both coefficients are positive then minimization of (6.3) together with (6.1) yields

$$M^\star = T/\{t_1 + (a_1^{-1} a_2 t_1 t_2)^{1/2}\}, \tag{6.5}$$

$$N^\star = T/\{t_2 + (a_1 a_2^{-1} t_1 t_2)^{1/2}\} \tag{6.6}$$

This solution may violate (6.2). Moreover, when we use particular values for a_1 and a_2 then these values imply either $M \leqq N$ or $M \geqq N$; see Table 5. Therefore M_0 and N_0 should be taken from Table 6. This table shows that if (6.6) yields $N^\star < N_p$ then N_0 is set equal to the pilot number N_p; if (6.6) results in $N^\star > T/(t_1 + t_2)$ violating $M \geqq N$ then an equal number of runs $M_0 = N_0 = T/(t_1 + t_2)$ is taken.

There is one more restriction not mentioned yet, viz. M and N must be integer. Therefore we examine the integer points in the neighbourhood of M_0 and N_0. We have to check if these integer pairs still satisfy all restrictions, i.e. (6.1), (6.2) and either $M \geqq N$ or $M \leqq N$. (If we find more than one admissible pair we select the pair yielding the smallest variance. Obviously in (6.1) we replace the equal sign by $\leqq$.)

Using (6.4) or Table 6 we have found M_0 and N_0. However these values were determined for a particular pair of values of the coefficients a_1 and a_2 in (6.3). Table 5 showed that a different pair holds for each alternative and for each case ($M \geqq N$, $M \leqq N$), together six pairs. Actually method (A) gives the same coefficients for $M \geqq N$ and $M \leqq N$ so that we have five instead of six different pairs. Each of the five pairs (a_1, a_2) gives a corresponding pair (M_0, N_0). Substituting M_0 and N_0 into (6.3) yields the minimum variance. In this way five minimum variances are found. Finally we select the *minimum* among these five minimal variances and determine the corresponding method ((A), (B) or (C)) and number of runs (M_0 and N_0).

Summarizing this section, our selection procedure takes into account the *covariances* created by the various alternatives as these covariances determine the coefficients a_1 and a_2. Moreover we can incorporate possible differences among the *computer times per run* as we can take values for t_1 and t_2 in (6.1) varying with (A), (B) and (C). The variance is minimized not only through selection of a suitable *variance reducing method* but also through an *optimal* combination of the *number of runs* per system. Note that once we have decided on the optimal number of runs per system we may obtain a more accurate estimate of $\text{var}(\bar{d})$ using the exact formulas in Table 2.

7. Estimation of the Coefficients in the Optimization Procedure

In this section we shall discuss how the pilot runs can be used to obtain estimates for the coefficients a_1, a_2, t_1 and t_2. The N_p pilot runs of the systems 1 and 2 should be generated using method (C) since this method creates all four covariances c_1 through c_4. We can estimate c_1, the covariance between antithetic runs of system 1, from

$$(7.1) \qquad \hat{c}_1 = \sum_{i=1}^{n_p} (x_{2i-1} - \bar{x})(x_{2i} - \bar{x}^\star)/(n_p - 1) \quad (n_p = N_p/2, N_p \text{ even})$$

where the x_{2i} are the n_p antithetic runs with average $\bar{x}^\star$, and the x_{2i-1} are the n_p normal runs with average $\bar{x}$. The covariance c_2 is estimated analogously. When estimating c_3, the positive cross- covariance between x_i and y_i $(i = 1, \ldots, N_p)$ we have to remember that e.g. the pair (x_1, y_1) is not independent of (x_2, y_2); see Table 1. Therefore we divide the N_p pilot runs into two groups as shown by Table 7, and estimate c_3 from (7.2) through (7.4).

$$(7.2) \qquad \hat{c}_3(1) = \sum_{i=1}^{n_p} (x_{2i-1} - \bar{x})(y_{2i-1} - \bar{y})/(n_p - 1),$$

$$(7.3) \qquad \hat{c}_3(2) = \sum_{i=1}^{n_p} (x_{2i} - \bar{x}^\star)(y_{2i} - \bar{y}^\star)/(n_p - 1),$$

$$(7.4) \qquad \hat{c}_3 = \{\hat{c}_3(1) + \hat{c}_3(2)\}/2.$$

Such grouping can also be used to estimate c_4, and the variances σ_1^2 and σ_2^2. The estimates of the variances and covariances are substituted into Table 5 to obtain estimates of a_1 and a_2 for (A), (B), (C) and $M \geqq N$ or $M \leqq N$.

If it takes T_p units of time to run the first system N_p times applying method (C), then we can estimate t_1 by

$$(7.5) \qquad \hat{t}_1 = T_p/N_p.$$

(In systems with a stochastic runlength T_p will be stochastic.) In the same way we can estimate t_2. Since the extra computer time for antithetic or common random numbers is usually negligible, we may decide to use the same $\hat{t}_1$ and $\hat{t}_2$ for (A), (B) and (C).

Notice that the optimum values of M, N and $\text{var}(\bar{d})$ are *nonlinear* functions of the variances, covariances and times per run. Hence, when using unbiased estimators for these variances etc., the corresponding estimated optimum M, N and $\text{var}(\bar{d})$ are still *biased.* This type of bias was studied by Fishman in [3]. We have not tried to determine this bias in our problem but we conjecture that it is of negligible importance; see also [4]. If this bias would not be negligible then the variance reduction might be less than maximal since we might not select the best VRT ((A), (B), (C)) and might not allocate computer time optimally (M_0, N_0). Note, however, that the estimator of $\bar{d}$ remains unbiased since the pilot runs with method (C) yield unbiased estimators

TABLE 7

Grouping N_p *pilot runs for the estimation of the covariance* c_3.

Group 1			Group 2		
Random numbers	Response of system		Random numbers	Response of system	
	1	2		1	2
$\mathbf{R}_1$	x_1	y_1	$I - \mathbf{R}_1$	x_2	y_2
$\mathbf{R}_2$	x_3	y_3	$I - \mathbf{R}_2$	x_4	y_4
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
$\mathbf{R}_{n_p}$	x_{N_p-1}	y_{N_p-1}	$I - \mathbf{R}_{n_p}$	x_{N_p}	y_{N_p}

and so do the remaining runs (independent of the pilot runs) using one of the methods (A), (B), (C).[4]

8. Applications and Comments

The optimization procedure was applied to the estimation of the difference in mean waiting time in several simple queuing systems; see Table 8. The optimum M and N were found to differ greatly from each other (unless a_1 is negative or $N^\star > T(t_1 + t_2)$). So taking an equal number of runs per system may be very suboptimal. Given an optimal choice of M and N we found point estimates of the variances suggesting that C is best and D is worst.

Our optimization procedure has the following disadvantages.

(i) It takes *time* to estimate the coefficients and to perform the necessary calculations with these coefficients. Nevertheless in a complicated simulation study this extra time is negligible and therefore the optimization may be worthwhile.

(ii) The procedure is based on *estimates* of the variances and covariances calculated from the N_p pilot runs. Hence if N_p increases then these estimates become more reliable and a more reliable selection from the methods (A), (B), (C) is possible. Unfortunately, if we augment N_p and we find that the best method is not (C) (the method applied in the pilot phase) but either (A) or (B), then most of the runs have already been generated with the inferior method (C)! To solve this dilemma we may decide to restrict N_p to, say, 10% of our a priori guess of M_0 and N_0.[5]

(iii) If, after the pilot-phase, we decide to switch from (C) to either (A) or (B), then we get "*nonhomogeneous*" output, as switching to (A) means that the cross-covariances

TABLE 8

Results of the application of the optimization procedure.

Systems compared	Minimal $\hat{\text{var}}(\bar{d}) \times 10^5$				Optimum M and N for method (C)★		Comment on M and N
	(A)	(B)	(C)	(D)			
1. Single-server systems							
a. Exponential distributions	162	76	52	257	100	100	$N^\star > T/(t_1 + t_2)$
b. Exponential distributions+	83	99	77	104	139	50	$N^\star < N_p$
c. Constant service times	9	3	2	14	150	150	a_1 negative
d. One exponential, one constant service time	114	145	113	180	164	54	
2. One single-server and one two-servers system	33	26	24	51	128	68	
3. Two two-servers systems	61	45	39	79	95	94	a_1 negative

+ 1a and 1b use different computer programs.

★ For the methods (A, (B) and (D) approximately the same values were found.

[4] More technically: $\bar{d} = w\,\bar{d}_1 + (1 - w)\bar{d}_2$ where $\bar{d}_1$ and $\bar{d}_2$ are the estimators from stages 1 and 2. In the pilot phase (C) yields an unbiased estimator $\bar{d}_1$. After the pilot phase either method (A), or (B) or (C) will be used to obtain $\bar{d}_2$ depending on their variances, say $\hat{\sigma}_i^2$ ($i = 1, 2, 3$), estimated in the pilot phase. Then $E(\bar{d}_2) = E[E(\bar{d}_2|\hat{\sigma}_i^2)]$ with $E(\bar{d}_2|\hat{\sigma}_i^2)$ being unbiased since the observations in phase 2 are statistically independent of phase 1.

[5] Actually we should not compare var($\bar{d}$) when applying (A), (B) or (C) in all M_0 and N_0 runs but instead var($\bar{d}$) when applying (A), (B) or (C) in ($M_0 - N_p$) and ($N_0 - N_p$) runs and (C) in N_p runs. When N_p is small compared with M_0 and N_0, as it should be, then for the sake of simplicity we may use the procedure of §6.

TABLE 9

Optimal values of M and N when applying antithetic variates only.

Side-conditions	M_0	N_0
$M^\star \geqq N_p$, $N^\star \geqq N_p$	$M^\star$	$N^\star$
$M^\star < N_p$	N_p	$(T - t_1N_p)/t_2$
$N^\star < N_p$	$(T - t_2N_p)/t_1$	N_p

c_3 and c_4 become zero, and switching to (B) means that c_1, c_2 and c_4 become zero. This complicates the *statistical analysis*. In [6] we show how we can still estimate var$(\bar{d})$ from formulas based on (3.2) (but differing from Table 2 where in all runs a single method is applied); the resulting confidence intervals hold only approximately. We also refer to [6] for a discussion of minimizing computer time subject to a fixed variance, and for generalizations to k ($\geqq 2$) systems.

Because of the above disadvantages we may decide not to use the procedure to select one of the methods (A), (B), or (C). Instead we may a priori choose *method* (A), i.e. the method that does not complicate the statistical analysis of the simulation results (taking the average of each antithetic pair gives independent observations). Our procedure remains useful for the *optimal* computer time allocation. Since the coefficients a_1 and a_2 are always positive for (A), and we need not distinguish between $M \geqq N$ and $M \leqq N$ in Table 5, we can replace Table 6 by Table 9.

References

1. Andréasson, I. J., "The Application of Two Variance-Reducing Techniques on a Queuing Simulation in SIMULA-67," Report NA 70.21, Department of Information Processing, The Royal Institute of Technology, Stockholm, 1970.
2. Emshoff, J. R. and Sisson, R. L., *Design and Use of Computer Simulation Models*, The MacMillan Company, New York, second printing, 1971.
3. Fishman, G. S., "Digital Computer Simulation: The Allocation of Computer Time in Comparing Simulation Experiments," RM-5288-1-PR, The RAND Corporation, Santa Monica, California, Oct. 1967. (Also published in *Operations Research*, Vol. 16, no. 2 (March–April 1968), pp. 280–295, erratum in Vol. 16, no. 5 (September–October 1968), p. 1087.)
4. Ghurye, S. G. and Robbins, H., "Two-Stage Procedures for Estimating the Difference between Means," *Biometrika*, Vol. 41, parts 1 and 2 (June 1954), pp. 146–152.
5. Halton, J. H., "A Retrospective and Prospective Survey of the Monte Carlo Method," *Siam Review*, Vol. 12, No. 1 (January 1970), pp. 1–63.
6. Kleijnen, J. P. C., *Statistical techniques in simulation, Parts I and II*, to be published by Marcel Dekker Inc., New York, ± Sept. 1974.
7. Moy, W. A., "Practical Variance-Reducing Procedures for Monte Carlo Simulations," in: *The Design of Computer Simulation Experiments*, edited by T. H. Naylor, Duke University Press, Durham, 1969.
8. Naylor, T. H., Wertz, K. and Wonnacott, T. H., "Methods for Analyzing Data from Computer Simulation Experiments," *Communications of the ACM*, Vol. 10 (November 1967), pp. 703–710.

LINKED CITATIONS

You have printed the following article:

Antithetic Variates, Common Random Numbers and Optimal Computer Time Allocation in Simulation
Jack P. C. Kleijnen
Management Science, Vol. 21, No. 10, Application Series. (Jun., 1975), pp. 1176-1185.
Stable URL:
http://links.jstor.org/sici?sici=0025-1909%28197506%2921%3A10%3C1176%3AAVCRNA%3E2.0.CO%3B2-3

References

[3] **The Allocation of Computer Time in Comparing Simulation Experiments**
George S. Fishman
Operations Research, Vol. 16, No. 2. (Mar. - Apr., 1968), pp. 280-295.
Stable URL:
http://links.jstor.org/sici?sici=0030-364X%28196803%2F04%2916%3A2%3C280%3ATAOCTI%3E2.0.CO%3B2-E

[4] **Two-Stage Procedures for Estimating the Difference Between Means**
S. G. Ghurye; Herbert Robbins
Biometrika, Vol. 41, No. 1/2. (Jun., 1954), pp. 146-152.
Stable URL:
http://links.jstor.org/sici?sici=0006-3444%28195406%2941%3A1%2F2%3C146%3ATPFETD%3E2.0.CO%3B2-S

[5] **A Retrospective and Prospective Survey of the Monte Carlo Method**
John H. Halton
SIAM Review, Vol. 12, No. 1. (Jan., 1970), pp. 1-63.
Stable URL:
http://links.jstor.org/sici?sici=0036-1445%28197001%2912%3A1%3C1%3AARAPSO%3E2.0.CO%3B2-X

NOTE: *The reference numbering from the original has been maintained in this citation list.*