S.G. Henderson and B.L. Nelson (Eds.), *Handbook in OR & MS, Vol. 13*

DOI: 10.1016/S0927-0507(06)13015-7

Chapter 15

Correlation-Based Methods for Output Analysis

David Goldsman

School of Industrial and Systems Engineering, Georgia Institute of Technology, USA
E-mail: sman@isye.gatech.edu

Barry L. Nelson

Department of Industrial Engineering and Management Sciences, Northwestern University, USA
E-mail: nelsonb@northwestern.edu

Abstract

We describe methods for estimating the variance of the sample mean of a steady-state simulation output process using correlation-based methods, all involving some form of "batching". Since no single method dominates all others across all measures of performance, we provide a framework for asymptotic comparison of such estimators. Research challenges are also identified.

1 Introduction

The input processes driving a simulation are usually random variables – examples include interarrival times of customers, resource service times, and machine breakdown times. Random input means that the output from the simulation is also apt to be random – for instance, customer waiting times, resource utilizations, and product cycle times. This in turn implies that runs of the simulation only yield *estimates* of the true measures of system performance – e.g., the mean customer waiting time, the long-run utilization, or the 80th percentile of product cycle time. Since these estimators are themselves random variables, they are therefore subject to sampling error which must be taken into account in a rigorous way if we are to make valid inferences or decisions concerning the performance of the underlying system. Thus, as part of a complete simulation study, we should always carry out a careful statistical analysis of the simulation's output.

The fundamental problem is that simulations almost never produce raw output that is independent and identically distributed (i.i.d.), much less normally distributed. For example, consecutive customer waiting times from a complicated queueing system pose a number of difficulties that hinder analysis via elementary statistical methods:

- Simulation data are not independent – typically, they are serially correlated. If one customer at the post office waits in line a long time, then the next customer is also likely to wait a long time.
- Simulation data are not identically distributed. Customers showing up early in the morning might have a much shorter wait than those who show up just before closing time.
- Simulation data are not normally distributed – waiting-time distributions are usually skewed to the right (and are certainly never less than zero).

These facts of life render it challenging to apply "classical" statistical techniques to the analysis of simulation output. So our purpose in this chapter is to present methods for statistical analysis of output from discrete-event computer simulations. We focus our discussion on data arising from *steady-state* (nonterminating) simulations, where the interest lies in the long-run behavior of the system. Presumedly this steady-state behavior is independent of the simulation's initial conditions. An example is a continuously running production line for which the experimenter is interested in some long-run performance measure. See Chapters 1 and 2 for examples and a precise definition of the "steady-state simulation" problem.

The main concern of the present chapter lies in studying methods for evaluating the variance of estimators derived from steady-state simulation output data. In particular, we concentrate on so-called "correlation-based" techniques, where we have knowledge of – and take advantage of – the correlation structure of certain stochastic processes. For a general introduction to estimation methods for simulation see Chapter 8.

Of course, many simulations are not run in steady state, e.g., the simulation of the ebb and flow of customers in a bank that opens and closes at certain times every day. Such *transient* (terminating) simulations use output analysis techniques that are primarily based on the method of independent replications (see Law and Kelton, 2000 and Chapter 8).

We henceforth assume access to steady-state data $Y_1, Y_2, \ldots, Y_n$, where Y_i might represent, e.g., the ith customer's waiting time in a specified queue. In this case, a good analysis might start off with, at the very least, an estimate of the unknown mean μ of the steady-state output process. Of course, the sample mean $\overline{Y}_n$ is the usual estimator for μ; but since the sample mean is a random variable, the experimenter should estimate the sample mean's variability as well. One such measure is simply $\sigma_n^2 \equiv n\,\mathrm{Var}(\overline{Y}_n)$, or almost equivalently, the *variance parameter*, $\sigma^2 \equiv \lim_{n\to\infty} \sigma_n^2$. Owing to the fact that outputs from steady-state simulations are almost never i.i.d., the "standard" sample-variance

estimator for σ^2 is almost guaranteed to be biased (usually on the low side for queueing simulations), and should therefore never be used.

The remainder of this chapter addresses a number of valid methods for estimating σ_n^2 or σ^2. There are many techniques in the literature concerning this key task. We begin in Section 2 with some additional motivational comments and basic material to keep the discussion self-contained. In Section 3, we devote much our time to a tutorial on the well-known methods of nonoverlapping batch means (NBM) and standardized time series (STS); afterwards, Section 4 introduces overlapping batch means (OBM) estimators and overlapping versions of the STS estimators. These, along with other popular techniques such as spectral analysis, regeneration, and autoregressive modeling, are prescribed in Chapter 16 and standard simulation texts such as Bratley et al. (1987) and Law and Kelton (2000). Finally, Section 5 gives a summary, conclusions, and some suggestions for future research areas.

2 Motivation

In this chapter, we consider a stationary stochastic process $Y_1, Y_2, \ldots, Y_n$, e.g., a simulation in steady state. In fact, we assume that the stochastic process satisfies a Functional Central Limit Theorem (FCLT). This assumption applies to a general class of stochastic processes, and will help give us the limiting properties of the various variance estimators considered herein.

Assumption 1 (FCLT). There exist constants μ and positive σ such that as $n \to \infty$,

$$X_n \Rightarrow \sigma \mathcal{W},$$

where $\mathcal{W}$ is a standard Brownian motion process, "$\Rightarrow$" denotes weak convergence as $n \to \infty$ (see Billingsley, 1968) and

$$X_n(t) \equiv \frac{\lfloor nt \rfloor (\overline{Y}_{\lfloor nt \rfloor} - \mu)}{\sqrt{n}} \quad \text{for } t \geqslant 0,$$

where $\overline{Y}_j \equiv \sum_{k=1}^{j} Y_k/j$, $j = 1, 2, \ldots$, and $\lfloor \cdot \rfloor$ is the greatest integer function.

Glynn and Iglehart (1990) list several different sets of sufficient conditions – usually in the form of moment and mixing conditions – for Assumption 1 to hold; see also Chapter 2. The constants μ and σ^2 in the assumption can be identified with the steady-state mean and variance parameters, respectively.

Let $R_k \equiv \text{Cov}(Y_1, Y_{1+k})$, $k = 0, \pm 1, \pm 2, \ldots$, denote the covariance function of the stationary stochastic process, and define the "center of gravity" quantity $\gamma \equiv -2 \sum_{k=1}^{\infty} k R_k$ (cf. Song and Schmeiser, 1995). In addition, the notation $p(n) = o(q(n))$ means that $p(n)/q(n) \to 0$ as $n \to \infty$. This is notation we will need later.

The rest of this chapter is devoted to motivating, defining and comparing classes of estimators of σ_n^2 and σ^2. To facilitate discussing what makes an estimator good, let $\hat{\sigma}_n^2$ and $\hat{\sigma}^2$ denote generic estimators of σ_n^2 and σ^2, respectively. In this section we focus on $\hat{\sigma}^2$ for convenience, but nearly everything applies to $\hat{\sigma}_n^2$ as well.

The most critical properties of $\hat{\sigma}^2$ are its variance and bias: $\text{Var}(\hat{\sigma}^2) = \text{E}\{(\hat{\sigma}^2 - \text{E}[\hat{\sigma}^2])^2\}$ and $\text{Bias}(\hat{\sigma}^2) = \text{E}[\hat{\sigma}^2] - \sigma^2$. Clearly an estimator with low bias – which implies that its distribution is centered at the desired value σ^2 – and low variance – so that it also tends to be close to the desired value – is what we want. A single measure that combines both measures is the mean squared error

$$\text{MSE}(\hat{\sigma}^2) = \text{Var}(\hat{\sigma}^2) + \text{Bias}^2(\hat{\sigma}^2).$$

As will become apparent when we look at different classes of estimators, there is often a bias-variance tradeoff so that doing something that decreases one of the measures increases the other. So why not simply use the variance estimator $\hat{\sigma}^2$ offering the lowest MSE? The problem is that it is not possible to derive a useful expression for the MSE of any reasonable variance estimator applied to all stationary output process that we might encounter in real life. As a result, the analysis and comparison of variance estimators typically takes one or more of the following forms:

Empirical evaluation: Variance estimators are applied to simulated processes with *known* σ^2's, allowing their bias, variance and MSE to be estimated. Since we "know the answer", we can see how the competing variance estimators compare against each other. Simple Markovian queueing models are often used for this type of comparison (e.g., Sargent et al., 1992).

Surrogate-processes evaluation: Simple output-process models, under which the bias, variance, and MSE of $\hat{\sigma}^2$ can be explicitly derived, are used as surrogates (stand-ins) for the more-general simulation output processes we might encounter. Time-series processes, such as low-order autoregressive and moving average, are often chosen for this purpose (cf. Sargent et al., 1992).

Asymptotic analysis: The bias, variance, and MSE measures of appropriately scaled versions of $\hat{\sigma}^2$ can sometimes be derived as the sample size n goes to infinity, and these limiting results are often free of the fine details of the specific simulation output process. We focus on this approach here by relying on the FCLT as our basic building block.

Unfortunately, it has never been proven, nor is it likely ever to be proven, that any one estimator has uniformly the smallest MSE across the space of all stationary output processes to which it could be applied.

The bias and variance of $\hat{\sigma}^2$ are the most critical distributional properties when determining its usefulness as an estimator. Secondarily, the distribution itself also matters, particularly when the goal is to use $\hat{\sigma}^2$ to form a confidence

interval for μ. The usual normal-theory $(1-\alpha)100\%$ confidence interval for μ takes the form

$$\overline{Y}_n \pm t_{1-\alpha/2,\nu}\sqrt{\frac{\hat{\sigma}^2}{n}}, \tag{1}$$

where $t_{1-\alpha/2,\nu}$ is the $1-\alpha/2$ quantile of the t distribution with ν degrees of freedom. The validity of this confidence interval depends on $\overline{Y}_n$ being approximately normally distributed and independent of $\hat{\sigma}^2$, which has a scaled chi-squared distribution with ν degrees of freedom. Therefore, if $\hat{\sigma}^2$ is to be incorporated into a confidence-interval procedure it is important to have some assurance that its distribution is approximately chi-squared and an appropriate degrees of freedom can be associated with it. For all practical purposes this property *never* precisely holds for steady-state simulation output processes, so the that best we can hope is that it is true in some appropriate asymptotic sense (cf. Alexopoulos et al., 2005a, who present histograms of the distributions of various variance estimators).

This chapter emphasizes asymptotic analysis, which has been exceptionally useful for deriving and comparing classes of estimators. But the estimator with the best asymptotic properties is not necessarily the best estimator in practice, because "practice" can include problems with sample sizes too small for the asymptotic properties of all available estimators to hold. Certain estimators with inferior asymptotic properties have been shown to be robust in practical problems. Nevertheless, asymptotic comparisons provide a level playing field at least for initial evaluation, as well as providing hints for how to improve estimator performance.

3 Estimators using nonoverlapping batches

This section examines a number of different estimators for σ^2 resulting from nonoverlapping batches of observations. Loosely speaking, this section examines estimators arising from the following recipe:

- Divide the run into contiguous, nonoverlapping *batches*,
- form an estimator from each batch, and
- take the average of the estimators.

In particular, we discuss the NBM, STS batched area, and STS batched Cramér–von Mises (CvM) estimators for σ^2 in Sections 3.1, 3.3 and 3.4, respectively; Section 3.2 gives a short primer on standardized time series. In each case, we give results on the expected value and variance of the estimator under consideration.

Throughout the entire section we will work with b contiguous, nonoverlapping batches of observations, each of length m, from the simulation output, $Y_1, Y_2, \ldots, Y_n$, where $n = bm$. Thus, the observations $Y_{(i-1)m+1}$, $Y_{(i-1)m+2}, \ldots, Y_{im}$ constitute batch i, for $i = 1, 2, \ldots, b$.

3.1 NBM estimator

The batches of observations can be depicted as follows.

$$\begin{aligned}
&\text{Batch 1: } Y_1, Y_2, \ldots, Y_m,\\
&\text{Batch 2: } Y_{m+1}, Y_{m+2}, \ldots, Y_{2m},\\
&\quad\vdots\\
&\text{Batch } b\text{: } Y_{(b-1)m+1}, Y_{(b-1)m+2}, \ldots, Y_n.
\end{aligned}$$

For each of these batches, we calculate the *batch mean*, $\overline{Y}_{i,m} \equiv m^{-1} \times \sum_{k=1}^{m} Y_{(i-1)m+k}$, for $i = 1, 2, \ldots, b$, which is the genesis of the method's name.

The NBM estimator for μ is simply the grand sample mean from the b batch means, $\overline{Y}_n \equiv b^{-1} \sum_{i=1}^{b} \overline{Y}_{i,m} = n^{-1} \sum_{\ell=1}^{n} Y_\ell$. Stationarity implies that $\mathrm{E}[\overline{Y}_n] = \mu$, so the grand mean is unbiased for μ; and the variance of the grand mean is, by definition, $\mathrm{Var}(\overline{Y}_n) = \sigma_n^2/n$.

The batch means $\overline{Y}_{i,m}$, $i = 1, 2, \ldots, b$, are often assumed to be i.i.d. normal random variables, at least for large enough batch size m; this is borne out by Equation (8) below. The i.i.d. assumption immediately suggests the NBM estimator for σ^2,

$$\mathcal{N}(b, m) \equiv \frac{m}{b-1} \sum_{i=1}^{b} (\overline{Y}_{i,m} - \overline{Y}_n)^2 \xrightarrow{\mathcal{D}} \frac{\sigma^2 \chi^2_{b-1}}{b-1} \tag{2}$$

as $m \to \infty$ with b fixed, and where χ^2_ν denotes a chi-squared random variable with ν degrees of freedom and "$\xrightarrow{\mathcal{D}}$" denotes convergence in distribution as $m \to \infty$ (see, e.g., Glynn and Whitt, 1991; Schmeiser, 1982; Steiger and Wilson, 2001). The NBM estimator, which is the sample variance of the batch means, is one of the most popular for σ^2, and is a benchmark for comparison with other estimators.

Under mild conditions, Chien et al. (1997), Goldsman and Meketon (1986) and Song and Schmeiser (1995) show that

$$\mathrm{E}\big[\mathcal{N}(b, m)\big] = \sigma^2 + \frac{\gamma(b+1)}{bm} + \mathrm{o}\left(\frac{1}{m}\right). \tag{3}$$

So as the batch size m increases, the bias of $\mathcal{N}(b, m)$ as an estimator of σ^2 goes to zero.

How does one prove a result such as Equation (3) for the expected value? It amounts to careful bookkeeping of covariance terms. First of all, assuming that the underlying process $\{Y_i\}$ is stationary and that all of the following sums

are well defined, we have

$$\begin{aligned}
\sigma_m^2 &\equiv m\,\mathrm{Var}(\overline{Y}_m) \\
&= \frac{1}{m}\sum_{i=1}^{m}\sum_{j=1}^{m}\mathrm{Cov}(Y_i, Y_j) \\
&= R_0 + 2\sum_{i=1}^{m-1}\left(1-\frac{i}{m}\right)R_i \quad \text{(after collecting like covariance terms)} \\
&= R_0 + 2\sum_{i=1}^{m-1} R_i - \frac{2}{m}\sum_{i=1}^{m-1} iR_i \\
&= R_0 + 2\sum_{i=1}^{\infty} R_i - 2\sum_{i=m}^{\infty} R_i - \frac{2}{m}\sum_{i=1}^{\infty} iR_i + \frac{2}{m}\sum_{i=m}^{\infty} iR_i \\
&= \sigma^2 + \frac{\gamma}{m} - 2\sum_{i=m}^{\infty}\left(1-\frac{i}{m}\right)R_i \\
&= \sigma^2 + \frac{\gamma}{m} + \mathrm{o}\left(\frac{1}{m}\right), \qquad (4)
\end{aligned}$$

where, as in Chien et al. (1997), we have implicitly assumed that $\sum_{j=1}^{\infty} j|R_j| < \infty$. Turning to the matter at hand, we have

$$\begin{aligned}
\mathrm{E}\big[\mathcal{N}(b,m)\big] &= \mathrm{E}\left[\frac{m}{b-1}\sum_{i=1}^{b}(\overline{Y}_{i,m}-\overline{Y}_n)^2\right] \\
&= \frac{m}{b-1}\,\mathrm{E}\left[\sum_{i=1}^{b}\overline{Y}_{i,m}^2 - b\overline{Y}_n^2\right] \\
&= \frac{m}{b-1}\left[\sum_{i=1}^{b}\mathrm{E}\big[\overline{Y}_{i,m}^2\big] - b\mathrm{E}\big[\overline{Y}_n^2\big]\right] \\
&= \frac{bm}{b-1}\big[\mathrm{E}\big[\overline{Y}_{i,m}^2\big] - \mathrm{E}\big[\overline{Y}_n^2\big]\big] \\
&\qquad \text{(by stationarity of the batch means)} \\
&= \frac{bm}{b-1}\big[\mathrm{Var}(\overline{Y}_{i,m}) - \mathrm{Var}(\overline{Y}_n)\big] \\
&\qquad \text{(since the common mean } \mu \text{ cancels)} \\
&= \frac{1}{b-1}\left[b\sigma^2 + \frac{b\gamma}{m} - \sigma^2 - \frac{\gamma}{n}\right] + \mathrm{o}\left(\frac{1}{m}\right) \\
&\qquad \text{(by Equation (4)),}
\end{aligned}$$

which gives (3) after a little algebra.

Obtaining the NBM estimator's variance takes more work. Chien et al. (1997), Glynn and Whitt (1991), Goldsman and Meketon (1986) and Song and Schmeiser (1995) find that, for fixed b,

$$\lim_{m\to\infty}(b-1)\,\mathrm{Var}\big(\mathcal{N}(b,m)\big) = 2\sigma^4 \tag{5}$$

which certainly makes sense in light of the distributional result given by Equation (2).

By combining Equations (3) and (5), we can obtain a simple expression for the mean squared error of the NBM estimator for σ^2,

$$\mathrm{MSE}\big[\mathcal{N}(b,m)\big] \doteq \left(\frac{\gamma(b+1)}{mb}\right)^2 + \frac{2\sigma^4}{b}, \tag{6}$$

where we have ignored some small-order terms. For large b and m, this expression can be minimized by taking the number of batches $b = cn^{2/3}$ and the batch size $m = n/b = n^{1/3}/c$, where we need to determine the constant c. Thus, expression (6) becomes

$$\mathrm{MSE}\big[\mathcal{N}(b,m)\big] \doteq \left(c^2\gamma^2 + \frac{2\sigma^4}{c}\right)\frac{1}{n^{2/3}}.$$

Minimizing this expression with respect to c, we obtain $c^\star = (\sigma^4/\gamma^2)^{1/3}$, and hence an (asymptotically) optimal mean squared error of

$$\mathrm{MSE}^\star\big[\mathcal{N}(b,m)\big] \doteq 3\left(\frac{\sigma^4\gamma}{n}\right)^{2/3}. \tag{7}$$

Of course, σ^2 and γ are not known in general, so that it will be problematic to actually obtain this minimal MSE in practice; however, Equation (7) still has practical value, since it can be used as a basis for comparison among other variance estimators. For instance, if another estimator has a minimal MSE of the same form as (7), but with a leading constant smaller than 3, then one can claim that it is a "better" estimator than NBM.

3.2 *STS primer*

Before presenting additional estimators for σ^2, we give a mini-tutorial on standardized time series, which will provide the necessary background. Schruben (1983) defines the *standardized time series* from batch i as

$$T_{i,m}(t) \equiv \frac{\lfloor mt\rfloor(\overline{Y}_{i,\lfloor mt\rfloor} - \overline{Y}_{i,m})}{\sigma\sqrt{m}}$$

for $0 \leqslant t \leqslant 1$ and $i = 1, 2, \ldots, b$, where

$$\overline{Y}_{i,j} \equiv \frac{1}{j}\sum_{k=1}^{j} Y_{(i-1)m+k}$$

for $j = 1, 2, \ldots, m$ and $i = 1, 2, \ldots, b$. Then we have the following theorem.

Theorem 1 (see Schruben (1983), Glynn and Iglehart (1990), Foley and Goldsman (1999), or Alexopoulos et al. (2005b), among others). *Define* $Z_i(m) \equiv \sqrt{m}(\overline{Y}_{i,m} - \mu)$, $i = 1, 2, \ldots, b$. *Then under Assumption* 1,

$$\begin{aligned}&\big(Z_1(m), \ldots, Z_b(m); \sigma T_{1,m}, \ldots, \sigma T_{b,m}\big)\\ &\quad\Rightarrow \big(\sigma Z_1, \ldots, \sigma Z_b; \sigma \mathcal{B}_0, \ldots, \sigma \mathcal{B}_{b-1}\big),\end{aligned} \tag{8}$$

where the Z_i*'s are i.i.d. standard normal random variables, and* $\mathcal{B}_u$ *denotes a Brownian bridge process on* $[u, u + 1]$, *i.e., for* $t \in [0, 1]$,

$$\mathcal{B}_u(t) = \mathcal{W}(u + t) - \mathcal{W}(u) - t\big[\mathcal{W}(u + 1) - \mathcal{W}(u)\big].$$

Notice that the scaling factor σ actually cancels out in the denominator of $T_{i,m}$, thus eliminating any unknowns and so allowing us to explicitly calculate all of the quantities on the left-hand side of Equation (8). Further, for those unfamiliar with Brownian bridges, it turns out that all finite-dimensional joint distributions of a Brownian bridge $\mathcal{B}_u$ are normal with $\mathrm{E}[\mathcal{B}_u(t)] = 0$ and $\mathrm{Cov}(\mathcal{B}_u(s), \mathcal{B}_u(t)) = \min(s, t) - st$, $0 < s, t < 1$. In addition, since Brownian motion has independent increments, it is easy to see that $\mathcal{B}_0, \mathcal{B}_1, \ldots, \mathcal{B}_{b-1}$ are independent Brownian bridges. These facts will allow us to evaluate properties of the random variables on the left-hand side of (8) using the analogous asymptotic distributions from the right-hand side of the equation.

3.3 Batched area estimator

This subsection deals with the (nonoverlapping) batched area estimator for σ^2 (Goldsman et al., 1990; Goldsman and Schruben, 1990).

We will work with the square of the weighted area under the standardized time series from the ith batch,

$$A_i(f; m) \equiv \left[\frac{1}{m}\sum_{k=1}^{m} f\left(\frac{k}{m}\right)\sigma T_{i,m}\left(\frac{k}{m}\right)\right]^2,$$

and its limiting functional

$$A_i(f) \equiv \left[\int_0^1 f(t)\sigma\mathcal{B}_{i-1}(t)\,\mathrm{d}t\right]^2,$$

for $i = 1, 2, \ldots, b$, where $f(t)$ is continuous on the interval $[0, 1]$ and normalized so that $\mathrm{Var}(\int_0^1 f(t)\mathcal{B}_0(t)\,\mathrm{d}t) = 1$. Under mild conditions (see Alexopoulos et al., 2005b), one can use the continuous mapping theorem (see Billingsley, 1968) to show that $A_i(f; m) \xrightarrow{\mathcal{D}} A_i(f)$, $i = 1, 2, \ldots, b$; and further, $A_1(f), \ldots, A_b(f)$ are i.i.d. $\sigma^2\chi^2_1$. This result motivates construction of the

batched area estimator for σ^2,

$$\mathcal{A}(f; b, m) \equiv \frac{1}{b} \sum_{i=1}^{b} A_i(f; m). \tag{9}$$

As the batch size $m \to \infty$ with a fixed number of batches b, the distribution of the batched area estimator converges to that of the average of the corresponding $A_i(f)$ functionals, i.e.,

$$\mathcal{A}(f; b, m) \xrightarrow{\mathcal{D}} \mathcal{A}(f; b) \equiv \frac{1}{b} \sum_{i=1}^{b} A_i(f) \sim \sigma^2 \frac{\chi_b^2}{b}.$$

The next theorem gives the expected value and variance of the area estimator.

Theorem 2 (see, e.g., Foley and Goldsman, 1999). *Suppose that* $\{Y_i, i \geqslant 1\}$ *is a stationary process for which Assumption* 1 *holds,* $\sum_{k=1}^{\infty} k^2 |R_k| < \infty$, *and* $\sigma^2 > 0$. *Further, suppose that* $\mathcal{A}^2(f; b, m)$ *is uniformly integrable* (*cf.* Billingsley, 1968). *If we define the quantities* $F \equiv \int_0^1 f(t)\,\mathrm{d}t$, $\overline{F} \equiv \int_0^1 \int_0^t f(s)\,\mathrm{d}s\,\mathrm{d}t$ *and* $F^\star \equiv [(F - \overline{F})^2 + \overline{F}^2]/2$, *then*

$$\mathrm{E}\big[\mathcal{A}(f; b, m)\big] = \sigma^2 + \frac{F^\star \gamma}{m} + \mathrm{o}\left(\frac{1}{m}\right)$$

and

$$\mathrm{Var}\big(\mathcal{A}(f; b, m)\big) \to \mathrm{Var}\left(\sigma^2 \frac{\chi_b^2}{b}\right) = 2\frac{\sigma^4}{b}$$

as $m \to \infty$. *Note that the limiting variance does not depend on the choice of the weighting function (as long as it is legal).*

Example 1. Schruben (1983) first considered the area estimator with constant weighting function $f_0(t) \equiv \sqrt{12}$ for $0 \leqslant t \leqslant 1$. For this choice, Theorem 2 implies that $\mathrm{E}[\mathcal{A}(f_0; b, m)] = \sigma^2 + 3\gamma/m + \mathrm{o}(1/m)$.

Example 2. If the selection of $f(t)$ turns out to give $F = \overline{F} = 0$, the resulting estimator is *first-order unbiased* for σ^2, i.e., its bias is $\mathrm{o}(1/m)$. An example of such a weighting function is the quadratic $f_2(t) \equiv \sqrt{840}(3t^2 - 3t + 1/2)$ (Goldsman et al., 1990; Goldsman and Schruben, 1990).

Example 3. Foley and Goldsman (1999) give an "orthonormal" sequence of first-order unbiased weights, $f_{\cos,j}(t) = \sqrt{8}\pi j \cos(2\pi j t)$, $j = 1, 2, \ldots$. It can be shown that the orthonormal estimators' limiting functionals $A_i(f_{\cos,1})$, $A_i(f_{\cos,2}), \ldots$ are i.i.d. $\sigma^2 \chi_1^2$. Thus, we have a new estimator for σ^2, since as

the batch size m becomes large,

$$\frac{1}{k}\sum_{j=1}^{k} A_i(f_{\cos,j}; m) \xrightarrow{\mathcal{D}} \sigma^2 \frac{\chi_k^2}{k}, \tag{10}$$

for some "reasonable" number of orthonormal weights k. This estimator is actually the average of different weighted area estimators from the *same batch*, whereas the estimator given in Equation (9) uses observations from *all batches*. Of course, one could also average estimators of the form given in Equation (10) over all batches to obtain even more degrees of freedom, but that estimator's performance properties have not yet been thoroughly evaluated.

Remark 1. Since the Z_i's and $\mathcal{B}_i$'s in Equation (8) turn out to be uncorrelated normal random variables, the standardized sample mean is asymptotically (as $m \to \infty$) independent of the standardized time series (Schruben, 1983). One can use this fact to argue that STS estimators such as $\mathcal{A}(f; b, m)$ are asymptotically independent of the NBM estimator $\mathcal{N}(b, m)$; and then we can take an appropriate linear combination of the STS and NBM estimators, e.g.,

$$\frac{b\mathcal{A}(f; b, m) + (b-1)\mathcal{N}(b, m)}{2b-1},$$

with the intent that the combined estimator will have more degrees of freedom – and so reduce variance without affecting bias too much.

3.4 Batched CvM estimator

This subsection discusses the weighted Cramér–von Mises estimator for σ^2 (see Goldsman et al., 1999). To begin with, the weighted area under the square of the STS from the ith batch and its limiting functional are given by

$$C_i(g; m) \equiv \frac{1}{m}\sum_{k=1}^{m} g\left(\frac{k}{m}\right)\sigma^2 T_{i,m}^2\left(\frac{k}{m}\right)$$

and

$$C_i(g) \equiv \int_0^1 g(t)\sigma^2 \mathcal{B}_{i-1}^2(t)\,\mathrm{d}t,$$

respectively. Here $g(t)$ is a weighting function normalized so that $\mathrm{E}[C_i(g)] = \sigma^2$ and possessing a continuous and bounded second derivative on $[0, 1]$.

Under mild assumptions, the continuous mapping theorem implies that $C_i(g; m) \xrightarrow{\mathcal{D}} C_i(g)$, $i = 1, 2, \ldots, b$. This leads to the *batched CvM* estimator for σ^2

$$\mathcal{C}(g; b, m) \equiv \frac{1}{b}\sum_{i=1}^{b} C_i(g; m).$$

As the batch size $m \to \infty$ with a fixed number of batches b, the distribution of the batched CvM estimator converges to that of the average of the corresponding $C_i(g)$ functionals, i.e.,

$$\mathcal{C}(g; b, m) \xrightarrow{\mathcal{D}} \mathcal{C}(g; b) \equiv \frac{1}{b} \sum_{i=1}^{b} C_i(g).$$

An immediate question to ask is: Why should we bother with yet another estimator $\mathcal{C}(g; b, m)$ for σ^2? Theorem 3 prepares us for an answer by presenting results on the expected value and variance of the weighted CvM estimator.

Theorem 3 (Goldsman et al., 1999). *Define* $G \equiv \int_0^1 g(t)\,\mathrm{d}t$. *Under conditions similar to those of Theorem 2,*

$$\mathrm{E}\big[\mathcal{C}(g; b, m)\big] = \sigma^2 + \frac{\gamma}{m}(G - 1) + \mathrm{o}\left(\frac{1}{m}\right)$$

and, for fixed b,

$$\begin{aligned} \lim_{m\to\infty} b\,\mathrm{Var}\big(\mathcal{C}(g; b, m)\big) &= b\,\mathrm{Var}\big(\mathcal{C}(g; b)\big) \\ &= \mathrm{Var}\big(C_1(g)\big) \\ &= 4\sigma^4 \int_0^1 g(t)(1-t)^2 \int_0^t g(s)s^2\,\mathrm{d}s\,\mathrm{d}t. \end{aligned} \tag{11}$$

Example 4. Consider the constant weighting function $g_0(t) \equiv 6$ for $0 \leqslant t \leqslant 1$. Theorem 3 shows that the resulting CvM estimator has expected value $\mathrm{E}[\mathcal{C}(g_0; b, m)] = \sigma^2 + 5\gamma/m + \mathrm{o}(1/m)$.

Example 5. Suppose we select a weighting function having $G = 1$ (in addition to the normalizing and derivative constraints). Then the theorem implies that the CvM estimator $\mathcal{C}(g; b, m)$ has bias $\mathrm{o}(1/m)$. For example, the quadratic weighting function $g_2^\star(t) \equiv -24 + 150t - 150t^2$ has this first-order unbiasedness property.

Now, back to the question concerning our interest in yet another estimator for σ^2. We see from Theorem 3 that the choice of weighting function $g(t)$ affects the variances of $\mathcal{C}(g; b, m)$ and $\mathcal{C}(g; b)$. This was not the case for the area estimator of Section 3.3, where the weighting function $f(t)$ affects the variance of $\mathcal{A}(f; b, m)$, but not that of the limiting functional $\mathcal{A}(f; b)$, which is always $\mathrm{Var}(\mathcal{A}(f; b)) = 2\sigma^4/b$. Thus, the CvM estimator gives us a tool to find reduced-variance estimators for σ^2.

Example 6. Theorem 3 shows that $\mathrm{Var}(\mathcal{C}(g_0; b)) = 0.8\sigma^4/b$ and $\mathrm{Var}(\mathcal{C}(g_2^\star; b)) = 1.73\sigma^4/b$, which are both smaller than the limiting ($m \to \infty$) variance

of any batched area estimator, as discussed immediately above. Even though $\mathrm{Var}(\mathcal{C}(g_2^\star; b)) > \mathrm{Var}(\mathcal{C}(g_0; b))$, the estimator $\mathcal{C}(g_2^\star; b, m)$ is first-order unbiased for σ^2, while $\mathcal{C}(g_0; b, m)$ does not have this nice property.

3.5 Comparison

We have seen that as the batch size $m \to \infty$, the NBM, batched area, and batched CvM estimators are all asymptotically unbiased for σ^2. Furthermore, the variances of these estimators are all more-or-less inversely proportional to the number of batches – though one has to be careful not to decrease the variance too much at the expense of relatively high bias or MSE (cf. Song and Schmeiser, 1995).

What happens if m and b *both* become large? Such a scenario might occur in sequential estimation procedures – for example, Steiger et al.'s (2005) ASAP3 procedure – which occasionally take additional observations to achieve user-specified estimator precision requirements. As m and $b \to \infty$, it has been shown in, e.g., Alexopoulos et al. (2000) and Chien et al. (1997) that the estimators under discussion herein are consistent in mean square.

Since NBM is regarded as the benchmark method, one could informally regard its bias and variance as straw men. The interesting result is that STS area and CvM estimators with certain well-chosen weighting functions can beat NBM in terms of large-sample bias; in addition, the CvM estimators studied in this article have smaller variance than does NBM. Better yet, we will see in the next section that the use of overlapping batches with respect to any particular estimator preserves its expected value, while reducing its variance – sometimes substantially. See Table 1 for an early preview.

Obviously, for fixed sample size $n = mb$, some estimators will tend to do better than others in terms of bias and variance. But for fixed n, decreasing one performance measure usually comes at the expense of increasing the other – the well-known trade-off that we have already mentioned. Thus, we ought to be interested in small-sample performance of the various estimators in addition to asymptotic performance. For various "toy" processes, it is actually possible to calculate exact results for finite sample sizes.

Example 7. Suppose that $Y_1, Y_2, \ldots, Y_n$ arise from a first-order moving average (MA(1)) process, that is, $Y_{i+1} = \theta\varepsilon_i + \varepsilon_{i+1}$, where the ε_i's are i.i.d. standard normal and $-1 < \theta < 1$. The MA(1) has covariance function $R_0 = 1 + \theta^2$, $R_{\pm 1} = \theta$, and $R_k = 0$ elsewhere, from which we can easily derive $\sigma^2 = \sum_{j=-\infty}^{\infty} R_j = (1+\theta)^2$ and $\gamma = -2\sum_{j=1}^{\infty} jR_j = -2\theta$.

After some tedious algebra (see, e.g., Goldsman et al., 2003), we have the following *exact* results:

$$\mathrm{E}\big[\mathcal{N}(b,m)\big] = \sigma^2 + \frac{\gamma(b+1)}{mb},$$
$$\mathrm{E}\big[\mathcal{A}(f_2;b,m)\big] = \sigma^2 + \frac{7(\sigma^2+6\gamma)}{2m^2} + \mathrm{o}\big(m^{-2}\big),$$
$$\mathrm{E}\big[\mathcal{C}\big(g_2^\star;b,m\big)\big] = \sigma^2 + \frac{4(\sigma^2+6\gamma)}{m^2} + \mathrm{o}\big(m^{-2}\big).$$

We see that these particular weighted area and CvM estimators are less biased than NBM as the batch size m becomes large; and the area estimator beats CvM for moderate m.

4 Estimators from overlapping batches

We now discuss the use of estimators based on *overlapping* batches, à la Meketon and Schmeiser (1984). Here we implement a slightly different recipe than that for nonoverlapping batches:

- Divide the run into a number of overlapping batches,
- form an estimator from each batch, and
- take the average of the estimators.

The presentation roughly follows that of the previous section. Section 4.1 discusses some necessary basics and provides the notation that we will use. Sections 4.2–4.4 give the OBM, STS overlapping area, and STS overlapping CvM estimators, respectively. As before, we are concerned with results on the expected values and variances of the various estimators.

4.1 Overlapping fundamentals

Suppose we have n observations $Y_1, Y_2, \ldots, Y_n$ on hand and that we form $n-m+1$ overlapping batches, each of size m. The overlapping batches can be depicted as follows.

$$\begin{aligned}
&\text{Overlapping batch 1:} \quad Y_1, Y_2, \ldots, Y_m,\\
&\text{Overlapping batch 2:} \quad Y_2, Y_3, \ldots, Y_{m+1},\\
&\qquad\vdots\\
&\text{Overlapping batch } n-m+1\text{:} \quad Y_{n-m+1}, Y_{n-m+2}, \ldots, Y_n
\end{aligned}$$

Specifically, the observations $Y_i, Y_{i+1}, \ldots, Y_{i+m-1}$ comprise batch i, $i = 1, 2, \ldots, n-m+1$. Further, we use the notation $b \equiv n/m$ as before, though b is no longer "the number of batches".

As in Section 3.2, the standardized time series from overlapping batch i is

$$T^{\mathrm{O}}_{i,m}(t) \equiv \frac{\lfloor mt \rfloor (\overline{Y}^{\mathrm{O}}_{i,\lfloor mt \rfloor} - \overline{Y}^{\mathrm{O}}_{i,m})}{\sigma\sqrt{m}}$$

for $0 \leqslant t \leqslant 1$ and $i = 1, 2, \ldots, n - m + 1$, where

$$\overline{Y}^{\mathrm{O}}_{i,j} \equiv \frac{1}{j} \sum_{k=0}^{j-1} Y_{i+k}$$

for $i = 1, 2, \ldots, n-m+1$ and $j = 1, 2, \ldots, m$. Under the same mild conditions as before,

$$\sigma T^{\mathrm{O}}_{\lfloor um \rfloor, m} \Rightarrow \sigma \mathcal{B}_u, \quad 0 \leqslant u \leqslant b - 1, \ u \text{ fixed}.$$

Roughly speaking, we will try to glean more information about σ^2 by considering the variance estimators from each of the $n - m + 1$ overlapping batches rather than from just the b nonoverlapping batches of the earlier discussion. The hope is that the additional information will somehow compensate for the fact that estimators arising from overlapping batches might be highly correlated.

4.2 OBM estimator

The ith overlapping batch mean is given by $\overline{Y}^{\mathrm{O}}_{i,m}$, $i = 1, 2, \ldots, n - m + 1$. The OBM estimator for σ^2 was originally studied by Meketon and Schmeiser (1984) (using a slightly different scaling coefficient), and is given by

$$\mathcal{O}(b, m) \equiv \frac{nm}{(n - m + 1)(n - m)} \sum_{i=1}^{n-m+1} \bigl(\overline{Y}^{\mathrm{O}}_{i,m} - \overline{Y}_n\bigr)^2.$$

Theorem 4. *Under mild conditions*, Goldsman and Meketon (1986) *and* Song and Schmeiser (1995) *show that*, *for large* b,

$$\mathrm{E}\bigl[\mathcal{O}(b, m)\bigr] \doteq \sigma^2 + \frac{\gamma}{m} + \mathrm{o}\left(\frac{1}{m}\right).$$

Further, Meketon and Schmeiser (1984), Damerdji (1995) *and* Alexopoulos et al. (2005b) *find that for large* b, *as* $m \to \infty$,

$$\mathrm{Var}\bigl(\mathcal{O}(b, m)\bigr) \doteq \frac{4\sigma^4}{3b}.$$

4.3 Overlapping area estimator

The square of the weighted area under the standardized time series from the ith overlapping batch is

$$A_i^{\mathrm{O}}(f;m) \equiv \left[\frac{1}{m}\sum_{k=1}^{m} f\left(\frac{k}{m}\right)\sigma T_{i,m}^{\mathrm{O}}\left(\frac{k}{m}\right)\right]^2,$$

$i = 1, 2, \ldots, n-m+1$. The *overlapping area estimator* for σ^2 is

$$\mathcal{A}^{\mathrm{O}}(f;b,m) \equiv \frac{1}{n-m+1}\sum_{i=1}^{n-m+1} A_i^{\mathrm{O}}(f;m).$$

Alexopoulos et al. (2005b) use the continuous mapping theorem to show that as $m \to \infty$,

$$\mathcal{A}^{\mathrm{O}}(f;b,m) \xrightarrow{\mathcal{D}} \mathcal{A}^{\mathrm{O}}(f;b) \equiv \frac{\sigma^2}{b-1}\int_0^{b-1}\left[\int_0^1 f(t)\mathcal{B}_u(t)\,\mathrm{d}t\right]^2 \mathrm{d}u. \tag{12}$$

It is easy to see that the expected value of the overlapping area estimator equals that of the corresponding batched area estimator. Thus, Theorem 2 gives the following.

Theorem 5. *Under mild conditions*, Alexopoulos et al. (2005a, 2005b) *show that*

$$\mathrm{E}\left[\mathcal{A}^{\mathrm{O}}(f;b,m)\right] = \sigma^2 + \frac{F^\star\gamma}{m} + \mathrm{o}\left(\frac{1}{m}\right).$$

Calculation of the variance of the overlapping area estimator can be undertaken using the right-hand side of Equation (12) along with some algebraic elbow grease. Some examples from Alexopoulos et al. (2005b) reveal that the limiting ($m \to \infty$) variance of the overlapping area estimator depends on the choice of weighting function.

Example 8. Consider the overlapping constant-weighted area estimator from Example 1. We have after some algebra that as $m \to \infty$,

$$\mathrm{Var}\left(\mathcal{A}^{\mathrm{O}}(f_0;b,m)\right) \to \mathrm{Var}\left(\mathcal{A}^{\mathrm{O}}(f_0;b)\right) = \frac{24b-31}{35(b-1)^2}\sigma^4.$$

This compares very nicely to the generic batched area estimator's asymptotic ($m \to \infty$) variance, $\mathrm{Var}(\mathcal{A}(f;b)) = 2\sigma^4/b$ (see Theorem 2).

Example 9. Consider the overlapping area estimator with first-order unbiased quadratic weighting function $f_2(t)$ from Example 2. This estimator has a limiting variance of

$$\operatorname{Var}\bigl(\mathcal{A}^{\mathrm{O}}(f_2; b)\bigr) = \frac{3514b - 4359}{4290(b-1)^2}\sigma^4.$$

Example 10. Consider the overlapping area estimators from the family of orthonormal first-order unbiased weights $f_{\cos,j}(t)$, $j = 1, 2, \ldots$. From Example 3, we have a limiting variance of about

$$\operatorname{Var}\bigl(\mathcal{A}^{\mathrm{O}}(f_{\cos,j}; b)\bigr) \doteq \frac{8\pi^2 j^2 + 15}{12\pi^2 j^2 b}\sigma^4.$$

Remark 2. One can average the orthonormal estimators $\mathcal{A}^{\mathrm{O}}(f_{\cos,j}; b, m)$, $j = 1, 2, \ldots$, and use knowledge of the covariances of these estimators to obtain estimators with even smaller variance (cf. Alexopoulos et al., 2005b).

4.4 Overlapping CvM estimator

We define the overlapping CvM estimator from overlapping batch i by

$$C_i^{\mathrm{O}}(g; m) \equiv \frac{1}{m}\sum_{k=1}^{m} g\left(\frac{k}{m}\right)\left[\sigma T_{i,m}^{\mathrm{O}}\left(\frac{k}{m}\right)\right]^2,$$

$i = 1, 2, \ldots, n - m + 1$. The *overlapping CvM* estimator for σ^2 is

$$\mathcal{C}^{\mathrm{O}}(g; b, m) \equiv \frac{1}{n-m+1}\sum_{i=1}^{n-m+1} C_i^{\mathrm{O}}(g; m).$$

Then it can be shown using the continuous mapping theorem that as $m \to \infty$,

$$\mathcal{C}^{\mathrm{O}}(g; b, m) \xrightarrow{\mathcal{D}} \mathcal{C}^{\mathrm{O}}(g; b) \equiv \frac{1}{b-1}\int_0^{b-1}\int_0^1 g(u)\sigma^2\mathcal{B}_s^2(u)\,\mathrm{d}u\,\mathrm{d}s. \tag{13}$$

Meanwhile, Theorem 3 implies

$$\mathrm{E}\bigl[\mathcal{C}^{\mathrm{O}}(g; b, m)\bigr] = \sigma^2 + \frac{\gamma(G-1)}{m} + \mathrm{o}\left(\frac{1}{m}\right).$$

So the expected value of the overlapping CvM estimator is the same as that of the batched CvM estimator.

Let us turn to the variance of the overlapping CvM estimator. After a great deal of algebra involving Equation (13), we have the following results.

Example 11. For the overlapping constant-weighted CvM estimator, we have

$$\mathrm{Var}\big(\mathcal{C}^{\mathrm{O}}(g_0; b, m)\big) \to \mathrm{Var}\big(\mathcal{C}^{\mathrm{O}}(g_0; b)\big) = \frac{88b - 115}{210(b-1)^2}\sigma^4.$$

This compares nicely to the batched constant-weighted CvM estimator's asymptotic ($m \to \infty$) variance, $\mathrm{Var}(\mathcal{C}(g_0; b)) = 4\sigma^4/(5b)$.

Example 12. For the overlapping CvM estimator with quadratic weight $g_2^\star(t)$ from Example 5, we have

$$\mathrm{Var}\big(\mathcal{C}^{\mathrm{O}}\big(g_2^\star; b\big)\big) = \frac{10768b - 13605}{13860(b-1)^2}\sigma^4 \doteq \frac{0.777}{b}\sigma^4.$$

This compares to the batched quadratic CvM estimator's asymptotic variance, $\mathrm{Var}(\mathcal{C}(g_2^\star; b)) = 121\sigma^4/(70b)$. In other words, $\mathrm{Var}(\mathcal{C}^{\mathrm{O}}(g_2^\star; b))/\mathrm{Var}(\mathcal{C}(g_2^\star; b)) \doteq 0.450$.

Example 13. We can even work with the quartic weighting function

$$g_4^\star(t) \equiv -\frac{1310}{21} + \frac{19270t}{21} - \frac{25230t^2}{7} + \frac{16120t^3}{3} - \frac{8060t^4}{3}.$$

Goldsman et al. (1999) show that this weighting function yields the first-order unbiased CvM estimator for σ^2 having the minimum variance over all quartic weights. After still more algebra involving Equation (13), we find that the corresponding overlapping quartic CvM estimator has

$$\mathrm{Var}\big(\mathcal{C}^{\mathrm{O}}\big(g_4^\star; b\big)\big) \doteq \frac{0.477}{b}\sigma^4,$$

which is quite competitive compared to the other estimators examined so far.

4.5 Comparison

Paralleling the discussion in Section 3.5, we see that as $m \to \infty$, the overlapping area, overlapping CvM, and OBM estimators are all asymptotically unbiased for σ^2. In addition, the variances of these estimators are all inversely proportional to the ratio $b = n/m$ (for sufficiently large batch size).

All of the overlapping estimators preserve the bias properties of their nonoverlapping counterparts. Thus, we found that the overlapping area and overlapping CvM estimators with certain "unbiased" weighting functions can beat OBM in terms of large-sample bias. The overlapping STS estimators also defeat their nonoverlapped counterparts as well as OBM in terms of variance. Table 1, abstracted from Alexopoulos et al. (2005a, 2005b), gives a more-complete synopsis of all of the asymptotic results.

We can carry out small-sample analysis for the various overlapping estimators, similar to what we did in Example 7.

Table 1.
Approximate asymptotic bias and variance for different estimators (abstracted from Alexopoulos et al., 2005a, 2005b)

Nonoverlapping	(m/γ) Bias	(b/σ^4) Var	Overlapping	(m/γ) Bias	(b/σ^4) Var
$\mathcal{A}(f;b,m)$	$F^\star$	2	$\mathcal{A}^{\mathrm{O}}(f;b,m)$	$F^\star$	various
$\mathcal{A}(f_0;b,m)$	3	2	$\mathcal{A}^{\mathrm{O}}(f_0;b,m)$	3	0.686
$\mathcal{A}(f_2;b,m)$	o(1)	2	$\mathcal{A}^{\mathrm{O}}(f_2;b,m)$	o(1)	0.819
$\mathcal{A}(f_{\cos,j};b,m)$	o(1)	2	$\mathcal{A}^{\mathrm{O}}(f_{\cos,j};b,m)$	o(1)	$(8\pi^2 j^2+15)/(12\pi^2 j^2)$
$\mathcal{C}(g;b,m)$	$G-1$	Eq. (11)	$\mathcal{C}^{\mathrm{O}}(g;b,m)$	$G-1$	various
$\mathcal{C}(g_0;b,m)$	5	0.8	$\mathcal{C}^{\mathrm{O}}(g_0;b,m)$	5	0.419
$\mathcal{C}(g_2^\star;b,m)$	o(1)	1.729	$\mathcal{C}^{\mathrm{O}}(g_2^\star;b,m)$	o(1)	0.777
$\mathcal{C}(g_4^\star;b,m)$	o(1)	1.042	$\mathcal{C}^{\mathrm{O}}(g_4^\star;b,m)$	o(1)	0.477
$\mathcal{N}(b,m)$	1	2	$\mathcal{O}(b,m)$	1	1.333

Example 14. Suppose that $Y_1, Y_2, \ldots, Y_n$ arise from the MA(1) process described in Example 7. Goldsman et al. (2003) derive the following *exact* results:

$$\mathrm{E}\big[\mathcal{O}(b,m)\big] = \sigma^2 + \frac{\gamma}{m(b-1)}\left[\frac{b^2+1}{b} - \frac{2}{mb-m+1}\right],$$

$$\mathrm{E}\big[\mathcal{A}^{\mathrm{O}}(f_2;b,m)\big] = \mathrm{E}\big[\mathcal{A}(f_2;b,m)\big] = \sigma^2 + \frac{7(\sigma^2+6\gamma)}{2m^2} + \mathrm{o}\big(m^{-2}\big)$$

and

$$\mathrm{E}\big[\mathcal{C}^{\mathrm{O}}\big(g_2^\star;b,m\big)\big] = \mathrm{E}\big[\mathcal{C}\big(g_2^\star;b,m\big)\big] = \sigma^2 + \frac{4(\sigma^2+6\gamma)}{m^2} + \mathrm{o}\big(m^{-2}\big).$$

First of all, we see that the bias results here reflect their counterparts from Example 7. Like that example, the overlapping area and CvM estimators are less biased than OBM as the batch size m becomes large; and the overlapping area estimator beats overlapping CvM for moderate m.

5 Summary and conclusions

This chapter has focused on the need to estimate the variance parameter σ^2 to give users information about the precision of the sample mean $\overline{Y}_n$ as an estimator for the steady-state mean μ. We presented a small selection of ways to go about estimating σ^2, illustrating various general tricks of the trade on certain variance estimators of interest along the way.

Certainly, mean squared error is an excellent criterion for comparing the performance of different estimators; but we could have also evaluated variance estimator performance when used within confidence interval estimators of the general form given by Equation (1) – it usually turns out that good point estimators yield good confidence intervals (cf. Sargent et al., 1992).

In terms of specifics, we found that

- overlapping estimators for σ^2 always seem to preserve bias and decrease variance vs. their nonoverlapping counterparts – sometimes by a great deal, especially in the case of the STS estimators;
- overlapping performs as advertised on the empirical MA(1) example.

So even though the overlapping estimators can be quite correlated from batch to batch, their overall performance is superior to that of the asymptotically i.i.d. estimators from nonoverlapping batches.

In the future, we can envision a focus on other variance estimators that reuse data over and over again, in the spirit of overlapping estimators. Examples of this ilk that immediately come to mind are the "orthonormal" estimators from Foley and Goldsman (1999) and the "permuted" estimators studied by Calvin and Nakayama (2002). Such estimators are apt to be computationally more complicated, but this should not present a problem, as a number of authors are addressing that particular issue, e.g., Damerdji et al. (1997).

Perhaps more-ambitious, long-term goals include those of enhancing automatic run-control procedures (e.g., Steiger et al., 2005), multivariate point and confidence interval estimation, and applications of variance estimation techniques to quality control, ranking-and-selection, optimization, and financial analysis problems.

Acknowledgement

Partial support for our research was provided by National Science Foundation Grant DMI-0400260.

References

Alexopoulos, C., Goldsman, D., Tokol, G. (2000). Properties of batched quadratic-form variance parameter estimators for simulations. *INFORMS Journal on Computing* 13, 149–156.

Alexopoulos, C., Argon, N.T., Goldsman, D., Steiger, N.M., Tokol, G., Wilson, J.R. (2005a). Overlapping variance estimators for simulation: Implementation and evaluation. Technical report, School of Industrial and Systems Engineering, Georgia Institute of Technology, Atlanta, GA.

Alexopoulos, C., Argon, N.T., Goldsman, D., Tokol, G., Wilson, J.R. (2005b). Overlapping variance estimators for simulation. Technical report, School of Industrial and Systems Engineering, Georgia Institute of Technology, Atlanta, GA.

Billingsley, P. (1968). *Convergence of Probability Measures*. Wiley, New York.

Bratley, P., Fox, B.L., Schrage, L.E. (1987). *A Guide to Simulation*, 2nd edition. Springer-Verlag, New York.

Calvin, J.M., Nakayama, M.K. (2002). A comparison of output-analysis methods for simulations of processes with multiple regeneration sequences. In: Yücesan, E., Chen, H., Snowdon, J.L. (Eds.), *Proceedings of the 2002 Winter Simulation Conference*. IEEE Press, Piscataway, NJ, pp. 328–335.

Chien, C., Goldsman, D., Melamed, B. (1997). Large-sample results for batch means. *Management Science* 43, 1288–1295.

Damerdji, H. (1995). Mean-square consistency of the variance estimator in steady-state simulation output analysis. *Operations Research* 43, 282–291.

Damerdji, H., Henderson, S.G., Glynn, P.W. (1997). Computational efficiency evaluation in output analysis. In: Andradóttir, S., Healy, K.J., Withers, D.H. (Eds.), *Proceedings of the 1997 Winter Simulation Conference*. IEEE Press, Piscataway, NJ, pp. 208–215.

Foley, R.D., Goldsman, D. (1999). Confidence intervals using orthonormally weighted standardized times series. *ACM Transactions on Modeling and Computer Simulation* 9, 297–325.

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

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

Goldsman, D., Meketon, M.S. (1986). A comparison of several variance estimators. Technical Report J-85-12, School of Industrial and Systems Engineering, Georgia Institute of Technology, Atlanta, GA.

Goldsman, D., Schruben, L.W. (1990). New confidence interval estimators using standardized time series. *Management Science* 36, 393–397.

Goldsman, D., Meketon, M.S., Schruben, L.W. (1990). Properties of standardized time series weighted area variance estimators. *Management Science* 36, 602–612.

Goldsman, D., Kang, K., Seila, A.F. (1999). Cramér–von Mises variance estimators for simulations. *Operations Research* 47, 299–309.

Goldsman, D., Kang, K., Kim, S.-H., Seila, A.F., Tokol, G. (2003). Combining standardized time series area and Cramér–von Mises variance estimators. Technical report, School of Industrial and Systems Engineering, Georgia Institute of Technology, Atlanta, GA.

Law, A.M., Kelton, W.D. (2000). *Simulation Modeling and Analysis*, 3rd edition. McGraw-Hill, New York.

Meketon, M.S., Schmeiser, B.W. (1984). Overlapping batch means: Something for nothing? In: Sheppard, S., Pooch, U.W., Pegden, C.D. (Eds.), *Proceedings of the 1984 Winter Simulation Conference*. IEEE Press, Piscataway, NJ, pp. 227–230.

Sargent, R.G., Kang, K., Goldsman, D. (1992). An investigation of finite-sample behavior of confidence interval estimators. *Operations Research* 40, 898–913.

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

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

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

Steiger, N.M., Wilson, J.R. (2001). Convergence properties of the batch-means method for simulation output analysis. *INFORMS Journal on Computing* 13, 277–293.

Steiger, N.M., Lada, E.K., Wilson, J.R., Joines, J.A., Alexopoulos, C., Goldsman, D. (2005). ASAP3: A batch means procedure for steady-state simulation analysis. *ACM Transactions on Modeling and Computer Simulation* 15, 39–73.