*Proceedings of the 1997 Winter Simulation Conference*
*ed. S. Andradóttir, K. J. Healy, D. H. Withers, and B. L. Nelson*

# OPTIMAL QUADRATIC-FORM ESTIMATOR OF THE VARIANCE OF THE SAMPLE MEAN

Wheyming Tina Song

Department of Industrial Engineering
National Tsing Hua University
Hsinchu, Taiwan, REPUBLIC OF CHINA

Neng-Hui Shih

Department of Business Administration
National Ping Tung Institute of Commerce
Ping Tung, Taiwan, REPUBLIC OF CHINA

Mingjian Yuan

Department of Industrial Management
National Yunlin Technology University
Touliu, Taiwan, REPUBLIC OF CHINA

## ABSTRACT

A classical problem of stochastic simulation is how to estimate the variance of the sample mean of dependent but stationary outputs. Many variance estimators, such as the batch means estimators and spectral estimators, can be classified as quadratic-form estimators. Necessary and sufficient conditions on the quadratic-form coefficients such that the corresponding variance estimator has good performance have been proposed. But no one has utilized these conditions to pursue optimal quadratic-form coefficients to form an optimal variance estimator. In this paper, we seek an optimal (minimum variance unbiased) variance estimator by searching for the optimal quadratic-form coefficients.

## 1 INTRODUCTION

Consider estimating the variance of a sample mean $\bar{Y}$ from a sample $\underline{Y} = (Y_1, Y_2, ..., Y_n)$ from covariance-stationary process. Various types of estimators of $\mathrm{Var}(\bar{\mathrm{Y}})$ have been proposed. For example, there are estimators based on classical spectral analysis (Priestly, 1981), spectral-based regression (Heidelberger and Welch, 1981; Damerdji, 1991), regenerative processes (Crane and Iglehart, 1975), ARMA time series (Schriber and Andrews, 1984, Yuan and Nelson 1994), standardized time series (Schruben, 1983; Goldsman, Meketon, and Schruben, 1990; Glynn and Iglehart, 1990; Muñoz and Glynn, 1991) batch means (Schmeiser, 1982; Meketon and Schmeiser, 1984; Welch, 1987; Fox, Goldsman, and Swain, 1991; Glynn and Whitt, 1991; Bischak, Kelton, and Pollock, 1993; Fishman and Yarberry, 1993; Pedrosa and Schmeiser, 1994; Chien, 1994; Damerdji, 1994; Song and Schmeiser, 1995; Sherman, 1995; Chien, Goldsman, and Melamed, 1996; Song, 1996; Muñoz and Glynn, 1997) and orthonormally weighted standardized time series area (Foley and Goldsman, 1988)

The batch-means, some spectral-analysis, and some standardized-time-series estimators are linear functions of the cross-products $Y_iY_j$. That is, they can be written as a quadratic-form, $\hat{V} \equiv \underline{Y}^t\mathbf{Q}\underline{Y} = \sum_{i=1}^{n}\sum_{j=1}^{n} q_{ij}Y_iY_j$, where $\mathbf{Q}$ is a constant, symmetric (quadratic-form coefficient) matrix with $(i,j)^{\mathrm{th}}$ entry $q_{ij}$, for $i = 1, 2, ..., n; j = 1, 2, ..., n$.

Necessary and sufficient conditions on the quadratic-form coefficients such that the corresponding variance estimator has good performance have been proposed (Song and Schmeiser, 1993). However, no one has utilized these conditions to pursue optimal quadratic-form coefficients to form an optimal variance estimator. In this paper, we seek an optimal variance estimator by searching for optimal quadratic-form coefficients.

We assume that the data $\underline{Y}$ are from a covariance stationary process with mean $\mu$, variance $R_0^2$, variance-covariance matrix $\Sigma$, and finite kurtosis $\alpha_4 = \mathrm{E}(Y-\mu)^4/R_0^2$. Moreover, we assume that the sum of autocorrelations $\sum_{h=-(n-1)}^{n-1} \rho_h \equiv \sum_{h=-(n-1)}^{n} \mathrm{corr}(Y_i, Y_{i+h})$ converges to a finite limit $\gamma_0$ as $n$ goes to infinity.

In Section 2, we review some properties of the general class of quadratic-form estimators. In Section 3, we introduce a theoretical optimal quadratic-form estimator of $\mathrm{Var}(\bar{\mathrm{Y}})$, discuss its properties, and compare its performance to the overlapping batch means (OBM) estimator with its optimal batch size. In Section 4, we discuss possible ways to estimate the theoretical optimal variance estimator proposed in Section 3. Section 5 is a summary.

## 2 QUADRATIC-FORM VARIANCE ESTIMATORS

We review some necessary and sufficient conditions on the quadratic-form coefficients such that the corresponding estimator of $\text{Var}(\bar{Y})$ satisfies four properties: nonnegativity, location invariance, data reversibility, and smoothness. Expressions for the bias and variance of $\hat{V}$ as functions of the $q_{ij}$ are also included in this section.

**Nonnegativity**: Since $\text{Var}(\bar{Y})$ is always nonnegative, it is reasonable to require an estimator of $\text{Var}(\bar{Y})$ to be nonnegative. By definition, $\hat{V}$ is nonnegative for all data realizations if and only if $\mathbf{Q}$ is positive semi-definite.

**Location Invariance**: An estimator is location invariant if it is not a function of the process location. Location invariance is appealing because $\text{Var}(\bar{\text{X}}) = \text{Var}(\bar{\text{Y}})$, when $X_i = Y_i - d$. If $\hat{V}$ is location invariant, then we can assume without loss of generality that the process mean is zero when studying properties of $\hat{V}$. A necessary and sufficient condition for location invariance is $\sum_{i=1}^{n} q_{ij} = 0, i = 1, 2, ..., n$, or equivalently $\sum_{j=1}^{n} q_{ij} = 0, j = 1, 2, ..., n$.

**Data Reversibility:** Define the reversed sample $\{X_i\}_{i=1}^{n}$ with $X_i = Y_{n-i+1}$. We call the estimator $\hat{V}$ data reversible if $\hat{V}$ has the same value after being applied to both $\underline{Y}$ and $\underline{X}$. If $\hat{V}$ is data reversible, then reversing the quadratic-form coefficients is equivalent to reversing the order of the data. Thus, an estimator is data reversible if and only if $q_{ij} = q_{n-i+1,n-j+1}$ for all $i$ and $j$. When the data are from a covariance-stationary process, data reversibility seems desirable because $R_h \equiv \text{cov}(Y_i, Y_{i+h}) \equiv \text{cov}(X_i, X_{i+h})$ for all $i$ and lags $h$.

**Smoothness**: We define $\hat{V}^{(S)} = \underline{Y}^t\mathbf{Q}^{(S)}\underline{Y}$ to be a smooth estimator of $\text{Var}(\bar{\text{Y}})$ if all coefficients $q_{ij}^{(S)}$ with common lag $h = |i - j|$ are equal. We can smooth any non-smooth estimator to reduce its variance and without increasing its bias. That is, suppose $\hat{V} = \sum_{i=1}^{n}\sum_{j=1}^{n} q_{ij}Y_iY_j$ and consider the corresponding smoothed estimator $\hat{V}^{(S)} = \sum_{i=1}^{n}\sum_{j=1}^{n} q_{ij}^{(S)}Y_iY_j$, where $q_h^{(S)} = (n-h)^{-1}\sum_{\{i,j:|i-j|=h\}} q_{ij}$ for $h = 1, 2, ..., n-1$. Then $\hat{V}^{(S)}$ has the same bias as $\hat{V}$, but smaller variance (Grenander and Rosenblatt, 1957).

**Bias**: Without loss of generality we assume that the data are $p$ dependent; that is $\rho_h = 0$ for $|h| > p$, where possibly $p$ is infinite. The bias of a location-invariant estimator $\hat{V}$, defined as $\text{E}(\hat{V}) - \text{Var}(\bar{Y})$, is

$$\text{bias}(\hat{V}) = b_0R_0 + 2\sum_{h=1}^{p*} b_hR_h \tag{1}$$

where $R_h \equiv \text{Cov}(Y_i, Y_{i+h})$,
$b_h = \sum_{i=1}^{n-h}\{q_{i,i+h} - n^{-1}(1 - \frac{h}{n})\}$,
$h = 0, 1, ..., n-1$ , and $p^* = \min(n-1, p)$.

**Variance**: Let $\{Y_i\}_{i=1}^{n}$ be independent identically distributed (iid) random variables. Then the variance of the location invariant quadratic-form estimator $\hat{V}$ is

$$R_0^2\left[(\alpha_4 - 3)\sum_{i=1}^{n} q_{ii}^2 + 2\sum_{i=1}^{n}\sum_{j=1}^{n} q_{ij}^2\right].$$

Therefore, the variance of $\hat{V}$ for independent identically distributed (iid) normal data is $2R_0^2\sum_{i=1}^{n}\sum_{j=1}^{n} q_{ij}^2$, which is proportional to the sum of all squared quadratic-form coefficients.

## 3 OPTIMAL QUADRATIC-FORM VARIANCE ESTIMATORS

### 3.1 Definition of $\mathbf{Q}^*$

Let $\hat{V} \equiv \underline{Y}^t\mathbf{Q}\underline{Y}$ be any location invariant quadratic-form estimator of $\text{Var}(\bar{Y})$. Therefore, we can assume without loss of generality that $\mu = 0$. The optimal quadratic-form coefficient matrix introduced in this section is obtained by minimizing an upper bound on $\text{Var}(\hat{V})$ provided that $\hat{V}$ is an unbiased estimator of $\text{Var}(\bar{Y})$.

We first review two results. Equation (2) states that the expected value of $\hat{V}$ is equal to the trace of $\mathbf{Q\Sigma}$ while Equation (3) shows that $\phi\text{tr}(\mathbf{Q\Sigma Q\Sigma})$ is an upper bound on $\text{Var}(\hat{V})$. The derivations of Equations (2) and (3) are given in Rao and Kleffe (1988). Specifically,

$$\text{E}(\underline{Y}^t\mathbf{Q}\underline{Y}) = \text{tr}(\mathbf{Q\Sigma}) \tag{2}$$

and

$$\text{Var}(\underline{Y}^t\mathbf{Q}\underline{Y}) \leq \phi\text{tr}(\mathbf{Q\Sigma Q\Sigma}), \tag{3}$$

where $\phi$ is a function of $\text{E}(Y^3)$ and $\text{E}(Y^4)$, but not of $\mathbf{Q}$.

Throughout this paper, we define the optimal estimator of Var($\bar{Y}$), $\hat{V}^{(*)}(\mathbf{Q}^*) \equiv \underline{Y}^t\mathbf{Q}^*\underline{Y}$, to be the minimum variance unbiased estimator for the upper bound on Var($\hat{V}$) in Equation (3). We call $\mathbf{Q}^*$ the optimal Q-F coefficients matrix. That is, $\mathbf{Q}^*$ can be obtained by solving the following problem:

$$\text{(P.1) minimize tr}(\mathbf{Q\Sigma Q\Sigma})$$

subject to:

$$\mathbf{Q}\underline{1} = \underline{0}, \tag{4}$$

$$\text{tr}(\mathbf{Q\Sigma}) = \text{Var}(\bar{Y}) \tag{5}$$

$$\mathbf{Q} \text{ is positive definite,} \tag{6}$$

where $\underline{1} = [1, 1, ..., 1]^t$. Equation (4) enforces the property of location invariance and Equation (5) guarantees unbiasedness. The solution of problem (P.1) is

$$\mathbf{Q}^* = \lambda \mathbf{C}^t\mathbf{\Sigma}^{-1}\mathbf{C}, \tag{7}$$

where

$$\lambda = \frac{\text{tr}(\mathbf{C}^t\mathbf{\Sigma}^{-1}\mathbf{C\Sigma})}{\text{Var}(\bar{Y})} \tag{8}$$

and

$$\mathbf{C} = \mathbf{I} - \underline{1}(\underline{1}^t\mathbf{\Sigma}^{-1}\underline{1})^{-1}\underline{1}^t\mathbf{\Sigma}^{-1}. \tag{9}$$

This result is a direct application of a theorem in Rao (1973), stated in the Appendix.

### 3.2 Comparison with the Optimal OBM Estimator

The OBM estimator is a smooth quadratic-form estimator with many nice properties such as smaller asymptotic variance than all other batch-means estimator while requiring only $O(n)$ computational effort. In this subsection, we use the OBM estimator with the optimal batch size in terms of the mean-squared-error (MSE) as a basis for comparing with the optimal Q-F variance estimator introduced in Section 3.1.

Let $\hat{V}^{(O)}(m^*) = \underline{Y}^t\ \mathbf{Q}^{(O)}\ \underline{Y}$ be the OBM estimator of Var($\bar{Y}$) with the MSE-optimal batch size $m^*$. That is $m^* = \arg\min_m \text{MSE}(\hat{V}^{(O)}(m))$. Table 1 compares $\hat{V}^{(*)}(\mathbf{Q}^*)$ with $\hat{V}^{(O)}(m^*)$ in terms of the bias, variance, and MSE for the first-order autoregressive (AR(1)) process with mean $\mu = 0$, lag-1 correlation $\phi = 0.8182$, and var($Y$) = 5.54885. The variance of the sample mean Var($\bar{Y}$) = 1; the sum of correlations $\gamma_0 = 10$, and the sample size $n = 50$ for this example.

Table 1. AR(1), $\phi = 0.8182$ and $n = 50$

| $\hat{V}$ | Bias | Variance | MSE |
|---|---|---|---|
| $\hat{V}^{(*)}(\mathbf{Q}^*)$ | -.011 (.006) | .038 (.002) | .039 (.002) |
| $\hat{V}^{(O)}(m^*)$ | -.481 (.004) | .150 (.004) | .383 (.003) |

The simulation results shown in Table 1 are based on 50 independent macro-replications. Each involves 50 independent micro-replications, each having sample size $n = 50$. Each macro-replication generates one estimator of the variance of the sample mean $\hat{V}$. One macro-replication generates one bias, variance, and MSE of $\hat{V}$, The standard error of the bias, variance and MSE are reported in the parentheses next to the corresponding estimates. Table 1 shows that the estimator $\hat{V}^{(*)}(\mathbf{Q}^*)$ has smaller bias (in fact zero bias) and smaller variance than the OBM estimator $\hat{V}^{(O)}(m^*)$. The MSE of $\hat{V}^{(*)}(\mathbf{Q}^*)$ is about 10 percent of the MSE of $\hat{V}^{(O)}(m^*)$ for the AR(1) process.

In practice, we are not able to obtain $\mathbf{Q}^*$ since it depends on the unknown parameter Var($\bar{Y}$) (see Equation (5)). But the huge MSE reduction encourages us to further investigate the estimator $\hat{V}^{(*)}(\mathbf{Q}^*)$.

### 3.3 Viewing $\mathbf{Q}^*$ Graphically

We consider three processes: (1) AR(1) as used in Section 3.2, (2) the second-order autoregressive AR(2) process, and (3) M/M/1-queue-wait-time (M/M/1-QWT) process. The parameters of these three processes are selected as follows: the mean $\mu = 0$; the variance of the sample mean Var($\bar{Y}$) = 1; the sum of correlations $\gamma_0 = 10$, and the sample size $n = 50$. Applying the results in (P.1), we derive the optimal Q-F coefficients matrix $\mathbf{Q}^*$ and present them in three-dimensional plots as a function of $i$ and $j$.

The three-dimensional plots of $\mathbf{Q}^*$ for AR(1) and M/M/1 are almost identical: the main-diagonal terms are positive, the first off-diagonal terms are negative, and the other terms are negligible. This pattern remains the same for AR(1) and M/M/1 processes for a broad range of parameters except those cases where $\gamma_0 \simeq 1$, which is close to the iid process. Figure 1 contains the three-dimensional plot of $\mathbf{Q}^*$ for AR(1). It can be seen from the plot that $\hat{V}^{(*)}(\mathbf{Q}^*)$ satisfies the four properties: nonnegativity, location invariance, reversibility and smoothness.

Since the main and first off-diagonal terms of $\mathbf{Q}^*$ play an important role in $\mathbf{Q}^*$, the ratio of the $q^*_{ii}$ to $q^*_{i,i+1}$ seems to be an important summary quantity of $\mathbf{Q}^*$. The ratio increases as $\gamma_0$ increases and converges to -2 as $\gamma_0 \to \infty$. The ratio approaches -2 at about $\gamma_0 = 10$. Figure 2 shows the ratio of $q^*_{ii}$ to $q^*_{i,i+1}$ versus $\gamma_0$ for AR(1). The analogous plot for M/M/1 is almost identical to Figure 2.

The three-dimensional plot of $\mathbf{Q}^*$ for AR(2) has the main-diagonal terms positive, the first off- and second off-diagonal terms negative, and the other terms are negligible. Figure 3 is the three-dimensional plot of $\mathbf{Q}^*$ for AR(2). We observe that the ratio $q^*_{ii}$ / $(q^*_{i,i+1} + q^*_{i,i+2})$ converges to -2. Again, $\hat{V}^{(*)}(\mathbf{Q}^*)$ satisfies the four properties: nonnegativity, location invariance, data reversibility and smoothness.

The three-dimensional plots of $\mathbf{Q}^*$ for AR(1) and AR(2) differ from that for OBM estimator, in which the $q_{ij}$ linearly decreases to zero as $|i-j|$ increases for $0 < |i-j| < m$. Figure 4 shows the three-dimensional plot of the quadratic-form coefficient $q_{ij}$ for the OBM estimator for $n = 50$ with batch size $m = 10$. One reason to explain the difference between the plots is in the bias expression. For $\hat{V}^{(*)}(\mathbf{Q}^*)$ to have zero bias, we have observed that $b_0R_0 = -2\sum_{h=1}^{p*} b_hR_h$ (see Equation (1)). For OBM to have low bias, $b_h$ must be close to 0 for all $h$. Specifically, the sum of the main diagonal should be $1/n$, with each successive off-diagonal sum decreasing to $n^{-1}(1-|h|/n))$ for all lags $|h|$ whose autocorrelation is nonzero. Thus, for OBM to have low bias requires a wide ridge when the data process has autocorrelation extending over many lags. For iid data, $\hat{V}^{(*)}(\mathbf{Q}^*) = \hat{V}^{(O)}(m = 1)$.

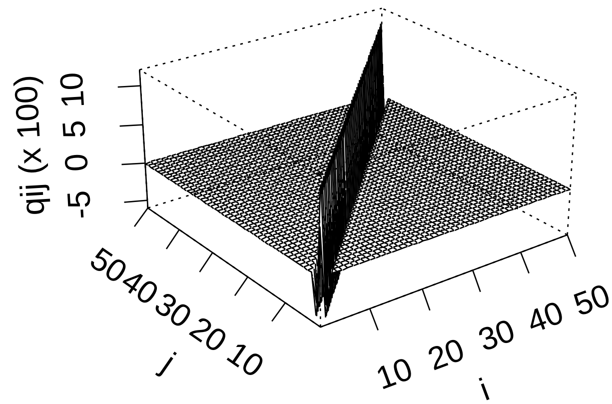


Figure 1: Three-dimensional Plot of $\mathbf{Q}^*$ for AR(1)

## 4 ESTIMATING THE OPTIMAL QUADRATIC-FORM COEFFICIENTS

We now investigate the statistical performance obtained when we use the data $\underline{Y}$ to estimate $\mathbf{Q}^*$. Let $\hat{\mathbf{Q}}^* = [(\hat{q}^*_{ij})]$ be the estimator of $\mathbf{Q}^*$. To obtain a particular method that is computationally reasonable, we assume that only the main-diagonal and first off-diagonal terms of $\mathbf{Q}^*$ are non-zero. That is,

$$\hat{q}^*_{ij} = 0 \ \text{ for } |i-j| \geq 2. \tag{10}$$

This structure is appropriate for AR(1) and M/M/1 processes.

We now define the main and the first off-diagonal terms. We assume that the output data has $\gamma_0 > 10$, so we can apply the result shown in Figure 2 that the ratio of the $q^*_{ii}$ to $q^*_{i,i+1}$ approaches -2. That is, we set

$$\hat{q}^*_{i,i+1} = \hat{q}^*_{i+1,i} = -\hat{q}^*_{22}/2 \tag{11}$$

for $i = 2, 3, ..., n-1$. To satisfy the invariance property, we set

$$\hat{q}^*_{11} = \hat{q}^*_{nn} = \hat{q}^*_{22}/2, \tag{12}$$

so that $\sum_{i=1}^{n} \hat{q}^*_{ij} = 0$ for $j = 1, 2, ..., n$. To satisfy unbiasedness, we plug Equations (10) to (12) into Equation (1) to enforce $\text{bias}(\hat{V}) = 0$ and obtain

$$\hat{q}^*_{ii} = n^{-1}(n-1)^{-1}(1-\hat{\rho}_1)^{-1}\frac{\hat{\text{Var}}(\bar{Y})}{\hat{R}_0} \tag{13}$$

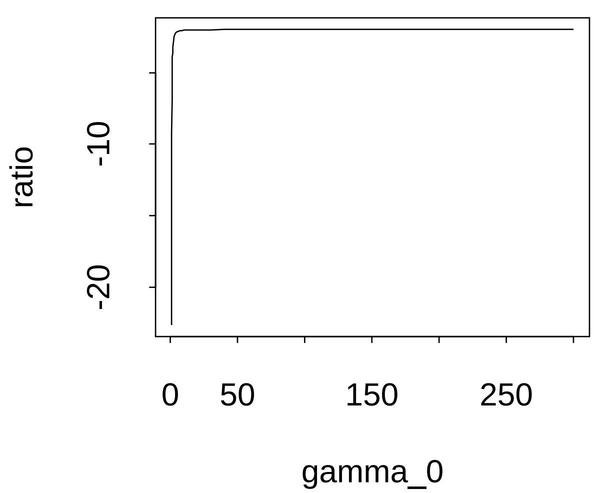


Figure 2: The Ratio of $q_{22}$ to $q_{12}$

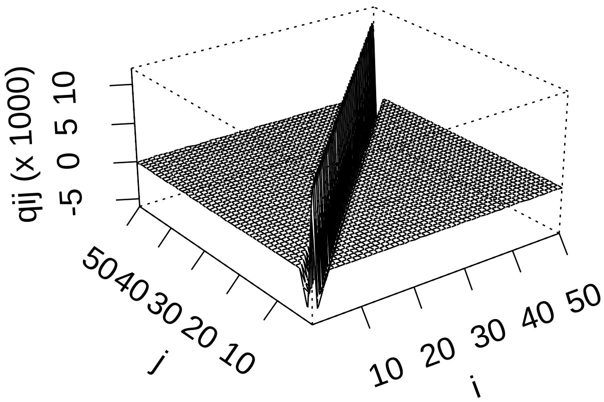


Figure 3: Three-dimensional Plot of $\mathbf{Q}^*$ for AR(2)

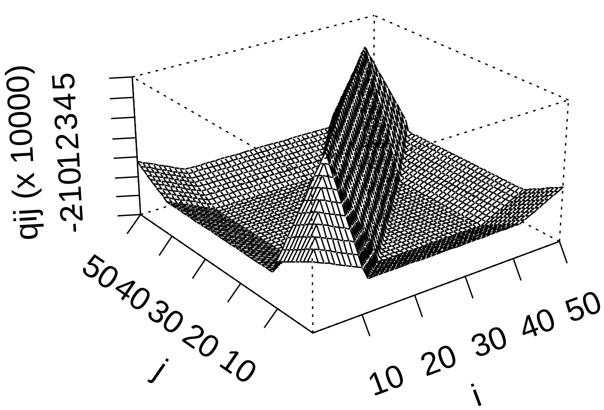


Figure 4: Three-dimensional Plot of $\mathbf{Q}^{(O)}$ for OBM

for $i = 2, 3, ..., n-1$, where $\hat{\rho}_1$, $\hat{R}_0$, and $\hat{\text{Var}}(\bar{Y})$ denote the estimators of the unknown parameters $\rho_1$, $R_0$, and $\text{Var}(\bar{Y})$, respectively. We will define these estimators in the next paragraph. In this setting, $\hat{V}^{(*)}(\hat{\mathbf{Q}}^*) = (\hat{q}^*_{22}/2)\sum_{i=1}^{n-1}(Y_i - Y_{i+1})^2$, where $\hat{q}^*_{22}$ is defined in Equation (13).

In the empirical study, we define $\hat{R}_0 \equiv (n-1)^{-1}\sum_{i=1}^{n}(Y_i - \bar{Y})^2$, $\hat{\rho}_1 \equiv \sum_{i=1}^{n-1}(Y_i - \bar{Y})(Y_{i+1} - \bar{Y})/\sum_{i=1}^{n}(Y_i - \bar{Y})^2$, $\hat{\text{Var}}(\bar{Y}) \equiv \hat{V}^{(O)}(m^{1-2-1})$, which is the OBM estimator using Pedrosa and Schmeiser's 1-2-1 OBM batch size (Pedrosa and Schmeiser, 1994).

We now compare four different estimators of $\hat{\text{Var}}(\bar{Y})$: $\hat{V}^{(*)}(\hat{\mathbf{Q}}^*)$, $\hat{V}^{(O)}(m^*)$, $\hat{V}^{(O)}(m^{1-2-1})$, and $\hat{V}^{(O)}(m^S)$. The first estimator is the estimated optimal Q-F estimator proposed above and the last three estimators are OBM estimators with different batch sizes, where $m^*$ is the MSE-optimal batch size, $m^{1-2-1}$ is the 1-2-1 OBM batch size (Pedrosa and Schmeiser, 1994), and $m^S$ is Song's batch size (Song, 1996). The empirical results are shown in Tables 2 and 3 in terms of bias, variance, and MSE for AR(1) data with $n = 500$ and M/M/1 data with $n = 5000$. In both cases $\gamma_0 = 10$ and $\text{Var}(\bar{Y}) = 1$.

Table 2. AR(1), n=500

| $\hat{V}$ | Bias | Variance | MSE |
|---|---|---|---|
| $\hat{V}^{(*)}(\hat{\mathbf{Q}}^*)$ | -0.29 (.01) | .05 (.01) | .13 (.01) |
| $\hat{V}^{(O)}(m^*)$ | -.14 (.01) | .06 (.006) | .08 (.005) |
| $\hat{V}^{(O)}(m^{1-2-1})$ | -.20 (.01) | .07 (.01) | .11 (.07) |
| $\hat{V}^{(O)}(m^S)$ | -.24 (.01) | .05 (.01) | .11 (.01) |

Table 3. M/M/1, n=5000

| $\hat{V}$ | Bias | Varance | MSE |
|---|---|---|---|
| $\hat{V}^{(*)}(\hat{\mathbf{Q}}^*)$ | -0.08 (.01) | .24 (.02) | .25 (.02) |
| $\hat{V}^{(O)}(m^*)$ | -.25 (.004) | .08 (.003) | .14 (.003) |
| $\hat{V}^{(O)}(m^{1-2-1})$ | -.08 (.01) | .24 (.02) | .25 (.02) |
| $\hat{V}^{(O)}(m^S)$ | -.09 (.01) | .24 (.01) | .25 (.01) |

As can be seen, the proposed estimated optimal Q-F estimator does not perform better nor worse than Pedrosa and Schmeiser's 1-2-1 OBM or Song's estimator. Both Pedrosa and Schmeiser's 1-2-1 OBM and Song's estimator have similar MSE, although the tradeoffs between bias and variance differ.

In the proposed simple algorithm, we use OBM estimator with 1-2-1 OBM as the batch size to estimate the unknown parameter $\text{Var}(\bar{Y})$ in Equation (13) as the initial value to estimate the optimal Q-F coefficients $\mathbf{Q}^*$. There are other ways to estimate the unknown parameter $\text{Var}(\bar{Y})$. For example, we can estimate individual correlations. One specific method is first fitting an autoregressive process and then estimating the corresponding parameters and finally computing the corresponding correlations (Yuan and Nelson, 1994).

## 5 SUMMARY

This paper proposes the idea of searching for the optimal quadratic-form estimator to estimate the variance of the sample mean for a stationary process. The optimal quadratic-form coefficients are obtained by minimizing an upper bound on the variance of the quadratic-form estimator of $\mathrm{Var}(\bar{Y})$. The statistical performance in terms of both bias and variance outperforms the optimal OBM estimator if the process is known. If the process is unknown, the proposed simple method is still competitive with two existing methods. The theoretical optimal quadratic-form expression provides a reasonable foundation to search for the optimal quadratic-form estimator and the proposed simple method encourages future research.

## APPENDIX

Theorem (Rao, 1973): Let $\mathbf{Q}$ , $\mathbf{V}$, and $\{\mathbf{U}_i, i = 1, 2, ..., k\}$ be positive definite and symmetric matrices. Let $\mathbf{B}$ be any arbitrary matrix and $\{p_i, i = 1, 2, ..., k\}$ be constants. The solution of the following minimization problem

$$\text{minimize } \mathrm{tr}(\mathbf{QVQV})$$

subject to:

$$\mathbf{QB} = 0$$

$$\mathrm{tr}(\mathbf{QU}_i) = p_i, i = 1, 2..., k$$

is

$$\mathbf{Q}^* = \sum_{i=1}^{k} \lambda_i \mathbf{C}^t \mathbf{V}^{-1} \mathbf{U}_i \mathbf{V}^{-1} \mathbf{C},$$

where $\lambda_1, \lambda_2, ..., \lambda_k$ are roots of $\sum_{i=1}^{k} \lambda_i \mathrm{tr}(\mathbf{C}^t \mathbf{V}^{-1} \mathbf{U}_i \mathbf{V}^{-1} \mathbf{C} \mathbf{U}_j) = p_j, j = 1, 2, ..., k$ and $\mathbf{C} = \mathbf{I} - \mathbf{B}(\mathbf{B}^t \mathbf{V}^{-1} \mathbf{B})^{-1} \mathbf{B}^t \mathbf{V}^{-1}$.

## ACKNOWLEDGMENTS

This work was supported by NSC Grant 86–2213–E–007–014 to the National Tsing Hua University. We thank Bruce Schmeiser, Barry Nelson, Bo-Ray Huang, and Chi-Wen Chen for helpful comments.

## REFERENCES

Bischak, D. P., W. D. Kelton and S. M Pollock. 1993. Weighted batch means for confidence intervals in steady-state simulations. *Management Science* 39: 1002–1019.

Chien, C. 1994. Batch size selection for the batch means method. In *Proceedings of the Winter Simulation Conference*, ed. J. D. Tew, S. Manivannan, D. A. Sadowski and A. F. Seila, 345–352. Piscataway, New Jersey: IEEE.

Chien, C., D. M. Goldsman and B. Melamed. 1996. Large-sample results for batch means. *Management Science*, forthcoming.

Crane, M. A. and D. L. Iglehart. 1975 Simulating stable stochastic systems, III: regenerative processes and discrete-event simulation. *Operations Research*, 33-45.

Damerdji, H. 1991. Strong consistency and other properties of the spectral variance estimator. *Management Science* 37: 1424–1440.

Fishman, G. S. and L. S. Yarberry. 1993. An implementation of the batch means method. Technical-Report UNC/OR/TR/93-1, Department of Operations Research, University of North Carolina, Chapel Hill, North Carolina.

Foley, R. D. and D. Goldsman. 1988. Confidence intervals using orthonormally weighted standardized time series. *Proceedings of the Winter Simulation Conference* (M. Abrams, P. Haigh, and J. Comfort, eds.), 422-424.

Fox, B. L., D. M. Goldsman and J. J. Swain. 1991. Spaced batch means. *Operations Research Letters* 10: 255-266.

Glynn, P. W. and D. L. Iglehart. 1990. Simulation output analysis using standardized time series. *Mathematics of Operations Research* 15:1–16.

Glynn, P. W. and W. Whitt. 1991. Estimating the asymptotic variance with batch means. *Operations Research Letters* 10:431–435.

Grenander, U. and M. Rosenblatt. 1957. *Time Series.* Chelsea Publishing Company, New York.

Goldsman, D., M. Meketon, and L. W. Schruben. 1990. Properties of standardized time series weighted area variance estimators. *Management Science* 36:602-612.

Heidelberger, P. and P. D. Welch. 1981. A spectral method for confidence interval generation and run length control in simulation. *Communications of the ACM* 24, 233-245.

Meketon, M. S. and B. Schmeiser. 1984. Overlapping batch means: something for nothing? In *Proceedings of the Winter Simulation Conference*, ed. S. Sheppard, U. Pooch and C. D. Pegden, 227–230. Piscataway, New Jersey: IEEE.

Muñoz, D. F. and P. W. Glynn. 1991. Multivariate standardized time series for steady-state simulation output analysis. Technical report, Department of Operations Research, Stanford University.

Muñoz, D. F. and P. W. Glynn. 1997. A batch means methodology for estimation of a nonlinear function of a steady-state mean. *Management Science* 43, forthcoming.

Pedrosa, A. and B. Schmeiser. 1994. Estimating the variance of the sample mean: Optimal Batch-Size Estimation and 1-2-1 Overlapping Batch Means. Technical Report SMS94-3, School of Industrial Engineering, Purdue University.

Priestley, M.B. 1981. *Spectral Analysis and Time Series*, Academic Press, Boston.

Rao, C. R. 1973. Linear Statistical Inference and Its Applications, New York: John Wiley.

Rao, C. R. and J. Kleffe. 1988. Estimation of Variance Components and Application. New York: Elsevier Science Publishing Company.

Schmeiser, B. 1982. Batch size effects in the analysis of simulation output. *Operations Research* 30:556–568.

Schriber, T. J. and R. W. Andrews. 1984. ARMA-based confidence interval procedures for simulation output analysis. *American Journal of Mathematical and Management Sciences*, 345-373.

Schruben, L. W. 1983. Confidence interval estimation using standardized time series. *Operations Research* 31: 1090–1108.

Sherman, M. 1995. On batch means in the simulation and statistics communities. In *Proceedings of the Winter Simulation Conference*, ed. C. Alexopoulos, K. Kang, W. Lilegdon, and D. Goldsman, 297–303. Piscataway, New Jersey: IEEE.

Song, W. T. 1996. On the estimation of optimal batch sizes in the analysis of simulation output analysis. *European Journal of Operational Research* 88:304–309.

Song, W. T. and B. Schmeiser. 1993. Variance of the sample mean: Properties and graphs of quadratic-form estimators. *Operations Research* 41:501–517.

Song, W. T. and B. Schmeiser. 1995. Optimal mean-squared-error batch sizes. *Management Science* 41:110–123.

Welch, P. D. 1987. On the relationship between batch means, overlapping batch means and spectral estimation. In *Proceedings of the Winter Simulation Conference*, ed. A. Thesen, H. Grant and W. D. Kelton, 309–317. Piscataway, New Jersey: IEEE.

Yuan, M. J. and B. Nelson. 1994. Autoregressive-output-analysis methods revisited. *Annals of Operations Research* 53, 391-418.

## AUTHOR BIOGRAPHIES

**WHEYMING TINA SONG** is a professor in the Department of Industrial Engineering at the National Tsing Hua University in Taiwan. She received a B.S. in statistics and an M.S. in industrial engineering from Cheng-Kung University in Taiwan, M.S. degrees in applied mathematics and industrial engineering from the University of Pittsburgh, and a Ph.D. in industrial engineering from Purdue University. Her research interests are the probabilistic and statistical aspects of stochastic simulation.

**NENG-HUI SHIH** is an associate professor in the Department of Business Administration at National Ping Tung Institute of Commerce in Taiwan. He received a Ph.D. in industrial engineering from Tsing Hua University. His research interests are the statistical aspects of simulation and reliability.

**MINGJIAN YUAN** is an associate professor in the Department of Industrial Management at the Yunlin Technology University in Taiwan. He received a B.S. in Industrial Engineering from Feng-Chia University and an M.S. and Ph.D. in Industrial and Systems Engineering from the Ohio State University. His research interests are the statistical aspects of simulation.