

A Splitting Scheme for Control Variates

Athanassios N. Avramidis

Cornell University, Ithaca, NY, 14853, USA

James R. Wilson

North Carolina State University, Raleigh, NC, 27695, USA

This paper details a new control-variate splitting scheme yielding an unbiased estimator of the mean response and an unbiased estimator of the variance of the first estimator. This scheme also yields an asymptotically exact confidence interval for the mean response. We present analytical and empirical performance comparisons of this scheme versus other control-variate procedures.

control variates; simulation experiments; splitting; variance reduction techniques

Address of the corresponding author:

James R. Wilson

Department of Industrial Engineering

North Carolina State University

Box 7906

Raleigh, NC 27695-7906, USA

Phone: (919) 515-6415

Fax: (919) 515-5281

Internet: jwilson@eos.ncsu.edu

1. Introduction

When applying the method of control variates (CV) to a simulation experiment with a univariate response Y whose mean is to be estimated, we identify a $1 \times q$ vector of concomitant random variables $\mathbf{C} = (C_1, \dots, C_q)$ having known, finite expectation $\boldsymbol{\mu}_{\mathbf{C}} \equiv E[\mathbf{C}]$ and a strong linear association with Y . To estimate $\theta \equiv E[Y]$, we try to predict the unknown deviation $Y - \theta$ as a linear function of the known deviation $\mathbf{C} - \boldsymbol{\mu}_{\mathbf{C}}$ and adjust the response accordingly:

$$Y_{\text{CV}}(\mathbf{a}) \equiv Y - \mathbf{a}(\mathbf{C} - \boldsymbol{\mu}_{\mathbf{C}})^{\text{T}}.$$

For any constant $1 \times q$ vector \mathbf{a} of control coefficients, the controlled response $Y_{\text{CV}}(\mathbf{a})$ is an unbiased estimator of θ . Let $\boldsymbol{\Sigma}_{Y\mathbf{C}} \equiv \text{cov}(Y, \mathbf{C})$ and $\boldsymbol{\Sigma}_{\mathbf{C}} \equiv \text{cov}(\mathbf{C})$, where we assume that all of the elements of these matrices are finite and that $\det(\boldsymbol{\Sigma}_{\mathbf{C}}) > 0$. The variance of $Y_{\text{CV}}(\mathbf{a})$ is minimized by the optimal control coefficient vector

$$\boldsymbol{\beta} = \boldsymbol{\Sigma}_{Y\mathbf{C}}\boldsymbol{\Sigma}_{\mathbf{C}}^{-1}; \tag{1}$$

see [10]. Even though in some applications $\boldsymbol{\Sigma}_{\mathbf{C}}$ may be known, $\boldsymbol{\Sigma}_{Y\mathbf{C}}$ is almost always unknown; and therefore $\boldsymbol{\beta}$ must be estimated.

We consider estimators of $\boldsymbol{\beta}$ and θ based on a random sample of n observations $\{(Y_i, \mathbf{C}_i) : i = 1, \dots, n\}$. The most commonly used control coefficient vector is the sample analog of $\boldsymbol{\beta}$,

$$\mathbf{b} = \mathbf{S}_{Y\mathbf{C}}\mathbf{S}_{\mathbf{C}}^{-1}, \tag{2}$$

computed from the given data set, where $\mathbf{S}_{Y\mathbf{C}}$ is the vector of sample covariances between the response and the controls, and $\mathbf{S}_{\mathbf{C}}$ is the sample covariance matrix of the controls. Using the control coefficient vector in (2), we obtain the classical control-variate estimator of θ

$$\hat{\theta}_{\text{CV}}(n) \equiv \bar{Y} - \mathbf{b}(\bar{\mathbf{C}} - \boldsymbol{\mu}_{\mathbf{C}})^{\text{T}}, \tag{3}$$

where \bar{Y} and $\bar{\mathbf{C}}$ are the sample means of $\{Y_i : i = 1, \dots, n\}$ and $\{\mathbf{C}_i : i = 1, \dots, n\}$ respectively. Without some additional assumptions about the joint distribution of each pair (Y_i, \mathbf{C}_i) , the controlled estimator $\hat{\theta}_{\text{CV}}(n)$ is, in general, biased. In addition, there is no known unbiased *internal* estimator of $\text{var}[\hat{\theta}_{\text{CV}}(n)]$ (that is, an unbiased estimator calculated from the same set of observations used to calculate $\hat{\theta}_{\text{CV}}(n)$); hence we must generate independent replications of $\hat{\theta}_{\text{CV}}(n)$ to obtain a reliable estimator of the sampling error in this statistic. Intuitively, this is not efficient, since each estimate of $\boldsymbol{\beta}$ is based on a small subset of the available observations.

Splitting is a well-known remedy that provides unbiased CV point estimators. The key idea, in its simplest form, is to split the complete sample into two groups, compute the estimator \mathbf{b} from the observations in the first group, and use this estimator to control the observations in the second group. By the independence of random variables computed from different groups, the split-CV estimator is unbiased. However, this means wasting the Y -observations in the first group. Tocher [16] and Nelson [12] considered other splitting schemes that use all of the Y -observations, but there appear to be difficulties with variance estimation in these schemes. One approach by Tocher uses two groups ([16], pp. 115–116) and lacks a variance estimator. Nelson’s scheme is a special case of another approach discussed by Tocher ([16], p. 116) and attributed to Tukey. Nelson gave a variance estimator that is generally biased and usually underestimates the variance. Nelson also constructed confidence intervals based on his split-CV point estimator and the associated variance estimator; and he found that these intervals typically have lower-than-nominal coverage for small sample sizes.

In this paper we propose a new control-variate estimation procedure based on splitting. In Section 2, we develop an unbiased point estimator of θ , an unbiased internal estimator of the variance of the first estimator, and an approximate confidence interval for θ . In Section 3, we calculate the variance of our split-CV point estimator when the response and the controls are jointly normal; and under much more general conditions on the joint distribution of the response and the controls, we derive some fundamental asymptotic properties of the point estimator, the variance estimator, and the confidence-interval estimator based on our splitting scheme. In Section 4 we present the results of an empirical performance comparison of our split-CV estimation procedure versus the classical CV procedure and Nelson’s procedure in the context of estimating the mean completion time for stochastic activity networks. Although this paper is based on [3], a precursor of the splitting scheme detailed in this paper was originally presented in [5].

2. Control-variate estimation with splitting

Let $\{(Y_i, \mathbf{C}_i) : i = 1, \dots, n\}$ be the observations obtained from n independent replications of the simulation experiment. The complete sample is split into m groups ($m \geq 2$) of $k \equiv n/m$ observations each so that the ℓ th group $\mathcal{H}_\ell \equiv \{(Y_i, \mathbf{C}_i) : i \in \mathcal{I}_\ell\}$ consists of the observations with indices in the set

$$\mathcal{I}_\ell \equiv \{(\ell - 1)k + 1, (\ell - 1)k + 2, \dots, \ell k\}, \quad \ell = 1, \dots, m.$$

In this paper, n is always assumed to be an integral multiple of m with $n = km$. Recommendations on how to choose m are given in Section 3.

Let $\mathbf{b}_\ell = \mathbf{b}_\ell(\mathcal{H}_\ell)$ be the control coefficient estimator (2) computed solely from the observations belonging to the ℓ th group \mathcal{H}_ℓ . Thus in terms of the sample statistics

$$\bar{Y}_\ell \equiv k^{-1} \sum_{i \in \mathcal{I}_\ell} Y_i \quad \text{and} \quad \bar{\mathbf{C}}_\ell \equiv k^{-1} \sum_{i \in \mathcal{I}_\ell} \mathbf{C}_i,$$

we have

$$\mathbf{b}_\ell = \left[\sum_{i \in \mathcal{I}_\ell} (Y_i - \bar{Y}_\ell)(\mathbf{C}_i - \bar{\mathbf{C}}_\ell) \right] \left[\sum_{i \in \mathcal{I}_\ell} (\mathbf{C}_i - \bar{\mathbf{C}}_\ell)^\top (\mathbf{C}_i - \bar{\mathbf{C}}_\ell) \right]^{-1}$$

for $\ell = 1, 2, \dots, m$. To simplify subsequent expressions, we define the *group-membership* function $\gamma(i) \equiv \lfloor (i-1)/k \rfloor + 1$, $i = 1, \dots, n$, where $\lfloor x \rfloor$ denotes the greatest integer $\leq x$ so that $(Y_i, \mathbf{C}_i) \in \mathcal{H}_{\gamma(i)}$ for $i = 1, \dots, n$. Using the auxiliary function $\delta(\ell) \equiv \ell \pmod{m} + 1$, $\ell = 1, \dots, m$, we define the *control-assignment* function $\tau(i) \equiv \delta[\gamma(i)]$, $i = 1, \dots, n$, specifying for each observation (Y_i, \mathbf{C}_i) the corresponding “control” group $\mathcal{H}_{\tau(i)}$ from which to compute the control coefficient vector $\mathbf{b}_{\tau(i)}$. Thus the i th controlled response is

$$Z_i \equiv Y_i - \mathbf{b}_{\tau(i)}(\mathbf{C}_i - \boldsymbol{\mu}_{\mathbf{C}})^\top \quad \text{for } i = 1, \dots, n. \quad (4)$$

Notice that \mathbf{C}_i and $\mathbf{b}_{\tau(i)}$ are independent since they are respective functions of two *disjoint* groups $\mathcal{H}_{\gamma(i)}$ and $\mathcal{H}_{\tau(i)}$ of independent identically distributed (i.i.d.) observations. It follows immediately that each controlled response (4) is an unbiased estimator of θ . Furthermore, we have the following basic property of the controlled responses.

Proposition 1 *If $m \geq 3$, then the controlled responses $\{Z_i : i = 1, \dots, n\}$ are pairwise uncorrelated.*

PROOF. For $1 \leq i \neq j \leq n$, we have

$$\begin{aligned} \text{cov}(Z_i, Z_j) &= \text{E}[Z_i Z_j] - \text{E}[Z_i] \text{E}[Z_j] \\ &= \text{E}[Y_i Y_j] - \text{E}[Y_i \mathbf{b}_{\tau(j)}(\mathbf{C}_j - \boldsymbol{\mu}_{\mathbf{C}})^\top] - \text{E}[Y_j \mathbf{b}_{\tau(i)}(\mathbf{C}_i - \boldsymbol{\mu}_{\mathbf{C}})^\top] \\ &\quad + \text{E}[\mathbf{b}_{\tau(i)}(\mathbf{C}_i - \boldsymbol{\mu}_{\mathbf{C}})^\top \mathbf{b}_{\tau(j)}(\mathbf{C}_j - \boldsymbol{\mu}_{\mathbf{C}})^\top] - \text{E}[Y_i] \text{E}[Y_j]. \end{aligned} \quad (5)$$

Observe that the first and the fifth terms on the right-hand side of (5) cancel each other. In the second term, we observe that $Y_i \mathbf{b}_{\tau(j)}$ is independent of \mathbf{C}_j since these two quantities are respective functions of two disjoint sets of mutually independent random vectors. Specifically, $Y_i \mathbf{b}_{\tau(j)}$ is a function of $\{(Y_i, \mathbf{C}_i)\} \cup \mathcal{H}_{\tau(j)}$ while \mathbf{C}_j is a function of $\{(Y_j, \mathbf{C}_j)\}$; and these two sets of i.i.d. observations are disjoint since $i \neq j$.

and $\mathcal{I}_{\tau(j)} \cap \mathcal{I}_{\gamma(j)}$ is empty. Thus for the second term on the right-hand side of (5), we have

$$\mathbb{E}[Y_i \mathbf{b}_{\tau(j)}(\mathbf{C}_j - \boldsymbol{\mu}_{\mathbf{C}})^{\mathsf{T}}] = \mathbb{E}[Y_i \mathbf{b}_{\tau(j)}] \mathbb{E}[(\mathbf{C}_j - \boldsymbol{\mu}_{\mathbf{C}})^{\mathsf{T}}] = 0. \quad (6)$$

A similar argument shows that the third term in (5) also vanishes. Finally we consider the fourth term in (5). Since $m \geq 3$, the definition of $\tau(\cdot)$ ensures that one of the following must occur: $\tau(j) \neq \gamma(i)$ or $\tau(i) \neq \gamma(j)$. If $\tau(i) \neq \gamma(j)$, then by an argument similar to that given for equation (6), we see that the quantities $\mathbf{b}_{\tau(i)}(\mathbf{C}_i - \boldsymbol{\mu}_{\mathbf{C}})^{\mathsf{T}} \mathbf{b}_{\tau(j)}$ and $(\mathbf{C}_j - \boldsymbol{\mu}_{\mathbf{C}})^{\mathsf{T}}$ are respective functions of two disjoint sets of mutually independent random vectors—namely, $\mathcal{H}_{\tau(i)} \cup \{(Y_i, \mathbf{C}_i)\} \cup \mathcal{H}_{\tau(j)}$ and $\{(Y_j, \mathbf{C}_j)\}$; and thus

$$\mathbb{E}[\mathbf{b}_{\tau(i)}(\mathbf{C}_i - \boldsymbol{\mu}_{\mathbf{C}})^{\mathsf{T}} \mathbf{b}_{\tau(j)}(\mathbf{C}_j - \boldsymbol{\mu}_{\mathbf{C}})^{\mathsf{T}}] = \mathbb{E}[\mathbf{b}_{\tau(i)}(\mathbf{C}_i - \boldsymbol{\mu}_{\mathbf{C}})^{\mathsf{T}} \mathbf{b}_{\tau(j)}] \mathbb{E}[(\mathbf{C}_j - \boldsymbol{\mu}_{\mathbf{C}})^{\mathsf{T}}] = 0. \quad (7)$$

On the other hand, if $\tau(j) \neq \gamma(i)$, then we apply the argument given for equation (7) with the roles of i and j reversed. ■

To build a control-variate procedure for estimating θ based on this splitting scheme with m groups and a total of n observations, we define the *split-control-variate* estimator of θ

$$\hat{\theta}_{\text{SP}}(m, n) \equiv n^{-1} \sum_{i=1}^n Z_i$$

and the variance estimator

$$S_{\text{SP}}^2(m, n) \equiv (n-1)^{-1} \sum_{i=1}^n [Z_i - \hat{\theta}_{\text{SP}}(m, n)]^2.$$

For simplicity, we will occasionally suppress the arguments m and n when referring to the estimators $\hat{\theta}_{\text{SP}}(m, n)$ and $S_{\text{SP}}^2(m, n)$. Now we are ready to state the main result of this section.

Theorem 1 *The statistic $\hat{\theta}_{\text{SP}}(m, n)$ is an unbiased estimator of θ . Furthermore, if $m \geq 3$, then $S_{\text{SP}}^2(m, n)/n$ is an unbiased estimator of $\text{var}[\hat{\theta}_{\text{SP}}(m, n)]$.*

PROOF. The unbiasedness of $\hat{\theta}_{\text{SP}}(m, n)$ follows from the unbiasedness of the controlled responses $\{Z_i : i = 1, \dots, n\}$ as discussed immediately following equation (4). If $m \geq 3$, then by Proposition 1 we have

$$\text{var}[\hat{\theta}_{\text{SP}}(m, n)] = \frac{1}{n^2} \sum_{i=1}^n \sum_{j=1}^n \text{cov}(Z_i, Z_j) = \frac{\text{var}(Z_1)}{n}. \quad (8)$$

It also follows from Proposition 1 and equation (8) that

$$\mathbb{E}[S_{\text{SP}}^2(m, n)/n] = \frac{1}{n(n-1)} \mathbb{E} \left[\sum_{i=1}^n Z_i^2 - n \hat{\theta}_{\text{SP}}^2(m, n) \right] = \frac{\text{var}(Z_1)}{n} = \text{var}[\hat{\theta}_{\text{SP}}(m, n)]. \quad \blacksquare$$

Restricting the proposed splitting scheme to $m = 2$ groups yields Tocher's two-group splitting scheme with $n_0 = n_1 = n/2$ as detailed on pp. 115–116 of [16]. (Due to a typographical error, n_0 and n_1 are changed to n_1 and n_2 respectively in the latter parts of Tocher's development.) The problem with Tocher's scheme is that the Z_i 's are, in general, correlated when $m = 2$; and this does not allow unbiased estimation of $\text{var}[\hat{\theta}_{\text{SP}}(m, n)]$. Similar difficulties arise with Tukey's generalized splitting scheme; and Tocher's formula for the variance of Tukey's split-CV estimator (namely, the last equation on p. 116 of [16]) appears to be incorrect. As Proposition 1 shows, these difficulties can be avoided by using $m \geq 3$ and a suitably chosen control-assignment function to guarantee that the Z_i 's are uncorrelated.

To construct an approximate $100(1 - \alpha)\%$ confidence interval for θ , we apply a standard technique for approximating the distribution of the variance estimator S_{SP}^2 using a chi-square distribution with “effective” degrees of freedom

$$\nu_{\text{eff}} \equiv \frac{2\text{E}^2[S_{\text{SP}}^2]}{\text{var}[S_{\text{SP}}^2]};$$

see p. 283 of [8]. If the Z_i 's were independent instead of just being uncorrelated, then the variance of S_{SP}^2 would be given by

$$\frac{n}{(n-1)^2} \left\{ \text{E}[(Z_1 - \theta)^4] - \text{E}^2[(Z_1 - \theta)^2] \right\}$$

to terms of order n^{-1} ([15], p. 338). This motivated the following heuristic formula for the effective degrees of freedom assigned to S_{SP}^2 :

$$\hat{\nu}_{\text{eff}} \equiv \left\lceil 2S_{\text{SP}}^4 / \frac{n}{(n-1)^2} \left\{ \frac{1}{n} \sum_{i=1}^n (Z_i - \hat{\theta}_{\text{SP}})^4 - S_{\text{SP}}^4 \right\} \right\rceil, \quad (9)$$

where $\lceil x \rceil$ denotes the smallest integer $\geq x$. (To simplify the notation, we suppress the dependence of $\hat{\nu}_{\text{eff}}$ on n and m .) In view of (9), an approximate $100(1 - \alpha)\%$ confidence interval for θ is

$$\hat{\theta}_{\text{SP}} \pm t_{1-\alpha/2}(\hat{\nu}_{\text{eff}}) \frac{S_{\text{SP}}}{\sqrt{n}}, \quad (10)$$

where $t_{1-\alpha/2}(\hat{\nu}_{\text{eff}})$ denotes the quantile of order $1 - \alpha/2$ for Student's t -distribution with $\hat{\nu}_{\text{eff}}$ degrees of freedom.

3. Properties of the splitting scheme

3.1 Efficiency of the splitting scheme under normality

In this subsection we derive the variance of the split-CV estimator under the assumption that the response and the controls observed on each replication have a multivariate Normal distribution:

$$\left\{ \begin{bmatrix} Y_i \\ \mathbf{C}_i^T \end{bmatrix} : i = 1, \dots, n \right\} \stackrel{\text{i.i.d.}}{\sim} N_{q+1} \left(\begin{bmatrix} \theta \\ \boldsymbol{\mu}_C^T \end{bmatrix}, \begin{bmatrix} \sigma_Y^2 & \boldsymbol{\Sigma}_{YC} \\ \boldsymbol{\Sigma}_{YC}^T & \boldsymbol{\Sigma}_C \end{bmatrix} \right). \quad (11)$$

For this case, Lavenberg, Moeller, and Welch [10] showed that the classical CV estimator $\hat{\theta}_{CV}(n)$ is unbiased and has variance

$$\text{var}[\hat{\theta}_{CV}(n)] = \frac{\sigma_{Y \cdot C}^2}{n} \cdot \frac{n-2}{n-(q+2)}, \quad \text{where } \sigma_{Y \cdot C}^2 = \sigma_Y^2 - \boldsymbol{\Sigma}_{YC} \boldsymbol{\Sigma}_C^{-1} \boldsymbol{\Sigma}_{YC}^T.$$

In view of (8), we can easily calculate $\text{var}[\hat{\theta}_{SP}(m, n)]$ from the variance of a single controlled response Z_1 . From (4) and the unbiasedness of Z_1 , we have

$$\begin{aligned} \text{var}(Z_1) &= E[Y_1^2] - 2E[Y_1 \mathbf{b}_2 (\mathbf{C}_1 - \boldsymbol{\mu}_C)^T] + E[(\mathbf{C}_1 - \boldsymbol{\mu}_C) \mathbf{b}_2^T \mathbf{b}_2 (\mathbf{C}_1 - \boldsymbol{\mu}_C)^T] - E^2[Y_1] \\ &= \sigma_Y^2 - 2E[\mathbf{b}_2] \boldsymbol{\Sigma}_{YC}^T + \text{tr}(\boldsymbol{\Sigma}_C E[\mathbf{b}_2^T \mathbf{b}_2]), \end{aligned} \quad (12)$$

where $\text{tr}(\cdot)$ denotes the trace operator. To continue the calculation of $\text{var}(Z_1)$, we need the following results:

$$E[\mathbf{b}_\ell] = \boldsymbol{\Sigma}_{YC} \boldsymbol{\Sigma}_C^{-1} \quad (13)$$

and

$$E[\mathbf{b}_\ell^T \mathbf{b}_\ell] = \boldsymbol{\Sigma}_C^{-1} \boldsymbol{\Sigma}_{YC}^T \boldsymbol{\Sigma}_{YC} \boldsymbol{\Sigma}_C^{-1} + \frac{\sigma_{Y \cdot C}^2}{k - (q+2)} \boldsymbol{\Sigma}_C^{-1}. \quad (14)$$

A sketch of the proof of (13) and (14) is given on p. 126 of [13]; a complete argument is given on pp. 25–28 of [3]. Substituting (13) and (14) into (12) and letting \mathbf{I}_q denote the $q \times q$ identity matrix, we obtain

$$\begin{aligned} \text{var}(Z_1) &= \sigma_Y^2 - 2\boldsymbol{\Sigma}_{YC} \boldsymbol{\Sigma}_C^{-1} \boldsymbol{\Sigma}_{YC}^T + \text{tr} \left[\boldsymbol{\Sigma}_{YC}^T \boldsymbol{\Sigma}_{YC} \boldsymbol{\Sigma}_C^{-1} + \frac{\sigma_{Y \cdot C}^2}{k - (q+2)} \mathbf{I}_q \right] \\ &= \frac{k-2}{k-(q+2)} \sigma_{Y \cdot C}^2. \end{aligned}$$

Using (8), we have

$$\text{var}[\hat{\theta}_{SP}(m, n)] = \frac{\sigma_{Y \cdot C}^2}{n} \cdot \frac{n-2m}{n-(q+2)m}. \quad (15)$$

We see that $\text{var}[\hat{\theta}_{\text{SP}}(m, n)] \geq \text{var}[\hat{\theta}_{\text{CV}}(n)]$ for any number of groups $m \geq 2$ and $\text{var}[\hat{\theta}_{\text{SP}}(m, n)]$ increases with m . This is to be expected: as the number of groups m increases, the estimate \mathbf{b}_ℓ from group \mathcal{H}_ℓ is based on fewer observations and therefore has larger variability. Recalling that we obtain an unbiased variance estimator only when $m \geq 3$, we recommend using $m = 3$ groups. This is based on the result (15) for the normal case as well as Monte Carlo experience for the nonnormal case.

3.2 Large-sample properties of the point estimators

In this subsection, we relax the normality assumption (11), and we derive some asymptotic properties of the splitting scheme as the total sample size, n , becomes large while the number of groups, m , remains fixed. Since we will be dealing with vectors, the notions of convergence with probability 1 (denoted by $\xrightarrow{\text{w.p.1}}$) and convergence in probability (denoted by $\xrightarrow{\mathcal{P}}$) are understood to be with respect to the usual Euclidean topology. First we show that the control coefficient vector \mathbf{b} defined in (2) is strongly consistent, i.e., it converges with probability 1 to the optimal control coefficient vector $\boldsymbol{\beta}$ in (1) as $n \rightarrow \infty$.

Lemma 1 *Let $\boldsymbol{\beta}$ and \mathbf{b} be as in (1) and (2) respectively, where $\det(\boldsymbol{\Sigma}_{\mathbf{C}}) > 0$. Then*

$$\mathbf{b} \xrightarrow{\text{w.p.1}} \boldsymbol{\beta} \text{ as } n \rightarrow \infty.$$

PROOF. Let \mathbf{A}_i (respectively, \mathbf{B}_i) denote the matrix formed from $\mathbf{S}_{\mathbf{C}}$ (respectively, $\boldsymbol{\Sigma}_{\mathbf{C}}$) by replacing the i th row of $\mathbf{S}_{\mathbf{C}}$ (respectively, $\boldsymbol{\Sigma}_{\mathbf{C}}$) with $\mathbf{S}_{Y\mathbf{C}}$ (respectively, $\boldsymbol{\Sigma}_{Y\mathbf{C}}$). Using Cramer's rule (§5-3 of [9]), we see that $b_i = \det(\mathbf{A}_i)/\det(\mathbf{S}_{\mathbf{C}})$ and $\beta_i = \det(\mathbf{B}_i)/\det(\boldsymbol{\Sigma}_{\mathbf{C}})$ for $i = 1, \dots, q$. By the strong law of large numbers (Theorem 22.1 of [7]), each element of $\mathbf{S}_{Y\mathbf{C}}$ (respectively, $\mathbf{S}_{\mathbf{C}}$) converges with probability 1 to the corresponding element of $\boldsymbol{\Sigma}_{Y\mathbf{C}}$ (respectively, $\boldsymbol{\Sigma}_{\mathbf{C}}$) as $n \rightarrow \infty$. Since $\det(\cdot)$ is a continuous function, we see that $\det(\mathbf{A}_i) \xrightarrow{\text{w.p.1}} \det(\mathbf{B}_i)$ for $i = 1, \dots, q$, and $\det(\mathbf{S}_{\mathbf{C}}) \xrightarrow{\text{w.p.1}} \det(\boldsymbol{\Sigma}_{\mathbf{C}}) > 0$ as $n \rightarrow \infty$. The desired conclusion follows immediately. ■

It is well known that the classical CV estimator $\hat{\theta}_{\text{CV}}(n)$ is asymptotically Normal with mean θ and variance parameter $\sigma_{Y, \mathbf{C}}^2$ as $n \rightarrow \infty$; see [12]. Next we show that the split CV estimator $\hat{\theta}_{\text{SP}}(m, n)$ with a fixed number of groups m has the same asymptotic distribution and that it is strongly consistent.

Theorem 2 *For any fixed m ,*

$$(i) \quad \hat{\theta}_{\text{SP}}(m, n) \xrightarrow{\text{w.p.1}} \theta \text{ as } n \rightarrow \infty, \text{ and}$$

(ii) $n^{1/2}[\hat{\theta}_{\text{SP}}(m, n) - \theta] \xrightarrow{\mathcal{D}} N(0, \sigma_{Y, \mathbf{C}}^2)$ as $n \rightarrow \infty$.

PROOF. To prove part (i), we note that in the ℓ th group \mathcal{H}_ℓ ($\ell = 1, \dots, m$), the statistics \bar{Y}_ℓ , $\bar{\mathbf{C}}_\ell$, and \mathbf{b}_ℓ respectively converge with probability 1 to θ , $\boldsymbol{\mu}_{\mathbf{C}}$, and $\boldsymbol{\beta}$ by the strong law of large numbers and Lemma 1; and the desired conclusion follows since $\hat{\theta}_{\text{SP}}(m, n)$ is the average of m terms of the form $\bar{Y}_\ell - \mathbf{b}_{\delta(\ell)}(\bar{\mathbf{C}}_\ell - \boldsymbol{\mu}_{\mathbf{C}})^\text{T}$.

To prove part (ii), we define the auxiliary quantities

$$\left. \begin{aligned} V_{\ell, n} &\equiv (n/m)^{1/2} (\bar{Y}_\ell - \theta) \\ \mathbf{W}_{\ell, n} &\equiv (n/m)^{1/2} (\bar{\mathbf{C}}_\ell - \boldsymbol{\mu}_{\mathbf{C}})^\text{T} \end{aligned} \right\} \text{ for } \ell = 1, \dots, m \text{ and } n = m, 2m, 3m, \dots$$

Moreover, we define

$$\left\{ \begin{pmatrix} V_\ell \\ \mathbf{W}_\ell \end{pmatrix} : \ell = 1, \dots, m \right\} \stackrel{\text{i.i.d.}}{\sim} N_{q+1}(\mathbf{0}_{q+1}, \boldsymbol{\Sigma}), \quad (16)$$

where $\mathbf{0}_{q+1}$ is a $(q+1)$ -dimensional column vector with all components equal to zero and $\boldsymbol{\Sigma}$ is the covariance matrix of (Y_1, \mathbf{C}_1) . Since convergence with probability 1 implies convergence in probability (Theorem 20.5(i) of [7]), Lemma 1 yields

$$[\mathbf{b}_1, \dots, \mathbf{b}_m]^\text{T} \xrightarrow{\mathcal{P}} [\boldsymbol{\beta}, \dots, \boldsymbol{\beta}]^\text{T} \text{ as } n \rightarrow \infty. \quad (17)$$

It follows from (17), the multivariate central limit theorem ([1], Theorem 3.4.3), and Theorem 4.4 of [6] that

$$\Upsilon_n \equiv \begin{bmatrix} V_{1, n} \\ \mathbf{W}_{1, n} \\ \vdots \\ V_{m, n} \\ \mathbf{W}_{m, n} \\ \mathbf{b}_1^\text{T} \\ \vdots \\ \mathbf{b}_m^\text{T} \end{bmatrix} \xrightarrow{\mathcal{D}} \Upsilon \equiv \begin{bmatrix} V_1 \\ \mathbf{W}_1 \\ \vdots \\ V_m \\ \mathbf{W}_m \\ \boldsymbol{\beta}^\text{T} \\ \vdots \\ \boldsymbol{\beta}^\text{T} \end{bmatrix} \text{ as } n \rightarrow \infty,$$

where $\xrightarrow{\mathcal{D}}$ denotes convergence in distribution [7]. If $\{v_1, \dots, v_m\}$ denote arbitrary real numbers, if $\{\mathbf{w}_1, \dots, \mathbf{w}_m\}$ denote arbitrary $q \times 1$ real vectors, and if $\{\mathbf{b}_1, \dots, \mathbf{b}_m\}$ denote arbitrary $1 \times q$ real vectors, then the function

$$\varrho(v_1, \mathbf{w}_1^\text{T}, \dots, v_m, \mathbf{w}_m^\text{T}, \mathbf{b}_1, \dots, \mathbf{b}_m) \equiv m^{-1/2} \sum_{\ell=1}^m [v_\ell - \mathbf{b}_{\delta(\ell)} \mathbf{w}_\ell]$$

is continuous everywhere in $m(2q+1)$ -dimensional Euclidean space; and thus the continuous mapping theorem ([7], Theorem 29.2) implies that

$$\begin{aligned} n^{1/2} [\hat{\theta}_{\text{SP}}(m, n) - \theta] &= n^{1/2} m^{-1} \sum_{\ell=1}^m [\bar{Y}_\ell - \mathbf{b}_{\delta(\ell)} (\bar{\mathbf{C}}_\ell - \boldsymbol{\mu}_{\mathbf{C}})^{\text{T}} - \theta] \\ &= m^{-1/2} \sum_{\ell=1}^m [V_{\ell, n} - \mathbf{b}_{\delta(\ell)} \mathbf{W}_{\ell, n}] \\ &= \varrho(\boldsymbol{\Upsilon}_n^{\text{T}}) \xrightarrow{\mathcal{D}} \varrho(\boldsymbol{\Upsilon}^{\text{T}}) \text{ as } n \rightarrow \infty; \end{aligned}$$

and in view of (16), we see that $\varrho(\boldsymbol{\Upsilon}^{\text{T}}) \sim N(0, \sigma_{Y, \mathbf{C}}^2)$. ■

Now we show that the variance estimator $S_{\text{SP}}^2(m, n)$ with m fixed is strongly consistent.

Theorem 3 *For any fixed m ,*

$$S_{\text{SP}}^2(m, n) \xrightarrow{\text{w.p.1}} \sigma_{Y, \mathbf{C}}^2 \text{ as } n \rightarrow \infty.$$

PROOF. By elementary algebra,

$$S_{\text{SP}}^2(m, n) = \frac{1}{n-1} \sum_{i=1}^n Z_i^2 - \frac{n}{n-1} [\hat{\theta}_{\text{SP}}(m, n)]^2. \quad (18)$$

By Theorem 2(i),

$$\frac{n}{n-1} [\hat{\theta}_{\text{SP}}(m, n)]^2 \xrightarrow{\text{w.p.1}} \theta^2 \text{ as } n \rightarrow \infty. \quad (19)$$

Moreover,

$$\begin{aligned} \frac{1}{n-1} \sum_{i=1}^n Z_i^2 &= \frac{n}{n-1} \cdot \frac{1}{m} \sum_{\ell=1}^m \frac{1}{n/m} \left\{ \sum_{i \in \mathcal{I}_\ell} Y_i^2 + \mathbf{b}_{\delta(\ell)} \left[\sum_{i \in \mathcal{I}_\ell} (\mathbf{C}_i - \boldsymbol{\mu}_{\mathbf{C}})^{\text{T}} (\mathbf{C}_i - \boldsymbol{\mu}_{\mathbf{C}}) \right] \mathbf{b}_{\delta(\ell)}^{\text{T}} \right. \\ &\quad \left. - 2\mathbf{b}_{\delta(\ell)} \left[\sum_{i \in \mathcal{I}_\ell} Y_i (\mathbf{C}_i - \boldsymbol{\mu}_{\mathbf{C}})^{\text{T}} \right] \right\} \\ &\xrightarrow{\text{w.p.1}} \theta^2 + \sigma_{Y, \mathbf{C}}^2, \text{ as } n \rightarrow \infty, \end{aligned} \quad (20)$$

where we have used Lemma 1 and the strong law of large numbers. Combining (19), (20), and (18), we obtain the desired result. ■

3.3 Asymptotic exactness of the confidence interval

In this section we show that, under some moment conditions, the confidence interval (10) is asymptotically exact, i.e., the probability that it covers θ converges to $1 - \alpha$ as $n \rightarrow \infty$.

Theorem 4 Suppose that the following expectations are finite:

$$\mathbb{E}[Y^4], \mathbb{E}[Y^3 C_i], \mathbb{E}[Y^2 C_i C_j], \mathbb{E}[Y C_i C_j C_k], \text{ and } \mathbb{E}[C_i C_j C_k C_l] \quad (21)$$

for $1 \leq i, j, k, l \leq q$. Then for any fixed number of groups m ,

$$p_n \equiv \Pr \left\{ \hat{\theta}_{\text{SP}} - t_{1-\alpha/2}(\hat{\nu}_{\text{eff}}) \frac{S_{\text{SP}}}{\sqrt{n}} \leq \theta \leq \hat{\theta}_{\text{SP}} + t_{1-\alpha/2}(\hat{\nu}_{\text{eff}}) \frac{S_{\text{SP}}}{\sqrt{n}} \right\} \rightarrow 1 - \alpha \text{ as } n \rightarrow \infty.$$

PROOF. We begin by establishing that

$$\hat{\nu}_{\text{eff}} \xrightarrow{\text{w.p.1}} \infty \text{ as } n \rightarrow \infty. \quad (22)$$

By Theorem 3, $S_{\text{SP}}^4 \xrightarrow{\text{w.p.1}} \sigma_{Y \cdot \mathbf{C}}^4$ as $n \rightarrow \infty$. Proceeding along the same lines as in the proof of Theorem 3 and exploiting the finiteness of the product moments (21), we can also show that $\frac{1}{n} \sum_{i=1}^n (Z_i - \hat{\theta}_{\text{SP}})^4$ converges to a finite quantity with probability 1. In view of the definition (9) of $\hat{\nu}_{\text{eff}}$, relation (22) follows. Thus $t_{1-\alpha/2}(\hat{\nu}_{\text{eff}}) \xrightarrow{\text{w.p.1}} z_{1-\alpha/2}$, where $z_{1-\alpha/2}$ denotes the quantile of order $1 - \alpha/2$ for the standard Normal distribution. It follows from Theorem 2(ii), Theorem 3, Slutsky's theorem ([14], p. 19), and the continuous mapping theorem that

$$\begin{aligned} \left| \frac{n^{1/2}(\hat{\theta}_{\text{SP}} - \theta)}{S_{\text{SP}}} \right| - t_{1-\alpha/2}(\hat{\nu}_{\text{eff}}) &= \left| \frac{n^{1/2}(\hat{\theta}_{\text{SP}} - \theta)}{\sigma_{Y \cdot \mathbf{C}}} \cdot \frac{\sigma_{Y \cdot \mathbf{C}}}{S_{\text{SP}}} \right| - t_{1-\alpha/2}(\hat{\nu}_{\text{eff}}) \\ &\xrightarrow{\mathcal{D}} |N(0, 1)| - z_{1-\alpha/2} \text{ as } n \rightarrow \infty. \end{aligned}$$

Finally we have

$$\begin{aligned} p_n &= \Pr \left\{ \left| \frac{n^{1/2}(\hat{\theta}_{\text{SP}} - \theta)}{S_{\text{SP}}} \right| - t_{1-\alpha/2}(\hat{\nu}_{\text{eff}}) \leq 0 \right\} \\ &\rightarrow \Pr\{|N(0, 1)| - z_{1-\alpha/2} \leq 0\} = 1 - \alpha \text{ as } n \rightarrow \infty. \quad \blacksquare \end{aligned}$$

4. Experimental evaluation

We considered the problem of estimating mean completion time of a stochastic activity network. We studied the behavior of four estimation procedures: (a) the direct-simulation procedure (DI) based on independent replications without controls; (b) the classical control-variate procedure (CCV); (c) Nelson's split-control-variate procedure (NSC) described in Section 6 of [12]; and (d) our split-control-variate procedure with $m = 3$ groups (SC3) as described in Section 2. For each of these four procedures,

we computed a point estimator and a confidence interval. The performance measures of interest were (i) point-estimator mean square error (MSE); (ii) confidence-interval coverage probability; and (iii) confidence-interval expected half-length.

For the experimental performance evaluation, we used two stochastic activity networks of realistic complexity. Network 1 was taken from page 190 of [2]; and the mean activity durations were also taken from the figure given in that reference. Network 2 was taken from page 245 of [11]; and the mean activity durations were taken to be the most-likely times shown in the figure given in that reference.

The simulation models for these activity networks were structured to operate as follows. For each nondummy activity duration V_i in a given network, the associated distribution was taken to be either (a) a normal distribution with a specified mean μ_i and standard deviation $\sigma_i = \mu_i/4$ whose tail was truncated below the value 0; or (b) an exponential distribution with a specified mean μ_i . We chose the exponential distribution as the nonnormal alternative for reasons elaborated in [4]. For network 1 the set of activities with durations as in (a) was taken to be $\{(1,3), (2,6), (2,4), (8,11), (10,13), (12,18), (16,17), (17,21), (17,23), (17,19), (18,19), (23,24)\}$. For network 2 the set of activities with durations as in (a) was taken to be $\{(1,2), (2,34), (2,23), (2,4), (2,32), (24,25), (20,22), (20,21), (21,31), (12,13), (13,14), (14,15), (5,11), (3,9), (4,10), (47,48), (41,42), (50,51)\}$.

The following rule was used for selecting control variates. Ranking the paths in decreasing order of expected duration, we chose the first three path times as the components of the control vector. We define *relative dominance* as the probability that the first of these paths is the critical path (i.e., the longest path) in a single realization of the network. For each of the selected networks, we simulated three variants exhibiting progressively greater relative dominance.

To estimate the mean completion time θ in each network with sufficient accuracy for use as the “true” estimand in reporting confidence-interval coverage probability and MSE in the main simulation study, we performed a large-scale pilot study using only the direct-simulation estimator $\hat{\theta}_{DI}$. Table 1 summarizes the results of the pilot study, displaying for each network the observed value of $\hat{\theta}_{DI}$ and the corresponding estimate of the standard error $SE(\hat{\theta}_{DI})$. Inspection of Table 1 indicates that in each network, the value of θ has been estimated to an accuracy of at least three significant figures. Thus we considered the estimates in Table 1 to be the “true” values of θ in the main simulation study. To ensure that the coverage probabilities reported for the main simulation study were not significantly contaminated by errors in the “true”

Table 1: Direct-simulation estimates of the mean completion time θ based on a pilot study

Network	Dominance	$\hat{\theta}_{\text{DI}}$	$\text{SE}(\hat{\theta}_{\text{DI}})$
1	0.38	980.584	0.1335
	0.65	1614.779	0.1084
	0.81	2729.791	0.0917
2	0.30	68.367	0.0233
	0.58	81.393	0.0241
	0.95	141.660	0.0132

value of θ for each network, we performed the following sensitivity analysis on the results of the main simulation study. Based on the standard errors given in Table 1, we constructed 95% confidence intervals $(\hat{\theta}_{\text{DI}}^L, \hat{\theta}_{\text{DI}}^U)$. The coverages reported for the main simulation study (Tables 2 and 4 below) were compared with the coverages obtained by successively using $\hat{\theta}_{\text{DI}}^L$ and $\hat{\theta}_{\text{DI}}^U$ as the “true” values of θ ; and the largest observed difference from the reported coverages was found to be less than 1.3%.

In previous experimentation [4], we found that the assumption of joint normality (11) between the response and the controls becomes increasingly untenable as the relative dominance increases; and this results in serious degradation in confidence-interval coverage. To assess the effect on confidence-interval coverage of departures from the normality assumption, we set the relative dominance at three broad levels (low, medium, and high); and to assess the effect of sample size n , we performed simulation experiments involving $n = 48, 96$, and 192 independent replications. Motivated by (15), we used $m = 3$ groups for all of the results reported here. For each combination of relative dominance and sample size, we generated 1024 independent experiments; and in each experiment we computed a point estimator of θ and a nominal 90% confidence interval for θ based on the DI, CCV, NSC, and SC3 procedures. By averaging the results across all 1024 experiments, we estimated point-estimator MSE and confidence-interval coverage probability and expected half-length for all four estimation procedures. The results of this experimental performance evaluation are summarized in Tables 2 through 5. In these tables, we have ensured that each entry has a relative error no larger than 5%.

The main conclusions that can be drawn from the results given in Tables 2 through 5 are the following:

1. All control-variate-based confidence intervals (CCV, NSC, and SC3) become

Table 2: Confidence-interval evaluation for network 1

Relative Dominance	Sample Size n	Coverage				Half-length			
		DI	CCV	NSC	SC3	DI	CCV	NSC	SC3
0.38	48	87.8	85.9	87.3	88.2	80.1	23.3	23.5	26.4
	96	88.4	87.5	88.7	90.2	56.5	16.3	16.4	17.4
	192	90.6	88.4	88.7	89.1	39.9	11.5	11.5	11.9
0.65	48	88.1	83.0	84.1	85.5	151.0	14.8	14.9	17.3
	96	88.6	87.3	88.0	88.7	106.4	10.5	10.5	11.5
	192	90.1	87.0	87.1	87.7	75.1	7.41	7.43	7.81
0.81	48	88.1	79.6	79.7	82.8	264.7	8.48	8.56	10.3
	96	89.0	83.6	84.1	85.9	186.4	6.13	6.15	6.99
	192	89.6	85.7	86.0	87.2	131.6	4.37	4.38	4.77

Table 3: Point-estimator MSE for network 1

Relative Dominance	Sample Size n	MSE			
		DI	CCV	NSC	SC3
0.38	48	2414.8	211.2	213.6	243.2
	96	1262.0	100.8	101.1	108.7
	192	574.7	52.0	52.0	55.2
0.65	48	8682.8	86.6	89.7	101.0
	96	4468.6	41.4	42.0	45.7
	192	2033.7	21.8	21.9	23.2
0.81	48	27022.2	32.5	34.6	40.9
	96	13782.5	15.3	15.7	16.3
	192	6246.8	7.44	7.53	7.89

Table 4: Confidence-interval evaluation for network 2

Relative Dominance	Sample Size n	Coverage				Half-length			
		DI	CCV	NSC	SC3	DI	CCV	NSC	SC3
0.30	48	88.6	87.6	87.0	88.5	4.36	3.84	3.81	4.38
	96	89.0	87.3	87.5	88.8	3.11	2.70	2.69	2.89
	192	89.4	89.5	89.9	90.4	2.19	1.89	1.89	1.97
0.58	48	88.1	85.1	85.5	86.5	4.58	2.99	2.98	3.50
	96	89.5	86.6	86.9	88.2	3.26	2.13	2.13	2.34
	192	89.2	88.6	88.6	90.1	2.29	1.50	1.50	1.59
0.95	48	87.9	62.9	65.1	69.9	7.55	0.82	0.85	1.13
	96	90.0	70.9	73.0	76.8	5.33	0.68	0.69	0.89
	192	89.1	77.6	78.5	82.0	3.75	0.51	0.51	0.63

Table 5: Point-estimator MSE for network 2

Relative Dominance	Sample Size n	MSE			
		DI	CCV	NSC	SC3
0.30	48	7.05	5.38	5.42	7.14
	96	3.59	2.79	2.78	2.98
	192	1.73	1.26	1.26	1.30
0.58	48	7.94	3.49	3.54	4.71
	96	3.86	1.80	1.80	1.95
	192	1.96	0.82	0.82	0.85
0.95	48	21.4	0.49	0.54	0.77
	96	10.2	0.27	0.28	0.32
	192	5.34	0.13	0.14	0.14

shorter on the average as relative dominance increases; however, this is achieved at the expense of some loss of coverage, which may be substantially lower than the nominal level in extreme cases (see network 2 with high relative dominance).

2. As the sample size increases, the CCV, NSC, and SC3 confidence intervals approach nominal coverage. This was expected, since all confidence intervals are asymptotically exact (see [10] for CCV; see Theorem 4 for SC3; the NSC confidence interval can also be shown to be asymptotically exact).
3. In order of both increasing coverage probability and increasing expected half-length of the confidence intervals, the procedures were almost always ranked as follows: CCV, NSC, and SC3.
4. In order of increasing point-estimator MSE, the procedures were almost always ranked as follows: CCV, NSC, and SC3.

Although the SC3 procedure does not appear to completely alleviate the loss of coverage in cases of extreme nonnormality, it consistently proved to be the most robust with respect to coverage probability in the examples presented here and in further Monte Carlo experiments we have conducted. The price paid was a slight increase in expected confidence-interval half-length and point-estimator MSE. When the controls are expected to be strongly correlated with the response (that is, for cases of large relative dominance in our experiments), one may expect all control-variate-based procedures to achieve large reductions in expected confidence-interval half-length and point-estimator MSE when compared to the direct-simulation procedure; and in

these situations the price seems to be well justified. In such cases, the SC3 procedure should be preferred, followed by NSC and CCV. When the controls are not expected to be strongly correlated with the response (cases of low relative dominance in our experiments), one may expect all control-variate-based procedures to achieve small MSE reductions and nearly nominal coverage probabilities; and in such cases, the CCV procedure should be preferred, followed by NSC and SC3.

An additional appealing feature of the SC3 procedure is the unbiasedness not only of the point estimator $\hat{\theta}_{\text{SP}}$ but also of the associated variance estimator S_{SP}^2/n , irrespective of any distributional assumptions. We would recommend SC3 for situations where joint normality of the response and the controls is suspect, especially when the controls are expected to be strongly correlated with the response and the sample size is relatively small.

Acknowledgments

This paper is partially based upon work supported by the National Science Foundation under Grant No. DMS-8717799. The U.S. Government has certain rights in this material. The authors thank George Fishman, David Goldsman, Barry Nelson, Bruce Schmeiser, and the anonymous referees for many constructive comments and suggestions.

References

- [1] T. W. Anderson, *An Introduction to Multivariate Statistical Analysis*, second ed., Wiley, New York, 1984.
- [2] J. M. Antill and R.W. Woodhead, *Critical Path Methods in Construction Practice*, Wiley, New York, 1982.
- [3] A. N. Avramidis, “Variance reduction techniques for simulation with applications to stochastic networks”, Ph.D. Thesis, Purdue University, School of Industrial Engineering, West Lafayette, IN, 1993.
- [4] A. N. Avramidis, K. W. Bauer, and J. R. Wilson, “Simulation of stochastic activity networks using path control variates”, *Naval Research Logistics* **38**, 183–201 (1991).

- [5] A. N. Avramidis and J. R. Wilson, “Control variates for stochastic network simulation”, in: O. Balci, R. P. Sadowski, and R. E. Nance (eds.), *Proceedings of the 1990 Winter Simulation Conference*, Institute of Electrical and Electronics Engineers, Piscataway, NJ, 1990, 323–332.
- [6] P. Billingsley, *Convergence of Probability Measures*, Wiley, New York, 1968.
- [7] P. Billingsley, *Probability and Measure*, second ed., Wiley, New York, 1986.
- [8] G. S. Fishman, *Concepts and Methods in Discrete Event Digital Simulation*, Wiley, New York, 1973.
- [9] G. Hadley, *Linear Algebra*, Addison-Wesley, Reading, MA, 1961.
- [10] S. S. Lavenberg, T. L. Moeller, and P. D. Welch, “Statistical results on control variables with application to queueing network simulation”, *Operations Research* **30**, 182–202 (1982).
- [11] J. L. McKenney and R. S. Rosenbloom, *Cases in Operations Management*, Wiley, New York, 1969.
- [12] B. L. Nelson, “Control variate remedies”, *Operations Research* **38**, 974–992 (1990).
- [13] B. D. Ripley, *Stochastic Simulation*, Wiley, New York, 1987.
- [14] R. J. Serfling, *Approximation Theorems of Mathematical Statistics*, Wiley, New York, 1980.
- [15] A. Stuart and J. K. Ord, *Kendall’s Advanced Theory of Statistics, Volume 1: Distribution Theory*, fifth ed., Oxford University Press, New York, 1987.
- [16] K. D. Tocher, *The Art of Simulation*, The English Universities Press, London, 1963.