Available online at www.sciencedirect.com
ScienceDirect
Operations Research Letters 35 (2007) 439–447
Operations Research Letters
www.elsevier.com/locate/orl

Linear combinations of overlapping variance estimators for simulation

Tûba Aktaran-Kalaycı[a], David Goldsman[b,*], James R. Wilson[c]

[a]Department of Industrial and Systems Engineering, University at Buffalo, Buffalo, NY 14260, USA
[b]H. Milton Stewart School of Industrial and Systems Engineering, Georgia Institute of Technology, Atlanta, GA 30332, USA
[c]Edward P. Fitts Department of Industrial and Systems Engineering, North Carolina State University, Raleigh, NC 27695, USA

Received 8 November 2005; accepted 9 August 2006
Available online 3 November 2006

Abstract

We estimate the variance parameter of a stationary simulation-generated process using a linear combination of overlapping standardized time series (STS) area variance estimators based on different batch sizes. We establish the linear-combination estimator's asymptotic distribution, presenting analytical and simulation-based results exemplifying its potential for improvements in accuracy and computational efficiency.

Keywords: Simulation; Stationary process; Variance estimation; Standardized time series; Area estimator; Overlapping area estimator

1. Introduction

Steady-state simulations are used to analyze a variety of complicated systems. If the output process $\{Y_i : i = 1, 2, \ldots, n\}$ is in steady state, then we might wish to estimate its mean μ. The sample mean $\bar{Y}_n = n^{-1}\sum_{i=1}^{n} Y_i$ is the natural, unbiased estimator for μ based on a sample of size n; but to complete the picture, we need to estimate the sample mean's variability as well. One such measure is the *variance parameter*, $\sigma^2 \equiv \lim_{n\to\infty} n\mathsf{Var}(\bar{Y}_n)$, which is difficult to estimate since outputs from steady-state simulations are almost never independent and identically distributed (i.i.d.) normal random variables. Well-known methods for estimating σ^2 include the following: the autoregressive method; nonoverlapping batch means (NBM); overlapping batch means (OBM); the regenerative method; spectral analysis; and standardized time series (STS) (see, for instance, Law and Kelton [11]). The current article combines in a synergistic way ideas involving STS and OBM to produce improved estimators of the variance parameter of a steady-state simulation output process.

* Corresponding author. Tel.: +1 404 894 2365; fax: +1 404 984 2301.
E-mail addresses: ta27@buffalo.edu (T. Aktaran-Kalaycı), sman@gatech.edu (D. Goldsman), jwilson@ncsu.edu (J.R. Wilson).

0167-6377/$ - see front matter © 2006 Published by Elsevier B.V.
doi:10.1016/j.orl.2006.08.007

Batching typically increases bias (a drawback), but decreases variance (an advantage). And an estimator computed from overlapping batches usually has smaller variance than that of the corresponding estimator computed from nonoverlapping batches. In particular, Meketon and Schmeiser [12] illustrate this phenomenon for variance estimators arising from NBM and OBM; they show that the variance of the OBM estimator is about $\frac{2}{3}$ that of the corresponding NBM estimator (where both methods are assumed to use the same batch size). Alexopoulos et al. [3,4] demonstrate even-more-dramatic savings when comparing a variety of overlapping STS estimators with their nonoverlapping counterparts.

In the present article, we first review some background basics on overlapping estimators in Section 2. Our main contribution, in Section 3, is to apply overlapping batch techniques to the STS area estimator using *different batch sizes* and then to construct a variance-optimal linear combination of these estimators. The resulting combined estimators will have lower bias and variance than their constituent estimators. We conduct a small Monte Carlo study in Section 3.5 to illustrate our points, before offering concluding remarks in Section 4. Some preliminary results on variance-optimal linear combinations of overlapping STS area variance estimators are also presented in Aktaran-Kalaycı and Goldsman [1].

2. Background

2.1. Setup for using STS variance estimators

We assume the stationary stochastic process $\{Y_i, i \geqslant 1\}$ has steady-state mean μ and variance parameter $\sigma^2 \in (0, \infty)$ such that the following functional central limit theorem (FCLT) holds.

Assumption FCLT. The STS

$$X_n(t) \equiv \frac{\lfloor nt \rfloor (\bar{Y}_{\lfloor nt \rfloor} - \mu)}{\sigma\sqrt{n}} \quad \text{for } t \in [0,1] \text{ and } n = 1, 2, \ldots \tag{1}$$

satisfies $X_n \underset{n\to\infty}{\Longrightarrow} \mathscr{W}$, where: $\lfloor \cdot \rfloor$ denotes the greatest integer function; $\mathscr{W}$ is a standard Brownian motion process on [0, 1]; and $\underset{n\to\infty}{\Longrightarrow}$ denotes weak convergence in $D[0,1]$, the space of functions on [0, 1] that are right-continuous with left-hand limits, as $n \to \infty$. See also [5,7].

Suppose we have n observations $Y_1, Y_2, \ldots, Y_n$ on hand, and we form $n - m + 1$ *overlapping batches*, each of size m. In particular, the observations $Y_i, Y_{i+1}, \ldots, Y_{i+m-1}$ constitute batch i, for $i = 1, 2, \ldots, n - m + 1$. Similar to the original definition in Schruben [14], the STS from overlapping batch i is

$$T^{\mathrm{O}}_{i,m}(t) \equiv \frac{\lfloor mt \rfloor (\bar{Y}^{\mathrm{O}}_{i,m} - \bar{Y}^{\mathrm{O}}_{i,\lfloor mt \rfloor})}{\sigma\sqrt{m}} \quad \text{for } 0 \leqslant t \leqslant 1 \text{ and } i = 1, 2, \ldots, n-m+1,$$

where

$$\bar{Y}^{\mathrm{O}}_{i,j} \equiv \frac{1}{j}\sum_{\ell=0}^{j-1} Y_{i+\ell} \quad \text{for } i = 1, 2, \ldots, n-m+1 \text{ and } j = 1, 2, \ldots, m.$$

Alexopoulos et al. [3] show that, under Assumption FCLT,

$$\sigma T^{\mathrm{O}}_{\lfloor sm \rfloor, m} \underset{m\to\infty}{\Longrightarrow} \sigma \mathscr{B}_{s,1} \quad \text{for fixed} \quad s \in [0, b-1],$$

where $b \equiv n/m > 1$ is a fixed ratio; and for $r \in [1, b)$ and $s \in [0, b - r]$, we let $\mathscr{B}_{s,r}(\cdot)$ denote a Brownian bridge process on the unit interval

$$\mathscr{B}_{s,r}(t) = \frac{t[\mathscr{W}(s+r) - \mathscr{W}(s)] - [\mathscr{W}(s+tr) - \mathscr{W}(s)]}{\sqrt{r}} \quad \text{for } t \in [0,1]. \tag{2}$$

2.2. Overlapping area variance estimator

The idea is to form a separate estimator for σ^2 from each overlapping batch—even though the resulting estimators are clearly dependent—and then to average those estimators. In the context of this article, the fixed quantity b may in general take any real value exceeding one while we let the batch size $m \to \infty$; and thus the total sample size is always taken to be $\lfloor bm \rfloor$. The area estimator from the ith overlapping batch is

$$A^{\mathrm{O}}_i(f; m) \equiv \left[\frac{1}{m}\sum_{\ell=1}^{m} f(\ell/m)\sigma T^{\mathrm{O}}_{i,m}(\ell/m)\right]^2 \quad \text{for } i = 1, 2, \ldots, \lfloor bm \rfloor - m + 1,$$