

MULTIVARIATE BATCH MEANS AND CONTROL VARIATES*

WEI-NING YANG AND BARRY L. NELSON

*National Taiwan Institute of Technology, Department of Information Management,
Taipei, Taiwan*
*Department of Industrial and Systems Engineering, The Ohio State University,
Columbus, Ohio 43210*

We consider applying the nonoverlapping batch means output analysis method in conjunction with the control-variate variance-reduction technique to estimate a steady-state multivariate mean vector. The effects of the number of batches and the number of control variates on the multivariate point and region estimators and the univariate point and interval estimators are considered. The results are experiment analysis guidelines in terms of an appropriate range of the number of batches to choose as a function of the number of responses and control variates. The results have implications for terminating simulations as well.

(VARIANCE REDUCTION; BATCH MEANS; CONTROL VARIATES; MULTIVARIATE STATISTICS)

1. Introduction

Computer simulation is frequently employed for the analysis of stochastic systems. There are many situations in which we are interested in several performance measures of a stochastic system simultaneously, possibly of several different systems. However, multivariate estimation procedures are rarely used in simulation output analysis because joint inference on multiple response variables is difficult when the variables are dependent.

Although simulation is often the only feasible method for estimating the parameters of a complex stochastic system, the computing cost for achieving acceptable precision can be a serious disadvantage. Variance reduction techniques can be used to reduce the variance of estimators, and variance reduction is particularly critical in multivariate estimation problems because joint inference procedures are more conservative than univariate procedures. Recent surveys of variance reduction that emphasize univariate estimation include Nelson (1987) and Wilson (1984).

This paper examines the effect of applying the control-variate variance-reduction technique, in conjunction with the batch-means output-analysis method, to estimate a multivariate mean vector; it extends results for the univariate case in Nelson (1989) to the multivariate case. The tradeoffs between using multivariate and univariate estimation procedures are also considered.

To be more precise, suppose the simulation output process is of the form $\mathbf{Z}_i = (\mathbf{Y}'_i, \mathbf{C}'_i)'$, for $i = 1, 2, \dots, n$, where \mathbf{Y} is a $p \times 1$ random vector, \mathbf{C} is a $q \times 1$ random vector, $'$ indicates the transpose of a matrix, and the output process is identically distributed and stationary. The specific problem we consider is estimating the p -variate mean vector $\boldsymbol{\theta} = E[\mathbf{Y}_i]$ when a q -variate control vector, \mathbf{C}_i , with known expectation, $\boldsymbol{\mu}_C$, can also be observed.

Such an output process can arise from either terminating or steady-state simulations. In terminating simulation, \mathbf{Z}_i could be a summary output from the i th independent replication. In steady-state simulation, after initial-condition effects have been removed (see, e.g., Schruben 1981), \mathbf{Z}_i , $i = 1, 2, \dots, n$, could be the outputs from within a single replication, which are typically dependent. Of course, to obtain a discrete-time process

* Accepted by James R. Wilson; received February 17, 1990. This paper has been with the authors 4 months for 2 revisions.

of the form assumed here we may have to transform the natural output process, possibly by batching by time. In either case, we want to form point and region estimators for Θ using the control-variate variance-reduction technique to improve the precision of the estimators.

Standard region estimation procedures assume that the output process consists of independent and identically distributed (i.i.d.) multivariate normal vectors. The assumption of normality is not necessarily true for the output process obtained from a terminating simulation. Both the assumptions of normality and independence may be violated for the output process from a steady-state simulation. Batching is an aid to realizing both assumptions: Batching makes the output processes from both terminating and steady-state simulations closer to normality due to central limit theorem effects, and the output process from a steady-state simulation less dependent for typical covariance structures.

The improvement from batching is obtained at the expense of loss of degrees of freedom. The approach taken in this paper is to assume the conditions of independence and normality are actually satisfied, and then to study the potential penalty for batching in terms of its effect on point and region estimator performance. That is, we directly assess the performance penalty for batching when it is not needed. We find that estimator performance is insensitive to the number of batches within a certain range. These results help to identify when the marginal improvement from additional batches is not worth the risk of significantly violating the assumptions of normality and independence; this is useful for experiment design and analysis because it limits the range within which we need to search for an acceptable number of batches.

The paper is organized as follows: We first review batching and control variates. Then we examine batch-size effects on the variance of the point estimator and the volume of the joint confidence region. Finally, we study the tradeoffs between using individual univariate estimation procedures and multivariate estimation procedures in terms of the properties of the half widths of the confidence intervals for individual univariate responses.

2. Review of Batching and Control Variates

To review the batch-means output-analysis method and control-variates variance-reduction technique, let the output of the simulation experiment \mathbf{Z} be as described above (we temporarily drop the subscript i). Let $\Theta = (\theta_1, \theta_2, \dots, \theta_p)'$ denote the unknown mean vector of interest. The variance-covariance matrix of \mathbf{Z} can be represented by

$$\text{Var}[\mathbf{Z}] = \Sigma = \begin{pmatrix} \Sigma_{YY} & \Sigma_{YC} \\ \Sigma_{CY} & \Sigma_{CC} \end{pmatrix}$$

where Σ_{YY} is the $p \times p$ matrix of $\text{Var}[\mathbf{Y}]$, Σ_{YC} is the $p \times q$ matrix of $\text{Cov}[\mathbf{Y}, \mathbf{C}]$, Σ_{CY} is the $q \times p$ matrix of $\text{Cov}[\mathbf{C}, \mathbf{Y}]$ and Σ_{CC} is the $q \times q$ matrix of $\text{Var}[\mathbf{C}]$.

The idea behind control variates is to identify a q -variate control vector \mathbf{C} that has known expectation, μ_C , and is strongly correlated with the p -variate response, \mathbf{Y} . The deviation $\mathbf{C} - \mu_C$ is then used to counteract the unknown deviation $\mathbf{Y} - \Theta$ by subtracting an appropriate linear transformation of $\mathbf{C} - \mu_C$ from the response. For any fixed $q \times p$ matrix of control coefficients, Φ , the control-variate estimator of Θ is

$$\hat{\Theta}(\Phi) = \mathbf{Y} - \Phi'(\mathbf{C} - \mu_C).$$

Letting $|\cdot|$ denote the determinant of a matrix, the generalized variance of the control-variate estimator is

$$|\text{Var}[\hat{\Theta}(\Phi)]| = |\Sigma_{YY} - 2\Phi'\Sigma_{CY} + \Phi'\Sigma_{CC}\Phi|$$

which is minimized by the optimal matrix of control coefficients

$$\Phi^* = \Sigma_{CC}^{-1} \Sigma_{CY}$$

(Venkatraman and Wilson 1986). The minimum generalized variance is

$$|\text{Var} [\hat{\theta}(\Phi^*)]| = |\Sigma_{YY}| \cdot \prod_{j=1}^{\nu} (1 - \rho_j^2)$$

where $\nu = \text{rank}(\Sigma_{YC})$, and the ρ_j 's are the canonical correlations between the response Y and the control vector C (for canonical correlation, see, e.g., Kshirsagar 1972).

Generalized variance is an important performance criterion when estimating a multivariate mean vector. However, practical decision-making is often based on univariate performance measures. If we are only interested in the individual univariate responses in a multivariate estimation problem, then the trace of the covariance matrix of the control-variate estimator is also an important criterion.

Consider the j th expected response, $\theta_j = E[Y_j]$, where Y_j denotes the j th element of Y . As a special case of the result above, the variance of the control-variate estimator of the univariate parameter θ_j alone is minimized by the optimal vector of control coefficients

$$\phi_j^* = \Sigma_{CC}^{-1} \Sigma_{CY_j}$$

yielding the minimum variance

$$\text{Var} [\hat{\theta}_j(\phi_j^*)] = [\Sigma_{YY}]_{jj} \cdot (1 - R_j^2),$$

where Σ_{CY_j} is the j th column of Σ_{CY} , $[\Sigma_{YY}]_{jj}$ is the j th diagonal element of Σ_{YY} and R_j^2 is the squared multiple correlation coefficient between C and Y_j .

Since ϕ_j^* minimizes the marginal variance of the control-variate estimator of θ_j , and since ϕ_j^* is the j th column of Φ^* , it follows that the control coefficient matrix Φ^* also minimizes the trace of the variance-covariance matrix of the control-variate estimator. The minimum trace is therefore

$$\text{tr}(\text{Var} [\hat{\theta}(\Phi^*)]) = \sum_{j=1}^p (1 - R_j^2) [\Sigma_{YY}]_{jj},$$

where $\text{tr}(\cdot)$ denotes the trace of a matrix. This result assumes that we use the same control vector C to estimate each univariate response. Later we discuss the possibility of using different control variates for each response.

In practice Σ_{CY} is often unknown, so Φ^* must be estimated. This results in an efficiency loss relative to the minimum generalized variance and trace; see §3 below.

Batching, as we use the term, means to partition the output process into k nonoverlapping batches of size $b = \lfloor n/k \rfloor$ and to compute the batch-mean vectors $\bar{Y}_j(k)$ and $\bar{C}_j(k)$, where

$$\bar{Y}_j(k) = \frac{1}{b} \sum_{i=(j-1)b+1}^{jb} Y_i \quad \text{and}$$

$$\bar{C}_j(k) = \frac{1}{b} \sum_{i=(j-1)b+1}^{jb} C_i$$

for $j = 1, 2, \dots, k$; from here on we assume k divides n evenly.

In the case of a terminating simulation, where the output process may be nonnormal, or a steady-state simulation, where the output process may be dependent and nonnormal, it is hoped that, at least approximately,

$$\bar{\mathbf{Z}}_j(k) = \begin{pmatrix} \bar{\mathbf{Y}}_j(k) \\ \bar{\mathbf{C}}_j(k) \end{pmatrix} \stackrel{i.i.d}{\sim} N_{p+q} \left(\begin{pmatrix} \boldsymbol{\theta} \\ \boldsymbol{\mu}_C \end{pmatrix}, \boldsymbol{\Sigma}(k) \right),$$

for $j = 1, 2, \dots, k$, where

$$\boldsymbol{\Sigma}(k) = \begin{pmatrix} \boldsymbol{\Sigma}_{YY}(k) & \boldsymbol{\Sigma}_{YC}(k) \\ \boldsymbol{\Sigma}_{CY}(k) & \boldsymbol{\Sigma}_{CC}(k) \end{pmatrix}$$

is analogous to $\boldsymbol{\Sigma}$ for the original output process. The approximations of independence and normality will tend to improve as k , the number of batches, decreases (the batch size b increases).

In the following sections we construct point and region estimators based on the control-variate point estimator formed from batch means, and we examine the effects of batch size and number of control variates on estimator performance.

3. Point Estimator

Let $\bar{\mathbf{Y}}$ and $\bar{\mathbf{C}}$ denote the sample mean vectors of the response and the control variates, respectively,

$$\begin{aligned} \bar{\mathbf{Y}} &= \frac{1}{k} \sum_{j=1}^k \bar{\mathbf{Y}}_j(k) = \frac{1}{n} \sum_{i=1}^n \mathbf{Y}_i, \\ \bar{\mathbf{C}} &= \frac{1}{k} \sum_{j=1}^k \bar{\mathbf{C}}_j(k) = \frac{1}{n} \sum_{i=1}^n \mathbf{C}_i, \end{aligned}$$

and let $\bar{\mathbf{Z}} = (\bar{\mathbf{Y}}', \bar{\mathbf{C}}')'$. Let $\hat{\boldsymbol{\Sigma}}_{YY}(k)$, $\hat{\boldsymbol{\Sigma}}_{CY}(k)$ and $\hat{\boldsymbol{\Sigma}}_{CC}(k)$ denote, respectively, the sample analogues of $\boldsymbol{\Sigma}_{YY}(k)$, $\boldsymbol{\Sigma}_{CY}(k)$ and $\boldsymbol{\Sigma}_{CC}(k)$, which are computed from the batch mean vectors as follows:

$$\begin{aligned} \hat{\boldsymbol{\Sigma}}_{YY}(k) &= \frac{1}{k-1} \sum_{j=1}^k (\bar{\mathbf{Y}}_j(k) - \bar{\mathbf{Y}})(\bar{\mathbf{Y}}_j(k) - \bar{\mathbf{Y}})', \\ \hat{\boldsymbol{\Sigma}}_{CY}(k) &= \frac{1}{k-1} \sum_{j=1}^k (\bar{\mathbf{C}}_j(k) - \bar{\mathbf{C}})(\bar{\mathbf{Y}}_j(k) - \bar{\mathbf{Y}})', \\ \hat{\boldsymbol{\Sigma}}_{CC}(k) &= \frac{1}{k-1} \sum_{j=1}^k (\bar{\mathbf{C}}_j(k) - \bar{\mathbf{C}})(\bar{\mathbf{C}}_j(k) - \bar{\mathbf{C}})'. \end{aligned}$$

Then the optimal control coefficient can be estimated by

$$\hat{\boldsymbol{\Phi}}^*(k) = \hat{\boldsymbol{\Sigma}}_{CC}^{-1}(k) \hat{\boldsymbol{\Sigma}}_{CY}(k)$$

and a control-variate point estimator of $\boldsymbol{\theta}$ is

$$\hat{\boldsymbol{\theta}}(k, p, q) = \bar{\mathbf{Y}} - (\hat{\boldsymbol{\Phi}}^*(k))'(\bar{\mathbf{C}} - \boldsymbol{\mu}_C).$$

The following theorem establishes the basic properties of this estimator when batching is not necessary:

THEOREM 3.1 (Venkatraman and Wilson 1986). *If $\mathbf{Z}_i, i = 1, 2, \dots, n$, are i.i.d. normal, then $E[\hat{\boldsymbol{\theta}}(n, p, q)] = \boldsymbol{\theta}$ and*

$$\frac{|\text{Var}[\hat{\boldsymbol{\theta}}(n, p, q)]|}{|\boldsymbol{\Sigma}_{\bar{\mathbf{Y}}\bar{\mathbf{Y}}}|} = \left(\frac{n-2}{n-q-2} \right)^p \cdot \prod_{j=1}^p (1 - \rho_j^2)$$

where the ρ_j 's are the canonical correlations between $\bar{\mathbf{Y}}$ and $\bar{\mathbf{C}}$.

Under the same assumptions as Theorem 3.1 we can readily show that

$$\frac{\text{tr} (\text{Var} [\hat{\Theta}(n, p, q)])}{\text{tr} (\Sigma_{\tilde{Y}\tilde{Y}})} = \left(\frac{n - 2}{n - q - 2} \right) \cdot \frac{\sum_{j=1}^p (1 - R_j^2) [\Sigma_{\tilde{Y}\tilde{Y}}]_{jj}}{\sum_{j=1}^p [\Sigma_{\tilde{Y}\tilde{Y}}]_{jj}}.$$

These results are for the case when there is no batching ($k = n$), and the original output process is i.i.d. normal. If the independence and normality assumptions are not valid, then we may batch the output process in hopes that, for some number of batches k small enough (equivalently, some batch size b large enough), the batch means are approximately i.i.d. normal.

Although the assumptions of independence and normality may never be precisely true, for the purpose of analysis we will assume that for all numbers of batches k the batch means are indeed i.i.d. normal. Therefore, $\Sigma(k)/k = \text{Var} [\bar{\mathbf{Z}}]$, which equals Σ/n in the special case that the original output process is i.i.d. The following results, which are similar to Nelson (1989) for the case of a single response ($p = 1$), assume that the batch means are i.i.d. normal.

THEOREM 3.2. *For fixed p and q , and $q + 2 < k$,*

$$\frac{|\text{Var} [\hat{\Theta}(k, p, q)]|}{|\Sigma_{\tilde{Y}\tilde{Y}}|} = \left(\frac{k - 2}{k - q - 2} \right)^p \cdot \prod_{j=1}^p (1 - \rho_j^2) \quad \text{and}$$

$$\frac{\text{tr} (\text{Var} [\hat{\Theta}(k, p, q)])}{\text{tr} (\Sigma_{\tilde{Y}\tilde{Y}})} = \left(\frac{k - 2}{k - q - 2} \right) \cdot \frac{\sum_{j=1}^p (1 - R_j^2) [\Sigma_{\tilde{Y}\tilde{Y}}]_{jj}}{\sum_{j=1}^p [\Sigma_{\tilde{Y}\tilde{Y}}]_{jj}}.$$

The proof of this theorem, and all other theorems, is in the appendix.

THEOREM 3.3. *For fixed p and q , and $q + 2 < k_1 < k_2$,*

$$\frac{|\text{Var} [\hat{\Theta}(k_1, p, q)]|}{|\text{Var} [\hat{\Theta}(k_2, p, q)]|} = \left[\frac{(k_1 - 2)(k_2 - q - 2)}{(k_2 - 2)(k_1 - q - 2)} \right]^p > 1 \quad \text{and}$$

$$\frac{\text{tr} (\text{Var} [\hat{\Theta}(k_1, p, q)])}{\text{tr} (\text{Var} [\hat{\Theta}(k_2, p, q)])} = \frac{(k_1 - 2)(k_2 - q - 2)}{(k_2 - 2)(k_1 - q - 2)} > 1.$$

Theorems 3.2 and 3.3 compare the control-variate point estimator to the sample mean, and to itself, for different numbers of batches, but always assuming that the batch means are i.i.d. normal. These are worst-case results for batching in the sense that the penalty is greatest because batching is not needed at all.

Theorem 3.2 shows that, for fixed p and q , increasing k decreases the generalized variance, especially for larger p , meaning that having a larger number of batches is more important when estimating more parameters. Similarly, increasing k decreases the trace of the covariance matrix, but the number of responses has no effect on the leading term of this ratio. For $p \leq 5$, the number of responses has a dramatic effect on the loss ratio $((k - 2)/(k - q - 2))^p$ when k is small, but little when $k \geq 80$.

Comparing control-variate estimators, more batches is better (the ratios in Theorem 3.3 are greater than 1), as would be expected. However, it is important to notice that the ratios are nearly 1 if $k_1 \geq 80$ when $p \leq 5$ and $q \leq 5$, no matter how large k_2 is. We can see this by letting $k_2 \rightarrow \infty$ in Theorem 3.3, which gives an upper bound on the penalty for using k_1 batches when we could have used k_2 batches. The resulting ratio is $((k_1 - 2)/(k_1 - q - 2))^p$, or just the loss ratio in Theorem 3.2. This means that the improvement from a larger number of batches is negligible beyond, say, 80, unless p is quite large.

We are considering different numbers of batches when the total number of observations, n , is fixed. It is interesting to contrast the batch-size effect with the effect of additional

sampling. Suppose that, under the same assumptions as Theorem 3.3, $\hat{\Theta}(k_1, p, q)$ is formed from k_1 batches of size b_1 , and $\tilde{\Theta}(k_2, p, q)$ is a control-variate estimator formed from $k_2 > k_1$ batches of the *same* batch size b_1 ; that is, $\tilde{\Theta}(k_2, p, q)$ is based on a larger total sample. The ratio of generalized variances $|\text{Var} [\hat{\Theta}(k_1, p, q)]|/|\text{Var} [\tilde{\Theta}(k_2, p, q)]|$ is then the ratio in Theorem 3.3 with the right-hand side multiplied by $(k_2/k_1)^p$. Thus, in the case of additional sampling, the improvement from more batches is magnified by $(k_2/k_1)^p$, compared to the case where n is fixed.

This is a basic theme of the paper: When the total sample size is fixed, the improvement in estimator performance from using a larger number of batches decreases rapidly as the number of batches increases, unlike the improvement from additional sampling. Our results help to identify the point where the marginal improvement from additional batches is not worth the risk of significantly violating the assumptions of normality and independence.

Next we examine the effect of the number of control variates. Consider two different sets of control variates containing q_1 and q_2 control variates. We add an argument (q) to ρ^2 , ν and R^2 to emphasize their dependence on the particular control variates.

THEOREM 3.4. *For fixed k and p ,*

$$|\text{Var} [\hat{\Theta}(k, p, q_2)]| < |\text{Var} [\hat{\Theta}(k, p, q_1)]|$$

if and only if

$$1 - \frac{\prod_{j=1}^{\nu(q_2)} (1 - \rho_j^2(q_2))}{\prod_{j=1}^{\nu(q_1)} (1 - \rho_j^2(q_1))} > 1 - \left(\frac{k - q_2 - 2}{k - q_1 - 2}\right)^p \quad \text{and}$$

$$\text{tr} (\text{Var} [\hat{\Theta}(k, p, q_2)]) < \text{tr} (\text{Var} [\hat{\Theta}(k, p, q_1)])$$

if and only if

$$1 - \frac{\sum_{j=1}^p (1 - R_j^2(q_2)) [\Sigma_{\tilde{v}\tilde{v}}]_{jj}}{\sum_{j=1}^p (1 - R_j^2(q_1)) [\Sigma_{\tilde{v}\tilde{v}}]_{jj}} > 1 - \frac{k - q_2 - 2}{k - q_1 - 2}.$$

A special case of Theorem 3.4 is adding control variates to a fixed set of q_1 control variates. Since $\prod_{j=1}^{\nu} (1 - \rho_j^2(q))$ and $\sum_{j=1}^p (1 - R_j^2(q))$ are nonincreasing in q when new control variates are added, Theorem 3.4 gives lower bounds on

$$1 - \frac{\prod_{j=1}^{\nu(q+1)} (1 - \rho_j^2(q+1))}{\prod_{j=1}^{\nu(q)} (1 - \rho_j^2(q))} \quad \text{and} \quad 1 - \frac{\sum_{j=1}^p (1 - R_j^2(q+1)) [\Sigma_{\tilde{v}\tilde{v}}]_{jj}}{\sum_{j=1}^p (1 - R_j^2(q)) [\Sigma_{\tilde{v}\tilde{v}}]_{jj}} \tag{1}$$

such that adding the $(q + 1)$ st control variate results in a reduction of the generalized variance and trace, respectively. Expression (1) is known in multivariate analysis as the squared partial multiple correlation coefficient.

Table 1 gives the lower bound for various values of p , q and k . For generalized variance, the lower bound increases as p or q increases, but becomes insensitive to q when $k \geq 80$ for $p \leq 5$. This means that, when $p \leq 5$, adding a new control variate is unlikely to increase the generalized variance when the number of batches is sufficiently large, and may reduce it. For the trace of the covariance matrix, the lower bound is given by the $p = 1$ entries, and thus does not depend on p ; it becomes insensitive to q when $k \geq 30$. The volume of a joint confidence region for Θ depends on the size of the generalized variance (see §4 below), while the variances of the univariate estimators of θ_j are reflected in the trace.

TABLE I

For fixed sample size n and q_1 control variates, the lower bound $1 - ((k - q_2 - 2)/(k - q_1 - 2))^p$ on the squared partial correlation for improving generalized variance of the point estimator by adding a control variate ($q_2 = q_1 + 1$).

	k	q_2	1	2	3	4	5
$p = 1$	8		0.167	0.200	0.250	0.333	0.500
	10		0.125	0.143	0.167	0.200	0.250
	16		0.071	0.077	0.083	0.091	0.100
	30		0.036	0.037	0.038	0.040	0.042
	60		0.017	0.018	0.018	0.018	0.019
	80		0.013	0.013	0.013	0.013	0.014
	100		0.010	0.010	0.010	0.011	0.011
	∞		0	0	0	0	0
$p = 3$	8		0.421	0.488	0.578	0.704	0.875
	10		0.330	0.370	0.421	0.488	0.578
	16		0.199	0.213	0.230	0.249	0.271
	30		0.103	0.107	0.111	0.115	0.120
	60		0.051	0.052	0.053	0.054	0.055
	80		0.038	0.038	0.039	0.039	0.040
	100		0.030	0.031	0.031	0.031	0.032
	∞		0	0	0	0	0
$p = 5$	8		0.600	0.672	0.763	0.868	0.969
	10		0.487	0.537	0.598	0.672	0.763
	16		0.310	0.330	0.353	0.379	0.410
	30		0.166	0.172	0.178	0.184	0.192
	60		0.083	0.085	0.086	0.088	0.089
	80		0.062	0.063	0.064	0.065	0.066
	100		0.050	0.050	0.051	0.052	0.052
	∞		0	0	0	0	0

4. Region Estimator

In this section the performance of a joint confidence region for Θ as a function of the number of control variates and batches is examined. Two performance measures of the joint confidence region are considered: expectation and standard deviation of the volume of the joint confidence region. Under the assumption that the batch means are independent and multivariate normally distributed, the joint confidence region achieves its nominal coverage probability. Thus, smaller values of both performance measures correspond to better performance.

If the batch means are i.i.d. normal, a $(1 - \alpha)100\%$ joint confidence region for Θ is (Wilson 1984)

$$\left\{ \mathbf{x} \in \mathbb{R}^p : [\hat{\Theta}(k, p, q) - \mathbf{x}]' \mathbf{G}^{-1}(k) [\hat{\Theta}(k, p, q) - \mathbf{x}] \leq \frac{p}{k - p - q} F_{\alpha}(p, k - p - q) \cdot (1 + T_q^2) \right\},$$

where $(k - 1)T_q^2 = (\bar{\mathbf{C}} - \mu_C)' (\hat{\Sigma}_{\bar{\mathbf{C}}\bar{\mathbf{C}}}^{-1}(k)) (\bar{\mathbf{C}} - \mu_C)$, $\hat{\Sigma}_{\bar{\mathbf{C}}\bar{\mathbf{C}}}(k) = \hat{\Sigma}_{\mathbf{C}\mathbf{C}}(k)/k$, $F_{\alpha}(p, k - p - q)$ is the $1 - \alpha$ quantile of the F distribution with p and $k - p - q$ degrees of freedom, and

$$\begin{aligned} \mathbf{G}(k) &= \frac{k - 1}{k} \{ \hat{\Sigma}_{\mathbf{Y}\mathbf{Y}}(k) - \hat{\Sigma}_{\mathbf{Y}\mathbf{C}}(k) \hat{\Sigma}_{\bar{\mathbf{C}}\bar{\mathbf{C}}}^{-1}(k) \hat{\Sigma}_{\mathbf{C}\mathbf{Y}}(k) \} \\ &= (k - q - 1) \hat{\Sigma}_{\bar{\mathbf{Y}}\bar{\mathbf{C}}}(k). \end{aligned}$$

Let $V_\alpha[\hat{\Theta}(k, p, q)]$ denote the volume of the $(1 - \alpha)100\%$ joint confidence region for Θ based on k batches and q control variates, and let $\Gamma(\cdot)$ denote the Gamma function.

THEOREM 4.1. *If the batch means are i.i.d. normal, then the expected volume of the joint confidence region is*

$$E [V_\alpha[\hat{\Theta}(k, p, q)]] = 2^{p/2} D(\alpha, k, p, q) \frac{\Gamma\left(\frac{k}{2}\right)}{\Gamma\left(\frac{k-p}{2}\right)} \cdot |\Sigma_{\bar{Y}} \cdot \bar{c}|^{1/2} \quad \text{where}$$

$$D(\alpha, k, p, q) = \frac{2\pi^{p/2}}{p\Gamma\left(\frac{p}{2}\right)} \left[\frac{p}{k-p-q} F_\alpha(p, k-p-q) \right]^{p/2} \quad \text{and}$$

$$\Sigma_{\bar{Y}} \cdot \bar{c} = \Sigma_{\bar{Y}\bar{Y}} - \Sigma_{\bar{Y}\bar{C}} \Sigma_{\bar{C}\bar{C}}^{-1} \Sigma_{\bar{C}\bar{Y}}$$

The standard deviation of the volume of the confidence region is

$$\sqrt{\text{Var} [V_\alpha[\hat{\Theta}(k, p, q)]]} = 2^{p/2} D(\alpha, k, p, q)$$

$$\times \left\{ \frac{\Gamma\left(\frac{k}{2}\right)}{\Gamma\left(\frac{k-2p}{2}\right)} \cdot \prod_{i=1}^p \left(\frac{k-q-i}{k-q-2i} \right) - \frac{\Gamma^2\left(\frac{k}{2}\right)}{\Gamma^2\left(\frac{k-p}{2}\right)} \right\}^{1/2} \cdot |\Sigma_{\bar{Y}} \cdot \bar{c}|^{1/2}.$$

Although the expected volume of the confidence region is the primary measure of region estimator performance, the stability of the confidence region is also an important criterion since a highly variable estimator may be far from its expectation in an application. The standard deviation of the volume of the confidence region is a measure of the stability of the confidence region.

4.1. *Fixed p and q*

Suppose the number of responses, p , the number of controls, q , and the particular control variates are fixed but the number of batches, k , varies. Based on the results in Theorem 4.1, the effects of the number of batches on $V_\alpha[\hat{\Theta}(k, p, q)]$ when the batch means are actually i.i.d. normal can be summarized as follows:

1. As k increases, both performance measures of the joint confidence region decrease but at a decreasing rate; so the gain from more batches decreases as the number of batches increases.
2. For larger p (respectively, q), decreases in the performance measures of the joint confidence region are still significant at larger values of k ; in other words, having k large is more valuable when estimating more parameters (respectively, when using more control variates).
3. With respect to the expected volume of the joint confidence region, there is little benefit from increasing the number of batches beyond $k = 80$ when $p \leq 5$ and $q \leq 5$, since the gain from more batches is insignificant; however, reducing the variability of the confidence region requires more batches, say $k = 100$.

Table 2 shows the number of batches k such that the marginal benefit, in terms of reduced expected volume, from 5 additional batches is just less than 5%. This is one way to define the number of batches at which increasing k further, with n fixed, has little additional benefit. Rows in the table show the effect of number of responses, while columns show the effect of number of control variates. Similar tables for the standard deviation in Yang (1989) show the same pattern, but with larger numbers of batches throughout.

TABLE 2
Number of Batches, k, Such that the Marginal Reduction in Expected Volume from 5 Additional Batches is Less than 5% when $\alpha = 0.05$

	<i>p</i>	1	2	3	4	5
<i>q</i>	0	14	22	29	37	44
	1	17	26	33	41	48
	2	20	28	38	44	51
	3	22	32	40	48	55
	4	24	34	42	51	58
	5	26	36	45	54	62

These summary results are similar to Nelson (1989) for $p = 1$. However, while Nelson found little additional benefit from $k > 60$ batches for $q \leq 5$, this upper limit increases as more parameters are estimated. With respect to these two performance measures of the joint confidence region, the results show that there is little additional benefit from $k > 100$ batches for $p \leq 5$ and $q \leq 5$.

4.2. Fixed p and k

For fixed p and k , assessing the effect of varying q is difficult since the units on $E [V_\alpha[\hat{\Theta}(k, p, q)]]$ and $\sqrt{\text{Var} [V_\alpha[\hat{\Theta}(k, p, q)]]} - |\Sigma_{\bar{Y}} \cdot \bar{c}|$ —change not only as q changes, but also with the particular control variates chosen. For simplicity, we only consider adding a new control variate to a fixed set of q control variates. We add an argument (q) to the subscript of $|\Sigma_{\bar{Y}} \cdot \bar{c}|$ to emphasize its dependence on the particular control variates. Since $|\Sigma_{\bar{Y}} \cdot \bar{c}(q)|$ decreases as control variates are added, we can make comparisons by considering

$$1 - \frac{\prod_{j=1}^{\nu(q+1)} (1 - \rho_j^2(q + 1))}{\prod_{j=1}^{\nu(q)} (1 - \rho_j^2(q))} = 1 - \frac{|\Sigma_{\bar{Y}} \cdot \bar{c}(q+1)|}{|\Sigma_{\bar{Y}} \cdot \bar{c}(q)|},$$

the squared partial multiple correlation coefficient of the $(q + 1)$ st control variate, given q control variates, which is always less than or equal to 1. If the squared partial multiple correlation coefficient is not large enough, then adding the $(q + 1)$ st control variate will degrade the performance of the region estimator by inflating $E [V_\alpha[\hat{\Theta}(k, p, q)]]$ and $\sqrt{\text{Var} [V_\alpha[\hat{\Theta}(k, p, q)]]}$.

Let $r\{M\}$ be the bound on the squared partial multiple correlation coefficient of the $(q + 1)$ st control variate, given q control variates, such that

$$1 - \frac{|\Sigma_{\bar{Y}} \cdot \bar{c}(q+1)|}{|\Sigma_{\bar{Y}} \cdot \bar{c}(q)|} \geq r\{M\}$$

implies that the performance measure M of the region estimator is no worse after adding the $(q + 1)$ st control variate, where M stands for $E [V_\alpha[\hat{\Theta}(k, p, q)]]$ or $\sqrt{\text{Var} [V_\alpha[\hat{\Theta}(k, p, q)]]}$. The performance measure M of the region estimator is improved if the squared partial multiple correlation coefficient is strictly greater than $r\{M\}$.

Table 3 gives the values of $r\{M\}$ for $\alpha = 0.05$, $p = 5$ and $k = 16, 30, 60, 80$ and 100 as $(q, q + 1)$ goes from $(0, 1)$ to $(4, 5)$. The results from this table and others in Yang (1989) can be summarized as follows:

1. As p increases, $r\{M\}$ increases for both performance measures; in other words, it is more difficult to improve the region estimator by adding a new control variate when

TABLE 3
Lower Bounds on the Squared Partial Multiple Correlation Coefficient for Adding a Control Variate When $p = 5$ and $\alpha = 0.05$

	$(q, q + 1)$	$r\{E[V_\alpha]\}$	$r\{\sqrt{\text{Var}[V_\alpha]}\}$
$k = 16$	0, 1	0.48	0.74
	1, 2	0.53	0.75
	2, 3	0.58	0.77
	3, 4	0.65	0.82
	4, 5	0.72	0.89
$k = 30$	0, 1	0.21	0.40
	1, 2	0.22	0.38
	2, 3	0.23	0.38
	3, 4	0.24	0.38
	4, 5	0.25	0.38
$k = 60$	0, 1	0.09	0.19
	1, 2	0.10	0.19
	2, 3	0.10	0.18
	3, 4	0.10	0.18
	4, 5	0.10	0.18
$k = 80$	0, 1	0.07	0.14
	1, 2	0.07	0.14
	2, 3	0.07	0.14
	3, 4	0.07	0.14
	4, 5	0.07	0.13
$k = 100$	0, 1	0.05	0.11
	1, 2	0.05	0.11
	2, 3	0.05	0.11
	3, 4	0.06	0.11
	4, 5	0.06	0.11

estimating more parameters, since improvement requires greater squared partial multiple correlation for the new control variate.

2. For fixed p , as k increases, $r\{M\}$ decreases and stabilizes for all values of q ; in other words, for large k it is easier to improve the region estimator by adding a new control variate no matter how many control variates have already been chosen.

3. For all p, q and k , $r\{\sqrt{\text{Var}[V_\alpha[\hat{\Theta}(k, p, q)]]}\} > r\{E[V_\alpha[\hat{\Theta}(k, p, q)]]\}$, meaning that it is more difficult to reduce the variability than the expected size of the confidence region by adding a control variate.

4. Comparing Table 3 to Table 1 indicates that improving region estimator performance requires greater squared partial multiple correlation than required to improve the generalized variance of the point estimator.

These summary results are similar to Nelson (1989) when $p = 1$. However, as the number of responses increases ($p > 1$), reducing $E[V_\alpha[\hat{\Theta}(k, p, q)]]$ and $\sqrt{\text{Var}[V_\alpha[\hat{\Theta}(k, p, q)]]}$ by adding a control variate is more difficult.

4.3. *The Special Case $q = 0$*

When there are no control variates, batching has no effect on the point estimator since it is just \bar{Y} . However, the number of batches does change the properties of the region estimator. The results below extend Schmeiser's (1982) results to $p > 1$.

TABLE 4
 Expectation and Standard Deviation of the Volume of the Confidence Region for $p = 5$ and $q = 0$

k	$\alpha = 0.10$		$\alpha = 0.05$		$\alpha = 0.01$	
	$E [V_\alpha]$	$\sqrt{\text{Var}} [V_\alpha]$	$E [V_\alpha]$	$\sqrt{\text{Var}} [V_\alpha]$	$E [V_\alpha]$	$\sqrt{\text{Var}} [V_\alpha]$
10	11,910	7,986	30,817	20,663	214,114	143,570
30	2,148	670	3,710	1,157	9,901	3,088
60	1,681	356	2,757	583	6,557	1,387
80	1,590	289	2,579	469	5,984	1,088
100	1,540	250	2,481	403	5,674	922
120	1,508	222	2,420	357	5,482	809
∞	274	0	427	0	933	0

The expected volume of the $(1 - \alpha)100\%$ confidence region for p -variate Θ based on k batches, but no control variates, is

$$E [V_\alpha[\hat{\Theta}(k, p, 0)]] = 2^{p/2} \cdot D(\alpha, k, p, 0) \cdot \frac{\Gamma\left(\frac{k}{2}\right)}{\Gamma\left(\frac{k-p}{2}\right)} \cdot \sqrt{|\Sigma_{\bar{Y}\bar{Y}}|}$$

The standard deviation of the volume of the confidence region is

$$\sqrt{\text{Var}} [V_\alpha[\hat{\Theta}(k, p, 0)]] = D(\alpha, k, p, 0) \times \left[\prod_{i=1}^p (k - i) - 2^p \cdot \frac{\Gamma^2\left(\frac{k}{2}\right)}{\Gamma^2\left(\frac{k-p}{2}\right)} \right]^{1/2} \cdot \sqrt{|\Sigma_{\bar{Y}\bar{Y}}|}$$

Table 4 shows the expectation and the standard deviation of the volume of the confidence region for different values of k , α , and $p = 5$. The units are $\sqrt{|\Sigma_{\bar{Y}\bar{Y}}|}$, which does not depend on k .

Consider the case when the number of responses, p , is fixed but the number of batches, k , varies. The results can be summarized as follows:

1. As k increases, both performance measures decrease but at a decreasing rate, meaning that the gain from additional batches is more when k is small. The $k = \infty$ entries are not attainable, since n is finite and $k \leq n$, but they are lower bounds on the performance measures as k increases.
2. As k increases, $\sqrt{\text{Var}} [V_\alpha]$ is affected more by k than is $E [V_\alpha]$.
3. For large p , significant decreases in the performance measures occur at larger values of k , meaning that having k large is more valuable when we are estimating more responses.
4. For small α , the rate of decrease in the performance measures as k increases is faster; so additional batches are more valuable for smaller values of α than for larger values of α .

These summary results are similar to Schmeiser (1982) when $p = 1$. However, while Schmeiser found little additional benefit beyond $k = 30$, this limit increases as we estimate more responses. With respect to the expected volume of the joint confidence region, there is little benefit from increasing the number of batches beyond $k = 44$ for $p \leq 5$, as shown in Table 2. As far as the stability of the joint confidence region is concerned, this number of batches is increased to $k = 75$ for $p \leq 5$; see Yang (1989).

4.4. Univariate Confidence Intervals

Even though constructing a joint confidence region for a multivariate response is important, practitioners often need to make inferences on each univariate response, which leads to simultaneous inference or multiple comparisons. Bonferroni's procedure and Scheffé's projection procedure are two approaches for obtaining multiple univariate confidence intervals. Both of these procedures are conservative in the sense that the actual confidence level may be greater than what is prespecified. This section considers batch-size effects on the efficiency of these procedures when simultaneously applying control variates.

Let $\hat{\theta}_j(k, p, q)$ be the j th component of $\hat{\theta}(k, p, q)$. When using k batches and q control variates, the confidence interval for θ_j based on Bonferroni's inequality that guarantees a confidence level of at least $1 - \alpha$ for all p intervals simultaneously is

$$\left\{ x \in \mathfrak{R} : [\hat{\theta}_j(k, p, q) - x]' [\mathbf{G}_j(k)]^{-1} [\hat{\theta}_j(k, p, q) - x] \leq \frac{1}{k - q - 1} \mathbf{F}_{\alpha/p}(1, k - q - 1) \cdot (1 + T_q^2) \right\},$$

which is equivalent to

$$\theta_j \in \hat{\theta}_j(k, p, q) \pm t_{\alpha/(2p), k-q-1} \cdot \left[\frac{1 + T_q^2}{k - q - 1} \mathbf{G}_j(k) \right]^{1/2},$$

where $t_{\alpha/(2p), k-q-1}$ is the $1 - \alpha/(2p)$ quantile of the t distribution with $k - q - 1$ degrees of freedom, and $\mathbf{G}_j(k)$ is the j th diagonal element of $\mathbf{G}(k)$. Notice that each interval individually has confidence level $1 - \alpha/p$.

Scheffé's projection procedure is used to construct confidence intervals for any linear combination of the mean vector and still achieve the overall confidence level. This projection procedure is very conservative when only the confidence intervals for the univariate means are constructed.

The confidence interval for θ_j using Scheffé's projection procedure is

$$\left\{ x \in \mathfrak{R} : [\hat{\theta}_j(k, p, q) - x]' [\mathbf{G}_j(k)]^{-1} [\hat{\theta}_j(k, p, q) - x] \leq \frac{p}{k - p - q} \cdot \mathbf{F}_\alpha(p, k - p - q) \cdot (1 + T_q^2) \right\}$$

(Miller 1981) which is equivalent to

$$\theta_j \in \hat{\theta}_j(k, p, q) \pm [\mathbf{F}_\alpha(p, k - p - q)]^{1/2} \cdot \left[\frac{p(1 + T_q^2)}{k - p - q} \mathbf{G}_j(k) \right]^{1/2}.$$

Let $H_j(B)$ and $H_j(S)$ denote the half widths of the confidence intervals for θ_j based on Bonferroni's procedure and Scheffé's procedure, respectively, with $1 - \alpha$ overall confidence level when simultaneously applying control variates and batching. Let $[\Sigma_{\tilde{y}} \cdot \tilde{c}]_{jj}$ denote the (j, j) element of $\Sigma_{\tilde{y}} \cdot \tilde{c}$. The next theorem gives the expected half widths of the univariate confidence intervals for θ_j .

THEOREM 4.2. *If the batch means are i.i.d. normal, then the expected half widths of the confidence intervals for θ_j based on Bonferroni's and Scheffé's procedures are*

$$E [H_j(B)] = \left[\frac{2}{k - q - 1} F_{\alpha/p}(1, k - q - 1) \right]^{1/2} \frac{\Gamma\left(\frac{k}{2}\right)}{\Gamma\left(\frac{k-1}{2}\right)} \cdot \sqrt{[\Sigma_{\bar{Y}} \cdot \bar{c}]_{jj}}, \quad \text{and}$$

$$E [H_j(S)] = \left[\frac{2p}{k - q - q} F_{\alpha}(p, k - p - q) \right]^{1/2} \frac{\Gamma\left(\frac{k}{2}\right)}{\Gamma\left(\frac{k-1}{2}\right)} \cdot \sqrt{[\Sigma_{\bar{Y}} \cdot \bar{c}]_{jj}}.$$

The ratio of the expected half width obtained by the Bonferroni procedure to that of Scheffé’s projection procedure can be expressed as

$$\frac{E [H_j(B)]}{E [H_j(S)]} = \left[\frac{F_{\alpha/p}(1, k - q - 1)}{F_{\alpha}(p, k - p - q)} \cdot \frac{k - p - q}{p(k - q - 1)} \right]^{1/2}.$$

The behavior of this ratio for $p \leq 5, q \leq 5$ and $k \leq 200$ can be summarized as follows:

1. For fixed p and q , the Bonferroni procedure dominates Scheffé’s projection procedure in the sense that the ratio is less than 1. The ratio increases as the number of batches, k , increases, meaning that the Bonferroni procedure is more sensitive to the number of batches.

2. For fixed k and p , the ratio decreases at an increasing rate as q increases, meaning that Scheffé’s projection procedure is more conservative than the Bonferroni procedure when more control variates are chosen. For larger k there is no significant decrease in the ratio as q increases.

3. For fixed k and q , the ratio decreases as p increases, meaning that Scheffé’s projection procedure is more conservative than the Bonferroni procedure when estimating more parameters.

Thus, using expected half width as the performance criterion, the Bonferroni procedure is superior to Scheffé’s procedure for constructing individual confidence intervals for the elements of Θ .

5. Discussion

The focus of this paper is on the design and analysis of single-replication experiments. While Nelson (1989) concluded that $10 \leq k \leq 60$ is reasonable for the case of $p = 1$ and $q \leq 5$, we modify those bounds to $60 \leq k \leq 100$ if $p \leq 5$ and $q \leq 5$. As a general principle, the more parameters that are to be estimated or the more control variates that are chosen the larger we would like k to be, provided that the assumptions of normality and independence are not violated. The number of batches at which the departure from independence and normality is significant is usually unknown. Keeping the number of batches small improves the approximations, but if the number of batches is not too small our results show that little is sacrificed in estimator performance due to the loss of degrees of freedom.

These results also apply to steady-state simulations when independent replications are employed and the total budget, n , is fixed. In this case we want to keep the number of replications, which corresponds to the number of batches in single-replication experiments, small to reduce the transient period deletion. Our results suggest dividing the budget into a modest number of long replications (60 to 100).

The results also apply to terminating experiments where independent replications are employed and the number of replications, n , is fixed. Although the observations from independent replications are independent, the normality assumption may be violated.

Batching improves the approximation of normality. Our results suggest batching the observations from n independent replications of terminating experiments into a modest number of batch means when $n > 100$.

Yang (1989) demonstrated that it is not always optimal to use the same set of control variates to estimate each univariate parameter of a multivariate parameter vector individually, even when it is optimal to use all of the control variates to minimize the generalized variance of the multivariate estimator. Similarly, it is not always optimal to batch all univariate output processes of a multivariate output process into the same number of batches. The value of our results is that they show that, beyond a certain number of batches, estimator performance is insensitive to the number of batches or to the number of control variates selected, and this number of batches is not very large. Thus, the benefits from individually selecting control variates or batch sizes are negligible when we can obtain a moderate number of batches.

This paper considers batching to improve the approximations of independence and normality. Although not specifically examined here, batching can also improve the performance of the control-variate point estimator in terms of bias, since unbiasedness is assured by normality of the output process (Nelson 1990).

One obvious direction for future work is to extend our results to the estimation of a multivariate metamodel. Perhaps a more fundamental problem is how to best batch a general multivariate output process—which may contain both discrete and continuous-time univariate processes—in order to obtain the type of stationary batch means process assumed here.¹

¹ This research was partially supported by National Science Foundation Grant No. ECS-8707634. The authors acknowledge many helpful comments by the Departmental Editor, Associate Editor and two anonymous referees.

Appendix

Proofs of the theorems are given in this appendix.

PROOF OF THEOREM 3.2. Since k is in the range such that the batch mean vectors $\bar{Z}_j(k), j = 1, 2, \dots, k$, are i.i.d. multivariate normally distributed, the results follow from Theorem 3.1. □

PROOF OF THEOREM 3.3. From Theorem 3.2 we have

$$|\text{Var} [\hat{\Theta}(k_l, p, q)]| = \left(\frac{k_l - 2}{k_l - q - 2} \right)^p \cdot \prod_{j=1}^p (1 - \rho_j^2) \cdot |\Sigma_{\hat{y}\hat{y}}|$$

for $l = 1$ or 2 , which implies that

$$\begin{aligned} \frac{|\text{Var} [\hat{\Theta}(k_1, p, q)]|}{|\text{Var} [\hat{\Theta}(k_2, p, q)]|} &= \left[\frac{(k_1 - 2)(k_2 - q - 2)}{(k_2 - 2)(k_1 - q - 2)} \right]^p \\ &= \left[\frac{(k_1 - 2)(k_2 - 2) - q(k_1 - 2)}{(k_2 - 2)(k_1 - 2) - q(k_2 - 2)} \right]^p > 1. \end{aligned}$$

The last inequality holds because $k_2 > k_1$. The trace result can be shown similarly. □

PROOF OF THEOREM 3.4. The proof is similar to the proof of Theorem 3.3. □

PROOF OF THEOREM 4.1. For notation, let $\vec{C}(k) = \{\bar{C}_i(k), i = 1, 2, \dots, k\}$. Rao (1966) showed that, conditional on $\vec{C}(k)$,

$$\begin{aligned} \hat{\Theta}(k, p, q) &\sim N_p(\Theta, (1 + T_q^2)\Sigma_{\hat{y}, \hat{e}}) \quad \text{and} \\ \mathbf{G}(k) &\sim W_p(k - q - 1, \Sigma_{\hat{y}, \hat{e}}) \end{aligned}$$

where $(k - 1)T_q^2$ has a Hotelling- T^2 distribution with $k - 1$ degrees of freedom and parameter q when $\Sigma_{\hat{c}\hat{c}} = \hat{\Sigma}_{cc}(k)/k$; and $W_p(k - q - 1, \Sigma_{\hat{y}, \hat{e}})$ denotes the p -dimensional Wishart distribution with $k - q - 1$ degrees of freedom on the covariance matrix $\Sigma_{\hat{y}, \hat{e}}$.

Therefore, conditional on $\bar{\mathbf{C}}(k)$, a $(1 - \alpha)100\%$ confidence region for Θ is

$$\left\{ \mathbf{x} \in \mathbb{R}^p : [\hat{\Theta}(k, p, q) - \mathbf{x}]' \mathbf{G}^{-1}(k) [\hat{\Theta}(k, p, q) - \mathbf{x}] \leq \frac{p}{k - p - q} \mathbf{F}_\alpha(p, k - p - q)(1 + T_q^2) \right\}$$

which yields the conditional volume of the confidence region (Anderson 1984, p. 263)

$$V_\alpha[\hat{\Theta}(k, p, q) | \bar{\mathbf{C}}(k)] = D(\alpha, k, p, q)(1 + T_q^2)^{p/2} |\mathbf{G}(k)|^{1/2}. \tag{2}$$

By Theorem 7.5.3 in Anderson (1984), given $\bar{\mathbf{C}}(k)$,

$$|\mathbf{G}(k)|^{1/2} \sim |\Sigma_{\bar{y} \cdot \bar{c}}|^{1/2} \cdot \chi_{k-q-1} \cdot \chi_{k-q-2} \cdots \chi_{k-q-p},$$

where $\chi_{k-q-1}, \chi_{k-q-2}, \dots, \chi_{k-q-p}$ are independent Chi random variables. Then we have

$$\begin{aligned} E [|\mathbf{G}(k)|^{1/2} | \bar{\mathbf{C}}(k)] &= |\Sigma_{\bar{y} \cdot \bar{c}}|^{1/2} \cdot \prod_{i=1}^p E [\chi_{k-q-i}] \\ &= 2^{p/2} \frac{\Gamma\left(\frac{k-q}{2}\right)}{\Gamma\left(\frac{k-p-q}{2}\right)} \cdot |\Sigma_{\bar{y} \cdot \bar{c}}|^{1/2} \quad \text{and} \end{aligned}$$

$$\begin{aligned} \text{Var} [|\mathbf{G}(k)|^{1/2} | \bar{\mathbf{C}}(k)] &= |\Sigma_{\bar{y} \cdot \bar{c}}| \cdot \left\{ \prod_{i=1}^p E [\chi_{k-q-i}^2] - \left(\prod_{i=1}^p E [\chi_{k-q-i}] \right)^2 \right\} \\ &= \left\{ \prod_{i=1}^p (k - q - i) - 2^p \frac{\Gamma^2\left(\frac{k-q}{2}\right)}{\Gamma^2\left(\frac{k-p-q}{2}\right)} \right\} \cdot |\Sigma_{\bar{y} \cdot \bar{c}}|. \end{aligned}$$

Thus, conditional on $\bar{\mathbf{C}}(k)$, the expected volume of the $(1 - \alpha)100\%$ confidence region for Θ is

$$\begin{aligned} E [V_\alpha[\hat{\Theta}(k, p, q) | \bar{\mathbf{C}}(k)]] &= D(\alpha, k, p, q)(1 + T_q^2)^{p/2} E [|\mathbf{G}(k)|^{1/2} | \bar{\mathbf{C}}(k)] \\ &= 2^{p/2} D(\alpha, k, p, q)(1 + T_q^2)^{p/2} \frac{\Gamma\left(\frac{k-q}{2}\right)}{\Gamma\left(\frac{k-p-q}{2}\right)} \cdot |\Sigma_{\bar{y} \cdot \bar{c}}|^{1/2} \end{aligned}$$

and the conditional variance of the volume of the confidence region is

$$\begin{aligned} \text{Var} [V_\alpha[\hat{\Theta}(k, p, q) | \bar{\mathbf{C}}(k)]] &= D^2(\alpha, k, p, q)(1 + T_q^2)^p \cdot \text{Var} [|\mathbf{G}(k)|^{1/2} | \bar{\mathbf{C}}(k)] \\ &= D^2(\alpha, k, p, q)(1 + T_q^2)^p \left[\prod_{i=1}^p (k - q - i) - 2^p \frac{\Gamma^2\left(\frac{k-q}{2}\right)}{\Gamma^2\left(\frac{k-p-q}{2}\right)} \right] \cdot |\Sigma_{\bar{y} \cdot \bar{c}}|. \end{aligned}$$

Since $(k - 1)T_q^2$ has a Hotelling- T^2 distribution with $k - 1$ degrees of freedom and parameter q , the random variable $(k - q)T_q^2/q$ has an F distribution with degrees of freedom q and $k - q$. By the relationship between the F distribution and the Beta distribution we have

$$\frac{1}{(1 + T_q^2)} \sim \text{Beta}\left(\frac{k - q}{2}, \frac{q}{2}\right)$$

so that

$$E [(1 + T_q^2)^{p/2}] = \frac{\Gamma\left(\frac{k}{2}\right)\Gamma\left(\frac{k-p-q}{2}\right)}{\Gamma\left(\frac{k-q}{2}\right)\Gamma\left(\frac{k-p}{2}\right)}, \tag{3}$$

$$E[(1 + T_q^2)^p] = \frac{\Gamma\left(\frac{k}{2}\right)\Gamma\left(\frac{k-2p-q}{2}\right)}{\Gamma\left(\frac{k-q}{2}\right)\Gamma\left(\frac{k-2p}{2}\right)} \quad \text{and}$$

$$\text{Var}[(1 + T_q^2)^{p/2}] = \frac{\Gamma\left(\frac{k}{2}\right)\Gamma\left(\frac{k-2p-q}{2}\right)}{\Gamma\left(\frac{k-q}{2}\right)\Gamma\left(\frac{k-2p}{2}\right)} - \frac{\Gamma^2\left(\frac{k}{2}\right)\Gamma^2\left(\frac{k-p-q}{2}\right)}{\Gamma^2\left(\frac{k-q}{2}\right)\Gamma^2\left(\frac{k-p}{2}\right)}$$

(Johnson and Kotz 1970). Thus, the expected volume of the $(1 - \alpha)100\%$ confidence region for θ is

$$E[V_\alpha[\hat{\theta}(k, p, q)]] = E[E\{V_\alpha[\hat{\theta}(k, p, q)|\tilde{C}(k)]\}]$$

$$= 2^{p/2}D(\alpha, k, p, q) E[(1 + T_q^2)^{p/2}] \frac{\Gamma\left(\frac{k-q}{2}\right)}{\Gamma\left(\frac{k-p-q}{2}\right)} \cdot |\Sigma_{\tilde{v}} \cdot \tilde{c}|^{1/2}$$

$$= 2^{p/2}D(\alpha, k, p, q) \frac{\Gamma\left(\frac{k}{2}\right)}{\Gamma\left(\frac{k-p}{2}\right)} \cdot |\Sigma_{\tilde{v}} \cdot \tilde{c}|^{1/2}, \tag{4}$$

and the variance of the volume of the confidence region is

$$\text{Var}[V_\alpha[\hat{\theta}(k, p, q)]] = E[\text{Var}\{V_\alpha[\hat{\theta}(k, p, q)|\tilde{C}(k)]\}] + \text{Var}[E\{V_\alpha[\hat{\theta}(k, p, q)|\tilde{C}(k)]\}]$$

$$= D^2(\alpha, k, p, q) E[(1 + T_q^2)^p] \left\{ \prod_{i=1}^p (k - q - i) - 2^p \frac{\Gamma^2\left(\frac{k-q}{2}\right)}{\Gamma^2\left(\frac{k-p-q}{2}\right)} \right\}$$

$$\times |\Sigma_{\tilde{v}} \cdot \tilde{c}| + 2^p D^2(\alpha, k, p, q) \text{Var}[(1 + T_q^2)^{p/2}] \frac{\Gamma^2\left(\frac{k-q}{2}\right)}{\Gamma^2\left(\frac{k-p-q}{2}\right)} \cdot |\Sigma_{\tilde{v}} \cdot \tilde{c}|$$

$$= 2^p D^2(\alpha, k, p, q) \left\{ \frac{\Gamma\left(\frac{k}{2}\right)}{\Gamma\left(\frac{k-2p}{2}\right)} \cdot \prod_{i=1}^p \left(\frac{k-q-i}{k-q-2i} \right) - \frac{\Gamma^2\left(\frac{k}{2}\right)}{\Gamma^2\left(\frac{k-p}{2}\right)} \right\} \cdot |\Sigma_{\tilde{v}} \cdot \tilde{c}|.$$

Results in Bauer and Wilson (1989) were useful for simplifying the expressions in this result. \square

PROOF OF THEOREM 4.2. The expected half width of the confidence interval for θ , when using the Bonferroni inequality is $\frac{1}{2}$ of the expected volume (4) specialized to the case $p = 1$ and confidence level $1 - \alpha/p$; that is, $\frac{1}{2}$ the expected volume as if θ , were estimated in isolation and we wanted a $1 - \alpha/p$ confidence region.

Conditional on $\tilde{C}(k)$, the expected half width of the confidence interval for θ , when using Scheffé's projection procedure is

$$E[H_j(S)|\tilde{C}(k)] = \left[\frac{p}{k-p-q} F_\alpha(p, k-p-q) \right]^{1/2} (1 + T_q^2)^{1/2} \cdot E[\sqrt{G_{jj}(k)}|\tilde{C}(k)].$$

Given $\tilde{C}(k)$, the conditional distribution of $G(k)$ is p -dimensional Wishart with $k - q - 1$ degrees of freedom on the covariance matrix $\Sigma_{\tilde{v}} \cdot \tilde{c}$; furthermore, given $\tilde{C}(k)$, the conditional distribution of $G_{jj}(k)/[\Sigma_{\tilde{v}} \cdot \tilde{c}]_{jj}$ is Chi-squared with $k - q - 1$ degrees of freedom. Thus, we have

$$E[\sqrt{G_{jj}(k)}|\tilde{C}(k)] = \sqrt{2} \frac{\Gamma\left(\frac{k-q}{2}\right)}{\Gamma\left(\frac{k-q-1}{2}\right)} \cdot \sqrt{[\Sigma_{\tilde{v}} \cdot \tilde{c}]_{jj}}$$

yielding

$$E [H_j(S) | \tilde{C}(k)] = \left[\frac{2p}{k-p-q} F_{\alpha}(p, k-p-q) \right]^{1/2} \frac{\Gamma\left(\frac{k-q}{2}\right)}{\Gamma\left(\frac{k-q-1}{2}\right)} (1 + T_q^2)^{1/2} \sqrt{[\Sigma_{\tilde{Y}} \cdot \tilde{c}]_j}.$$

Then using equation (3), the expected half width of the confidence interval for θ_j using Scheffé's projection procedure is

$$E [H_j(S)] = \left[\frac{2p}{k-p-q} F_{\alpha}(p, k-p-q) \right]^{1/2} \frac{\Gamma\left(\frac{k}{2}\right)}{\Gamma\left(\frac{k-1}{2}\right)} \cdot \sqrt{[\Sigma_{\tilde{Y}} \cdot \tilde{c}]_j}. \quad \square$$

References

- ANDERSON, T. W., *An Introduction to Multivariate Statistical Analysis*, Wiley, New York, 1984.
- BAUER, K. W. AND J. R. WILSON, "Control-variate Selection Criteria," Technical Report, School of Industrial Engineering, Purdue University, 1989.
- JOHNSON, N. L. AND S. KOTZ, *Continuous Univariate Distributions 2*, Houghton Mifflin Co., Boston, 1970.
- KSHIRSAGAR, A. M., *Multivariate Analysis*, Marcel Dekker, Inc., New York, 1972.
- MILLER, R. G., *Simultaneous Statistical Inference*, Springer-Verlag, New York, 1981.
- NELSON, B. L., "A Perspective on Variance Reduction in Dynamic Simulation Experiments," *Comm. Stat.*, B16 (1987), 385-426.
- , "Batch Size Effects on the Efficiency of Control Variates in Simulation," *European J Oper Res*, 43 (1989), 184-196.
- , "Control-Variate Remedies," *Oper. Res.*, 38 (1990), 974-992.
- RAO, C. R., "Least Squares Theory Using an Estimated Dispersion Matrix and Its Application to Measurement of Signals," *Berkeley Sympos. Math. Statist. and Probability Proc.*, 1, 1966, 355-372.
- SCHMEISER, B., "Batch Size Effects in the Analysis of Simulation Output," *Oper. Res.*, 30 (1982), 556-568.
- SCHRUBEN, L., "Control of Initialization Bias in Multivariate Simulation Response," *Comm ACM*, 24 (1981), 246-252.
- VENKATRAMAN, S. AND J. R. WILSON, "The Efficiency of Control Variates in Multiresponse Simulation," *Oper Res. Lett.*, 5 (1986), 37-42.
- WILSON, J. R., "Variance Reduction Techniques for Digital Simulation," *Amer. J. Math. and Management Sci.*, 4 (1984), 277-312.
- YANG, W., "Multivariate Estimation and Variance Reduction for Terminating and Steady-Simulation," Ph.D. Dissertation, Department of Industrial and Systems Engineering, The Ohio State University, Columbus, OH, 1989.