DOC. 9

AUTOMATIC BATCHING IN SIMULATION OUTPUT ANALYSIS

A Thesis

Submitted to the Faculty

of

Purdue University

by

António Manuel de Carvalho Pedrosa

In Partial Fulfillment of the

Requirements for the Degree

of

Doctor of Philosophy

May 1994

To Lúcia, my wife, Sarinha and Cati, my daughters, Ana and Alberto, my parents, and to the memory of Fernanda and José, my grandparents.

ACKNOWLEDGMENTS

It is a pleasure to express my thanks to all the people and organizations that made this work possible.

First mention goes to my advisor, Professor Bruce Schmeiser, for his excellent guidance and encouragement throughout this research. His suggestions, insight, and willingness to discuss my research whenever I needed were most helpful.

During this research I also had the benefit of lengthy discussions with the other members of my committee, Professors Dimitris Politis, Colm O'Cinneide, and Kevin Healy, and a number of colleagues and friends at Purdue University. Professor Politis deserves special recognition for providing references and suggestions that were helpful in improving the quality of this thesis. Among the friends at Purdue University I thank, in particular, Jin Wang for his constructive comments.

I also would like to thank two other people: Connie Good, a staff member of the School of Industrial Engineering, Purdue University, for helping me in administrative procedures; and Wheyming Tina Song, National Tsing Hua University, Taiwan, for providing the computer code to analyze her algorithm in Chapter 7 of this thesis.

I shall also refer Professor Augusto Medina, director of Escola Superior de Biotecnologia, Universidade Católica Portuguesa, for his valuable support, and Escola Superior de Biotecnologia, Junta Nacional de Investigação Cientifica e Tecnológica, and National Science Foundation, Grant No. DDM-93-00058, for providing the funding of my studies at Purdue University.

Finally, I would like to thank my family. My parents and grand-parents always encouraged my education. My wife, Lúcia, and daughters, Sarinha and Cati, have always given their love, support, and encouragement.

TABLE OF CONTENTS

Page

LIST OF TABLES

LIST OF FIGURES

ABSTRACT

Pedrosa, António Manuel de Carvalho. Ph.D., Purdue University, May 1994. Automatic Batching in Simulation Output Analysis. Major Professor: Bruce Schmeiser.

We develop methodology and propose an algorithm to estimate the variance of the sample mean of stationary stochastic processes. The algorithm, denoted by 1–2–1 OBM, is (i) automatic, in that it does not require the user to provide any parameter, (ii) robust, in that it can be applied to any stationary stochastic process and is not sensitive to reasonable violations of assumptions, (iii) computationally efficient, in that the computation time is proportional to the sample size, and (iv) statistically efficient, in that the mean squared error performance is good. We also provide a FORTRAN implementation.

Our approach minimizes the mean squared error of estimators of the variance of the sample mean parameterized by batch size by estimating the optimal batch size from a given data sample. The optimal batch size is a simple function of the correlation structure of the data: the sum of autocorrelations, γ_0, and the weighted sum of autocorrelations, γ_1. We develop estimators $\widehat{\gamma}_0$ and $\widehat{\gamma}_1$ of γ_0 and γ_1. Both $\widehat{\gamma}_0$ and $\widehat{\gamma}_1$ are obtained from overlapping-batch-means (OBM) estimators, but analogous development can be considered for other estimators parameterized by batch size. We develop theoretical and empirical guidelines to estimate the optimal batch size from $\widehat{\gamma}_0$ and $\widehat{\gamma}_1$. We prove that our optimal batch size estimator converges in probability to the optimal batch size.

We also develop asymptotically unbiased methodology to estimate the variance of the sample mean based on classical linear regression of OBM estimators. This approach motivates our derivation of asymptotic results regarding bias, covariance, and correlation of OBM estimators. We derive analogous results for Bartlett estimators, and we generalize these results for non-zero frequencies, i.e., for Bartlett estimators of the spectral density. We prove that OBM estimators and Bartlett estimators of the variance of the sample mean are asymptotically equivalent.

In addition, we develop DPSS, a four-parameter family of stochastic processes with good analytical, computational, and statistical properties for evaluating simulation methods. The four parameters are related to the mean, variance, lag-one autocorrelation, sum of autocorrelations, and weighted sum of autocorrelations through simple closed-form equations. The stationary marginal distribution is discrete uniform. Variates are easily and efficiently generated. The autocorrelation structure has a damped oscillating behavior allowing for a wide range of dependency characteristics.

1. INTRODUCTION

1.1 Motivation and Topic

Models of complex stochastic systems, which can not be analyzed with analytical or numerical methods, are often simulated. Realizations from random numbers are generated and observations are collected to estimate θ, the performance of the system, usually expressed as an expected value. The simulation point estimator $\widehat{\theta}$ is a random variable, the randomness being caused by the choice of a random number seed: choose a different seed and the value of the point estimator changes. Estimating the magnitude of the sampling error — the deviation between $\widehat{\theta}$ and the unknown true value θ — is a fundamental problem of simulation experimentation.

There exists an extensive literature describing output-analysis methods for estimating sampling error. Almost all approaches measure, directly or indirectly, sampling error using the standard error, the standard deviation of $\widehat{\theta}$. The standard-error estimate can be used alone, to construct a confidence interval, to test a hypothesis, or to determine whether more sampling is needed. For tractability, we focus on estimating $V = \mathrm{var}(\widehat{\theta})$ rather than its square root, the standard error.

Simulation output-analysis methods require statistical expertise to be used correctly and efficiently; for example, with batch-means methods a batch size must be selected. We are interested in methodology that, regardless of the type of data and of the type of property being estimated, can be applied automatically. By "automatic", we mean that the simulation output-analysis software, without intervention from the user, provides a good estimate of V.

Our primary measure of goodness of an estimator $\widehat{V}$ is the mean squared-error (MSE). The choice of squared-error loss is arbitrary; any loss function could be used. The MSE combines in a simple measure the variability of the estimator (precision) and its bias (accuracy), since $\mathrm{mse}(\widehat{V}) = \mathrm{E}(\widehat{V} - V)^2 = \mathrm{var}(\widehat{V}) + \mathrm{bias}^2(\widehat{V})$, where $\mathrm{bias}(\widehat{V}) = \mathrm{E}(\widehat{V}) - V$. Lehmann [1991, pp. 7-8] provides a good discussion of MSE as a measure of estimator quality.

Often in simulation the measure of performance, θ, is a population mean and the point estimator, $\widehat{\theta}$, is the sample mean. This dissertation considers the estimation of variance of the sample mean. Our main objective is to develop automatic (no parameters for the user to specify), robust, and computationally and statistically efficient methodology to estimate the variance of the sample mean. The primary tangible product of this dissertation is the 1–2–1 OBM algorithm for estimating V, which is based on the overlapping-batch-means (OBM) estimator.

1.2 Organization and Summary

This dissertation consists of nine chapters. Chapters 3, 4, 6, and 7 are based on papers submitted for publication.

Chapter 2 is background information. We introduce notation, review batch-means estimators of the variance of the sample mean and Bartlett estimators of the spectral density, and discuss the location-invariance property of OBM and Bartlett estimators. We also review three stochastic processes that are used in our Monte Carlo experimentation.

In Chapter 3 we develop DPSS, a four-parameter family of stochastic processes. The genesis of the process is the (s, S) inventory system with Bernoulli demand. The motivation is to create a test-bed process whose properties differ in fundamental ways from other test-bed processes such as the MA(1), AR(1), and EAR(1), and that can be used for evaluating simulation methods. Both variate generation and the process properties are tractable. The stationary marginal distribution is discrete uniform. The stationary mean, marginal variance, and autocorrelation structure are simple functions of the parameters; the parameters are simple functions of the stationary marginal mean, marginal variance, lag-one autocorrelation, sum of autocorrelations, and weighted sum of autocorrelations. The autocorrelation structure is cyclic, a more complex behavior than that of other test-bed processes. The DPSS process is used in the Monte Carlo experiments of Chapters 4 and 7.

In Chapters 4, 5, and 6 we derive theoretical guidelines. Some of these results are central to the 1–2–1 OBM algorithm developed in Chapter 7 and to the methodology based on regression of OBM estimators developed in Chapter 8. In Chapter 4 we derive the asymptotic covariances and correlations of OBM and Bartlett estimators when they are applied to a common sample from a stationary general linear process. The asymptotic correlations, which are identical for both types of estimators, are functions of the OBM batch size(s) and the Bartlett lag-window size(s). We are motivated by the possibility of using linear combinations of these estimators, which requires choosing p estimators and p corresponding weights. Both for determining appropriate linear-combination weights and for choosing batch/window sizes, the covariances and correlations are needed. Our Monte Carlo study suggests that the asymptotic correlation formula provides a good approximation to the true finite-sample correlation if (1) the sample size n is at least several multiples of γ_0 and (2) the both batch sizes are between γ_0 and $n/2$, where γ_0 is the sum of all autocorrelations.

In Chapter 5 we generalize to non-zero frequencies the results of Chapter 4. We consider the asymptotic covariance between two Bartlett estimators of the spectral density function applied to common data from a stationary general linear process. The asymptotic covariance is the product of the asymptotic variance of the Bartlett estimator with the smaller window size and a simple function of the window sizes. As with zero frequency in Chapter 4, the asymptotic correlation and the ratios of the asymptotic variances are independent of the spectral density, and therefore asymptotically independent of the application.

In Chapter 6 we derive the expected values and biases of the OBM estimator of the variance of the sample mean and of the Bartlett spectral estimator at zero frequency. We show that the OBM estimator is asymptotically equivalent to the Bartlett estimator. Setting the OBM batch size to the Bartlett window size, we show that the estimators' correlation is asymptotically one and that the estimators converge in mean squared error even after rescaling to estimate $n\,V$, where n is the sample size.

In Chapter 7 we develop computationally and statistically efficient methodology for estimating the optimal batch size of OBM estimators. We propose the 1–2–1 OBM algorithm for estimating the variance of the sample mean of stationary processes. This $O(n)$ (computation time proportional to n) algorithm has good theoretical properties. Theoretical results suggest, and Monte Carlo experimentation shows, that the MSE performance of 1–2–1 OBM approaches the idealistic performance that could be obtained if the correlation structure of the data were known. We also derive $O(n)$

estimators $\widehat{\gamma}_0$ and $\widehat{\gamma}_1$ of γ_0 (sum of autocorrelations) and γ_1 (weighted sum of autocorrelations), respectively, and state conditions for $\widehat{\gamma}_0$ and $\widehat{\gamma}_1$ to be MSE-consistent.

In Chapter 8 we derive asymptotically unbiased estimators of V, γ_0, and γ_1, using linear regression of OBM estimators, which is a special case of using linear combinations of classical estimators. We focus on estimating V and we discuss statistical efficiency, computational efficiency, and implementation aspects of regression based algorithms.

Finally, in Chapter 9, we present a summary of the results of this research and our conclusions. We also suggest some directions for future research.

The methodology that we develop is based on OBM estimation. Analogous development can be considered for other estimators parameterized by batch size, such as non-overlapping batch means and standardized time series. Also, algorithms structurally different from the 1–2–1 OBM (for example, with a different number of estimation stages) could be based on the theoretical results derived in this dissertation.

2. BACKGROUND

Here we summarize background information about the variance of the sample mean, overlapping batch means, and Bartlett spectral estimators. We also review three stochastic data processes that are used in our Monte Carlo experimentation.

2.1 Variance of the Sample Mean

For stationary time series $\{X_i\}$ a natural unbiased estimator of the population mean μ is the sample mean

$$\bar{X} = \frac{1}{n} \sum_{i=1}^{n} X_i, \qquad [2.1]$$

where n is the number of observations. The variance of any sample mean is

$$V = \frac{1}{n^2} \sum_{i=1}^{n} \sum_{j=1}^{n} \operatorname{cov}(X_i, X_j). \qquad [2.2]$$

For stationary time series, $\operatorname{cov}(X_i, X_{i+h}) = R(h)$ is a finite constant, yielding

$$V = \frac{R(0)}{n} \sum_{h=-n}^{n} \left(1 - \frac{|h|}{n}\right) \rho(h), \qquad [2.3]$$

where $R(0) = \operatorname{var}(X)$ and $\rho(h) = \operatorname{corr}(X_i, X_{i+h})$. Define

$$\gamma_0 = \sum_{h=-\infty}^{\infty} \rho(h). \qquad [2.4]$$

If $0 < \gamma_0 < \infty$ then

$$\lim_{n \to \infty} n\, V = \gamma_0\, R(0), \qquad [2.5]$$

as shown in, for example, Anderson [1971, p. 460]. For independent and identically distributed (iid) data, $\gamma_0 = 1$. Equation 2.5 implies that γ_0 can be thought of as the number of contiguous observations that carry the same information as one independent observation. The rate of convergence in Equation 2.5 depends on

$$\gamma_1 = \sum_{h=-\infty}^{\infty} |h| \; \rho(h) = 2\sum_{h=1}^{\infty} h \; \rho(h), \qquad [2.6]$$

the weighted sum of the correlations. In particular, Song and Schmeiser [1994] show that if $\gamma_1 < \infty$ then

$$n \; V \quad = \quad \gamma_0 \; R(0) \quad - \quad \frac{\gamma_1 R(0)}{n} \quad + \; o\left(\frac{1}{n}\right). \qquad [2.7]$$

The notation $O\,(g(n))$ and $o\,(g(n))$ has the usual meaning, i.e., (1) a function $f(n)$ is $O\,(g(n))$ if there is a positive finite constant M such that $|f(n)|/|g(n)| \leq M$ for n large enough; and (2) a function $f(n)$ is $o\,(g(n))$ if $f(n)/g(n) \to 0$ as $n \to \infty$.

2.2 OBM Estimator of V

The batch means methodology is based on dividing the n observations X_1, X_2, ..., X_n into b batches of size m with lag l. Thus batch i with lag l consists of observations $X_{l(i-1)+1}$, ..., $X_{l(i-1)+m}$. For $l = 1$ there is full overlap and the estimator is called overlapping batch means (OBM) [Meketon and Schmeiser, 1984] and for $l = m$ there is no overlap and the estimator is the nonoverlapping batch means (NBM) [Schmeiser, 1982]. Welch [1987] discusses partial overlapping, $1 < l < m$. Fox, Goldsman and Swain [1991] discuss spaced batch means, $l > m$.

The main concept underlying batch means is to include the effects of autocorrelation within, rather than between, the batch means. Equivalently, the main concept is that whatever correlation that exists between batch means is due to common data rather than to the autocorrelation structure, as discussed in a more general context in Schmeiser, Avramidis, and Hashem [1990].

The OBM estimator of V is defined as

$$\widehat{V}^{(O)} = \frac{m}{(n-m+1)(n-m)} \sum_{i=1}^{n-m+1} \left(\bar{X}_i - \bar{X}\right)^2 , \qquad [2.8]$$

where $1 \leq m \leq n-1$, and

$$\bar{X}_i \;=\; \frac{1}{m} \sum_{j=i}^{i-1+m} X_j. \qquad [2.9]$$

This estimator is unbiased for iid data for any sample size n and any batch size m, but autocorrelation can cause bias. Computation is $O(n)$ [Meketon and Schmeiser, 1984].

The OBM is a quadratic-form estimator since it can be written as

$$\widehat{V}^{(O)} = \sum_{i=1}^{n}\sum_{j=1}^{n} p_{ij}^{(O)} X_i X_j \qquad [2.10]$$

for constant coefficients $p_{ij}^{(O)}$,

$$p_{ij}^{(O)} = p_{ji}^{(O)} = \frac{m}{(n-m+1)(n-m)} \left[\frac{a_{ij}}{m^2} - \frac{a_{ii}+a_{jj}}{m\, n} + \frac{n-m+1}{n^2}\right], \qquad [2.11]$$

where a_{ij}, the number of batches that includes both X_i and X_j, is defined by

$$a_{ij} \;=\; \min\left\{n-m+1, \max(0, m-|j-i|), \min(i,j), n-\max(i,j)+1\right\}. \qquad [2.12]$$

The first term, $n-m+1$, is the number of batches; the second reflects the batch size m and lag of cross product; the third and fourth terms are end effects [Song and Schmeiser, 1993]. For later use, we rewrite the OBM quadratic-form coefficients as

$$p_{ij}^{(O)} = \begin{cases} O\left(\frac{m}{n^3}\right), & |i-j| > m, \\ O\left(\frac{\min(i,j)}{m\, n^2}\right), & i < m \text{ and } j < m, \\ O\left(\frac{n-\max(i,j)+1}{m\, n^2}\right), & i > n-m \text{ and } j > n-m, \\ \frac{1}{n^2}\left(1-\frac{|i-j|}{m}\right)\left[1+O\left(\frac{m}{n}\right)\right], & |i-j| \leq m,\ i \text{ or } j \geq m, \text{ and } i \text{ or } j \leq n-m. \end{cases} \qquad [2.13]$$

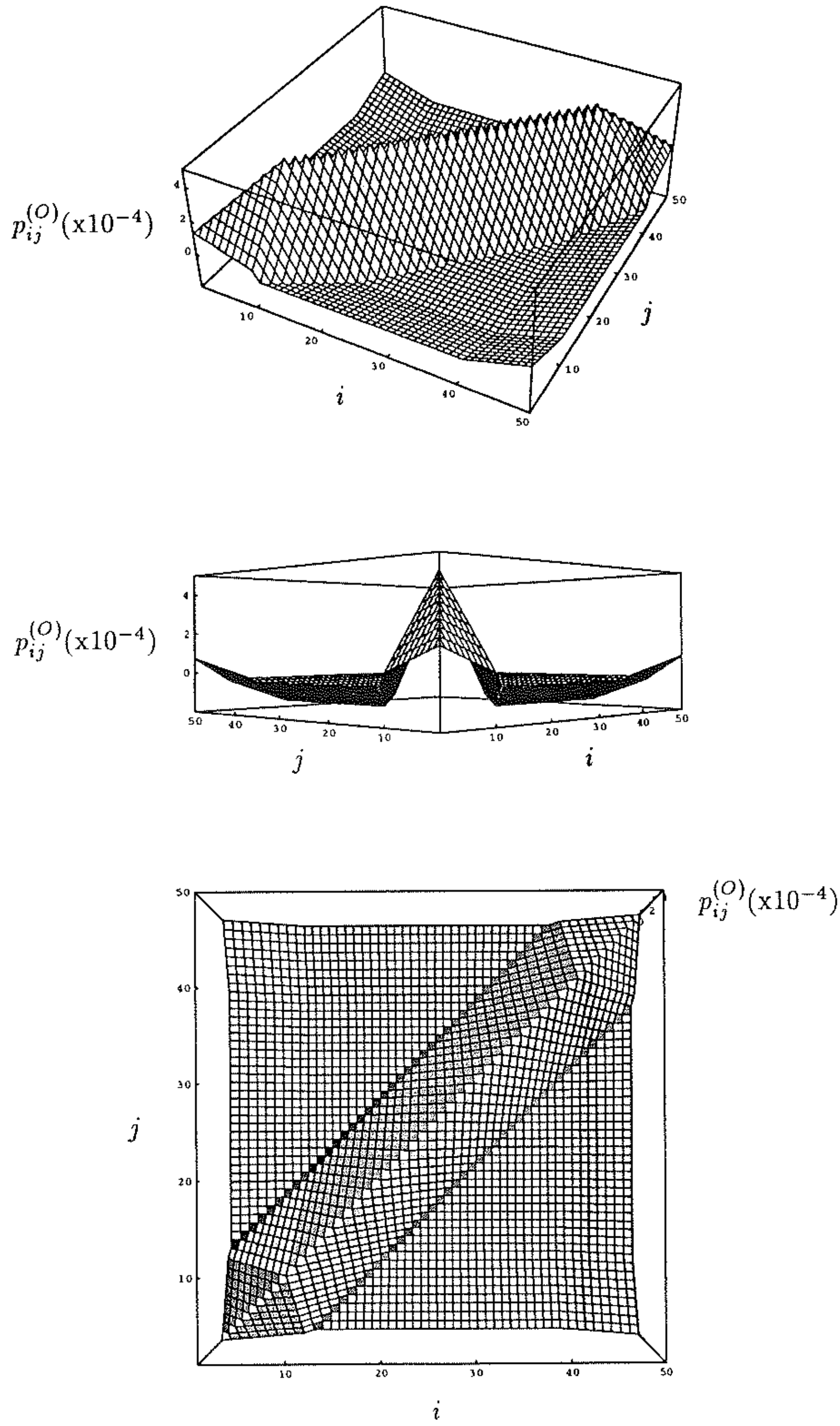

Figure 2.1 OBM quadratic-form coefficients for $n = 50$ and $m = 10$.

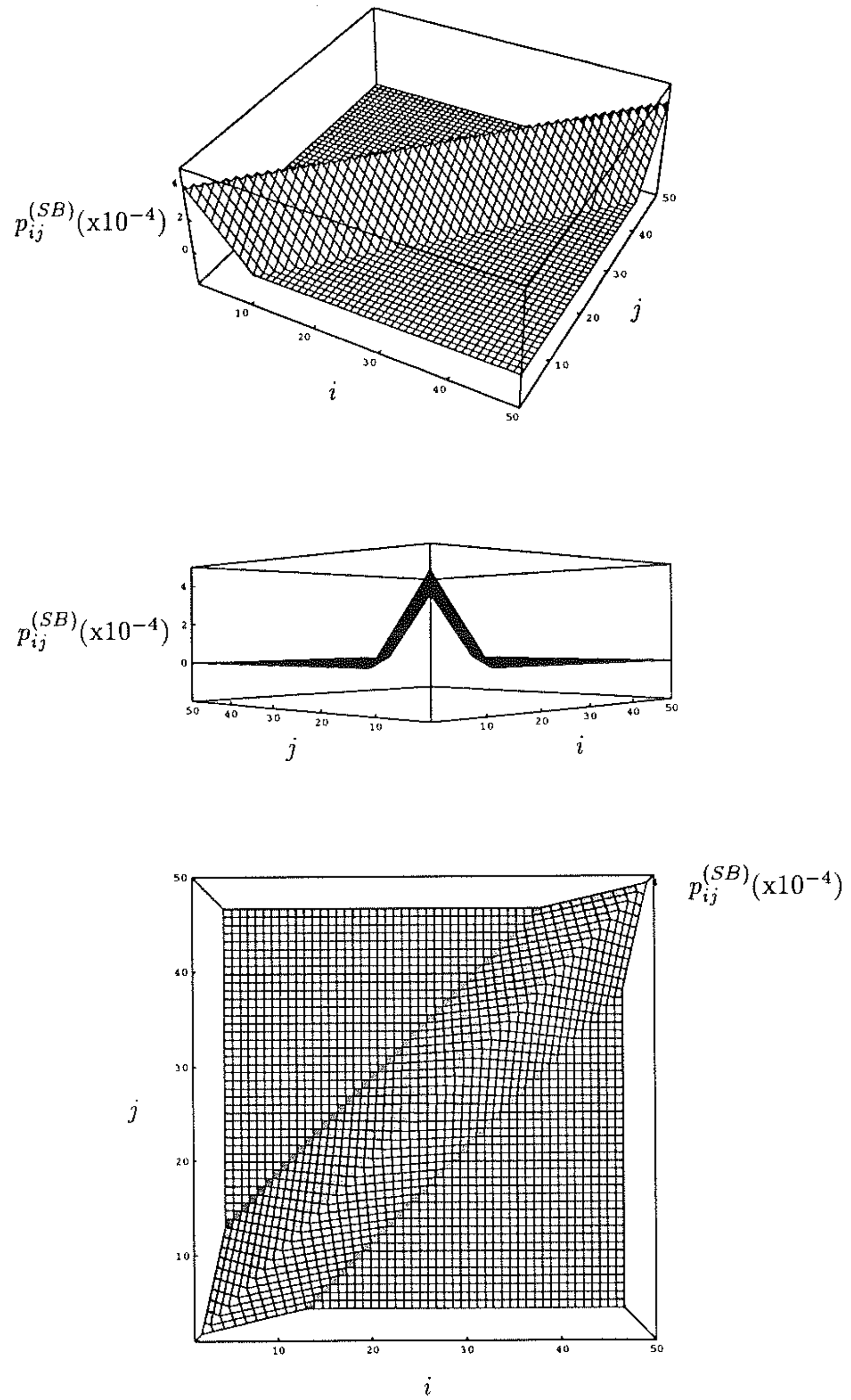

Figure 2.2 Simplified Bartlett quadratic-form coefficients for $n = 50$ and $m = 10$.

2.3 Bartlett Estimator of V

The Bartlett estimator [e.g., Priestley, 1992, p. 439] of the variance of the sample mean is

$$\widehat{V}^{(B)} = \frac{1}{n} \sum_{h=-m}^{m} (1 - \frac{|h|}{m})\, \widehat{R}(h), \qquad [2.14]$$

where m is the truncation point of the spectral window and m is smaller than the number of observations n. $\widehat{R}(h)$, the usual lag-h sample autocovariance, is

$$\widehat{R}(h) = \frac{1}{n} \sum_{i=1}^{n-|h|} \left(X_i - \bar{X}\right) \left(X_{i+|h|} - \bar{X}\right), \qquad 0 \leq |h| \leq n-1. \qquad [2.15]$$

Substituting Equation 2.15 into Equation 2.14 yields the quadratic-form expression for the Bartlett estimator

$$\widehat{V}^{(B)} = \sum_{i=1}^{n} \sum_{j=1}^{n} q_{ij}^{(B)} \left(X_i - \bar{X}\right) \left(X_{i+|h|} - \bar{X}\right), \qquad [2.16]$$

where

$$q_{ij}^{(B)} = q_{ji}^{(B)} = \begin{cases} \frac{1}{n^2} \left(1 - \frac{|i-j|}{m}\right), & |i-j| \leq m, \\ 0, & |i-j| > m. \end{cases} \qquad [2.17]$$

Equation 2.16 can be rewritten as

$$\widehat{V}^{(B)} = \sum_{i=1}^{n} \sum_{j=1}^{n} p_{ij}^{(B)} X_i X_j, \qquad [2.18]$$

where, as discussed in Song and Schmeiser [1993], $p_{ij}^{(B)} = q_{ij}^{(B)} + O(m/n^3)$, i.e.,

$$p_{ij}^{(B)} = p_{ji}^{(B)} = \begin{cases} \frac{1}{n^2} \left(1 - \frac{|i-j|}{m}\right) \left[1 + O\left(\frac{m}{n}\right)\right], & |i-j| \leq m, \\ O\left(\frac{m}{n^3}\right), & |i-j| > m. \end{cases} \qquad [2.19]$$

If the process mean is assumed to be known and zero, the simplified Bartlett estimator

$$\widehat{V}^{(SB)} = \sum_{i=1}^{n} \sum_{j=1}^{n} p_{ij}^{(SB)} X_i X_j, \qquad [2.20]$$

where $p_{ij}^{(SB)} = q_{ij}^{(B)}$ (Equation 2.17), is obtained from Equations 2.14 and 2.15 by setting $\bar{X} = 0$. This simplified form of the Bartlett estimator is often used for analysis because it is more tractable.

Figures 2.1 and 2.2 show three-dimensional plots of $p_{ij}^{(O)}$ and $p_{ij}^{(SB)}$, respectively, as functions of i and j, for $n = 50$ observations and batch/window size $m = 10$. Each figure contains three different views (front and up, left-hand corner, and top, respectively) of the same estimator. These graphs illustrate the similarities between OBM and simplified Bartlett estimators.

2.4 General Linear Process and Bartlett Estimator of Spectral Density

Suppose that the observations $\{X_i\}$ are from a stationary time series and these data can be expressed by a general linear process of the form

$$X_i = \sum_{h=-\infty}^{+\infty} b_h \, \varepsilon_{i-h}, \tag{2.21}$$

where the b_h's are constants and $\{\varepsilon_i\}$ is a sequence of iid random variables with finite variance σ_ε^2. The spectral density function is defined by

$$f(\omega) = \frac{1}{2\,\pi} \sum_{h=-\infty}^{\infty} R(h) \, e^{-jh\omega}, \qquad -\pi \le \omega \le \pi, \tag{2.22}$$

where j denotes the imaginary unit $\sqrt{-1}$ and $R(h)$, the lag-h autocovariance, is

$$R(h) = {\sigma_\varepsilon}^2 \sum_{i=-\infty}^{+\infty} b_i \, b_{i+|h|}. \tag{2.23}$$

At zero frequency the spectral density function is

$$f(0) = \frac{{\sigma_\varepsilon}^2}{2\pi} \left(\sum_{i=-\infty}^{\infty} b_i \right)^2 . \tag{2.24}$$

The Bartlett estimator of the spectral density function is

$$\hat{f}(\omega) = \frac{1}{2\,\pi} \sum_{h=-m}^{m} (1 - \frac{|h|}{m}) \, \hat{R}(h) \, e^{-jh\omega}, \tag{2.25}$$

where m is the truncation point of the spectral window and m is smaller than the number of observations n. $\widehat{R}(h)$ is the usual lag-h sample autocovariance (Equation 2.15). Substituting Equation 2.15 into Equation 2.25 yields the quadratic-form expression for the Bartlett spectral estimator

$$\widehat{f}(\omega) = \frac{n}{2\pi} \sum_{r=1}^{n} \sum_{s=1}^{n} q_{rs} \left(X_r - \bar{X}\right) \left(X_s - \bar{X}\right) e^{-j(r-s)\omega}, \qquad [2.26]$$

where q_{rs} is defined by Equation 2.17, i.e.,

$$q_{rs} = q_{sr} = \begin{cases} \frac{1}{n^2}\left(1 - \frac{|r-s|}{m}\right), & |r - s| \leq m, \\ 0, & |r - s| > m. \end{cases}$$

Equation 2.26 can be rewritten as

$$\widehat{f}(\omega) = \frac{n}{2\pi} \sum_{r=1}^{n} \sum_{s=1}^{n} p_{rs} X_r X_s e^{-j(r-s)\omega}, \qquad [2.27]$$

where, as shown in Appendix A, p_{rs} is defined by Equation 2.19, i.e.,

$$p_{rs} = p_{sr} = \begin{cases} \frac{1}{n^2}\left(1 - \frac{|r-s|}{m}\right)\left[1 + O\left(\frac{m}{n}\right)\right], & |r - s| \leq m, \\ O\left(\frac{m}{n^3}\right), & |r - s| > m. \end{cases}$$

Notice that

$$\widehat{f}(0) = \frac{n}{2\pi} \widehat{V}^{(B)}, \qquad [2.28]$$

where $\widehat{V}^{(B)}$ is defined by Equation 2.18.

If the process mean is assumed to be known and zero, the simplified Bartlett spectral estimator

$$\widehat{f}^{SB}(\omega) = \frac{n}{2\pi} \sum_{r=1}^{n} \sum_{s=1}^{n} q_{rs} X_r X_s e^{-j(r-s)\omega}, \qquad [2.29]$$

is obtained from Equation 2.26 by setting $\bar{X} = 0$. This simplified form of the Bartlett spectral estimator is often used for analysis because it is more tractable.

2.5 Location Invariance of OBM and Bartlett Estimators

OBM estimators of V (Equations 2.8 through 2.13), Bartlett estimators of V (Equations 2.14 through 2.19), and Bartlett estimators of the spectral density (Equations 2.25, 2.26, and 2.27) are location invariant. Therefore, we can assume without loss of generality that the process has zero mean. To see this, define the process $\{Y_i\} = \{X_i - d\}$, for any constant d. Let $\bar{Y}$ and $\bar{Y}_i$ be defined by Equations 2.1 and 2.9, respectively. The difference $\bar{Y}_i - \bar{Y}$ equals $\bar{X}_i - \bar{X}$. By Equation 2.8, the OBM estimates for both $\{X_i\}$ and $\{Y_i\}$ are the same. Similarly, $Y_i - \bar{Y} = X_i - \bar{X}$, and, by Equations 2.16 and 2.26, the Bartlett estimates for both $\{X_i\}$ and $\{Y_i\}$ are the same. Therefore, OBM and Bartlett estimators do not depend on the process mean, implying no loss of generality if zero mean is assumed.

2.6 Standard Results for the Fourth Moment

The results of Chapters 4 and 5 involve the evaluation of fourth-order functions of $\{X_i\}$. We use the standard result for quadrivariate distributions

$$\begin{aligned} \mathrm{E}(X_i X_{i+r} X_{i+s} X_{i+t}) &= \mathrm{E}(X_i\, X_{i+r})\,\mathrm{E}(X_{i+s}\, X_{i+t}) + \mathrm{E}(X_i\, X_{i+s})\,\mathrm{E}(X_{i+r}\, X_{i+t}) \\ &\quad + \mathrm{E}(X_i\, X_{i+t})\,\mathrm{E}(X_{i+r}\, X_{i+s}) + K_4(X_i, X_{i+r}, X_{i+s}, X_{i+t}), \end{aligned} \qquad [2.30]$$

where, as discussed, e.g., in Priestley [1992, p. 325], $K_4(X_i, X_{i+r}, X_{i+s}, X_{i+t})$ is the fourth joint cumulant of the distribution of $[X_i, X_{i+r}, X_{i+s}, X_{i+t}]$. For a general linear process the fourth joint cumulant is

$$K_4(X_i, X_{i+r}, X_{i+s}, X_{i+t}) = K_{\varepsilon,4} \sum_{h=-\infty}^{+\infty} b_h\, b_{h+r}\, b_{h+s}\, b_{h+t}, \qquad [2.31]$$

where the b_h's are the constants defined in Equation 2.21 and $K_{\varepsilon,4}$ is the fourth cumulant of ε_i [Rosenblatt, 1985, p. 47]. This quantity is related to the central moments by

$$K_{\varepsilon,4} \;=\; \mu_{\varepsilon,4} \;-\; 3\,\sigma_\varepsilon^2 \qquad [2.32]$$

where $\mu_{\varepsilon,4} = \mathrm{E}(\varepsilon_i - 0)^4$ [e.g., Kendall and Stuart, 1969, p. 70]. Let $\alpha_{\varepsilon,4}$ denote the kurtosis of ε_t:

$$\alpha_{\varepsilon,4} \;=\; \frac{\mu_{\varepsilon,4}}{\sigma_\varepsilon^2}. \qquad [2.33]$$

Then

$$K_{\varepsilon,4} \;=\; {\sigma_\varepsilon}^4\,(\alpha_{\varepsilon,4} \;-\; 3). \qquad [2.34]$$

2.7 Data Processes

We briefly review three data processes: iid-normal, AR(1) normal, and M/M/1-QWT. They have different correlation structures and different marginal distributions, but all are Markov processes: the distribution of the next value depends (at most) on the current value.

The iid-normal process has "no memory" since its value at time t is independent of all past values. Therefore $\rho(h) = 0$ for all nonzero values of h, $\gamma_0 = 1$ and $\gamma_1 = 0$.

The AR(1) normal time series is $X_t \;=\; \phi\, X_{t-1} \,+\, \varepsilon_t$, where $|\phi| \,<\, 1$ and $\{\varepsilon_t\}$ is a sequence of iid normal random variables with zero mean and variance ${\sigma_\varepsilon}^2$. The autocovariance and autocorrelation functions are $R(h) = {\sigma_\varepsilon}^2\phi^{|h|}/(1-\phi^2)$, and $\rho(h) = \phi^{|h|}$. This geometrically decreasing correlation structure implies that the sum of correlations is $\gamma_0 \;=\; (1+\phi)/(1-\phi)$ and the weighted sum of autocorrelations is $\gamma_1 \;=\; 2\phi/(1-\phi)^2 \;=\; (\gamma_0-1)(\gamma_0+1)/2$. For $\phi \,=\, 0$, the AR(1) reduces to the iid-normal process.

The M/M/1-QWT is a single-queue, single-server system with independent exponential interarrival times (A_i) and exponential service times (S_i). The queue waiting time for the i^{th} customer is $W_i = \max\{0, W_{i-1} + S_{i-1} - A_i\}$, and the expected waiting time is $\mathrm{E}(W_i) = \lambda/[\nu(\nu - \lambda)]$, where λ is the arrival rate and ν is the service rate. The marginal variance $R(0)$, the sum of autocorrelations γ_0, the traffic intensity τ, and the arrival and service rates are related through

$$\tau = \frac{\lambda}{\nu},$$

$$\lambda = \frac{\tau}{1-\tau}\left(\frac{\tau(2-\tau)}{R(0)}\right)^{\frac{1}{2}},$$

and

$$\tau^3 - 4\tau^2 + 5\tau - 2\,\frac{\gamma_0 - 1}{\gamma_0 + 1} = 0,$$

as discussed in Daley [1968, pp. 696-697]. For a given value of γ_0, the M/M/1-QWT has a correlation structure with heavier tails than the AR(1) process.

The weighted sum of autocorrelations can be approximated by

$$\gamma_1 = \left(\frac{\gamma_0^2 - 1}{2}\right)\left(\frac{1}{0.807 + \exp(-\sqrt[3]{\gamma_0})/2}\right). \qquad [2.35]$$

We derived this formula empirically. This formula is exact at $\tau = 0$ (which corresponds $\gamma_0 = 1$ and $\gamma_1 = 0$). We verified computationally for many different values of γ_0 between 1.05 and 200 that Equation 2.35 yields a relative error less than 1%.

The kurtosis is

$$\alpha_4 = \frac{3(\tau^2 - 2\tau + 4)}{\tau(2-\tau)}.$$

The time series $\{X_i\}$, where $X_i = W_i - \mathrm{E}(W_i) + \mu$, is a generalization of $\{W_i\}$ to obtain a desired mean μ.

3. DPSS: THE d-STATE BERNOULLI-DEMAND (s,S)-INVENTORY MARKOV CHAIN

3.1 Introduction

We develop the DPSS stochastic process, a four-parameter Markov chain based on a simple (s, S) inventory model. Our motivation is to obtain a new family of processes for evaluating simulation methods. In particular, we wish to be able to obtain process parameters as simple functions of desired process properties (of the stationary distribution and of the autocovariance structure) and to be able to generate variates [Devroye, 1986] easily. For example, we want to easily obtain various processes with specified mean, variance, and sum of autocorrelations, as well as to easily generate pseudo-random realizations.

Several such test-bed processes exist [Schmeiser and Song, 1989]. Available stationary marginal distributions include the normal (e.g., ARMA processes), exponential (e.g., Lawrance and Lewis [1981, 1982], Dewald, Lewis, and McKenzie [1989]), gamma (e.g., Lewis, McKenzie and Hugus [1989], and Schmeiser and Lal [1982]), and continuous uniform $(0,1)$ (e.g. Melamed [1991]). Many of these processes have the geometrically decreasing correlation structure of the AR(1) (first-order autoregressive) processes; the more-general ARMA(p, q) processes yield arbitrarily complex correlation structure, but require $p + q + 1$ parameters, and solving for parameter values as a function of specified properties is relatively difficult.

A genesis of the DPSS process is as a model of inventory position in discrete time. The acronym is based on the four process parameters: the number of states d, the probability of zero demand p, the lower inventory level s, and the upper inventory level S. Solving for these process parameters in terms of various process properties is closed form. The stationary distribution is discrete uniform. Important dependency properties, such as the lag-h autocorrelation, the sum of autocorrelations, and the autocorrelation center of gravity are easily computed. The lag-h autocorrelation is the sum of $2(d-1)$ sinusoids ($d-1$ cosines and $d-1$ sines) whose frequencies are proportional to the lag h; each sinusoid is damped by its own exponential.

This chapter is organized as follows. In Section 3.2 we define the general discrete (s, S) inventory process, develop its model, and determine some properties. We specialize the results to Bernoulli demands in Section 3.3. In Section 3.4 we provide the equations to determine the process parameters given the desired process properties. Section 3.5 is a summary. Appendix B contains two FORTRAN subroutines: RDPSS generates pseudo-random DPSS(d,p,s,S) realizations and SDPSS calculates the weighted sum of autocorrelations, γ_1, given the sum of autocorrelations, γ_0.

3.2 Process, Model and Properties

In this section we analyze the general process, describe the model as a Markov chain [Ross, 1983 and 1989, and Hoel, 1972], and derive the steady-state distribution and autocorrelation structure.

3.2.1 The Process

Suppose that the product demand $D^{(t)}$ at time t, for $t = 0, 1, \ldots$, has discrete probability distribution defined by

$$\mathrm{P}\left(D^{(t)} = \Delta\ k\right) \;=\; p_k, \qquad \forall\ k \;= 0, 1, \ldots, d-1,$$

and

$$\mathrm{P}\left(D^{(t)} > \Delta\ (d-1)\right) \;=\; p_d,$$

where Δ is a scale parameter. The store follows an $(s,\ S)$ ordering policy, where s and S are minimum and maximum inventory level, respectively. The inventory level at time t, $X^{(t)}$, defines the state of the system. If the inventory level goes below s, then an order is placed and is assumed to be immediately satisfied. Thus $X^{(t)}$ is given by

$$X^{(t)} = \begin{cases} X^{(t-1)} - D^{(t)}, & 0 \leq\ D^{(t)}\ \leq\ X^{(t-1)} - s, \\ S, & D^{(t)}\ >\ X^{(t-1)} - s. \end{cases}$$

We study the relationship between the system parameters (s, S and demand distribution) and the following properties: inventory-level stationary distribution, mean and lag-h autocovariance and autocorrelation for $h \in\ \{0, 1, 2, \ldots\}$.

3.2.2 The Model

The process can be modeled as a discrete Markov chain with d states, where

$$d \;=\; \frac{S - s}{\Delta}\; +\; 1.$$

The state space is

$$E \;=\; \{x_1,\ x_2,\ x_3, \ldots,\ x_d\},$$

where

$$x_k \;=\; S - (k-1)\,\Delta, \qquad \forall\ k \;= 1, 2, \ldots, d.$$

The transition probability matrix is

$$P = \begin{bmatrix} p_0 + p_d & p_1 & p_2 & p_3 & \dots & p_{d-2} & p_{d-1} \\ \sum_{k=d-1}^{d} p_k & p_0 & p_1 & p_2 & \dots & p_{d-3} & p_{d-2} \\ \sum_{k=d-2}^{d} p_k & 0 & p_0 & p_1 & \dots & p_{d-4} & p_{d-3} \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ \sum_{k=2}^{d} p_k & 0 & 0 & 0 & \dots & p_0 & p_1 \\ \sum_{k=1}^{d} p_k & 0 & 0 & 0 & \dots & 0 & p_0 \end{bmatrix}.$$

3.2.3 Properties

Since the Markov chain is irreducible, aperiodic and positive recurrent, it has the unique steady-state distribution

$$\lim_{t \to \infty} \mathrm{P}\left(X^{(t)} = x_k\right) = \pi_k, \qquad \forall\ k = 1, 2, \dots, d.$$

This limiting probability π_k is the long-run fraction of time the chain spends in state k. The steady-state distribution $(\pi_1, \pi_2, \ \dots\ , \pi_d)$ also satisfies the stationarity conditions

$$\begin{cases} (\pi_1, \pi_2, \dots, \pi_d)\ P &= (\pi_1, \pi_2, \dots, \pi_d) \\ \pi_1 \ + \ \pi_2 \ + \ \dots \ + \ \pi_d &= 1, \end{cases}$$

and therefore does not depend on the initial distribution. Grassmann et al. [1985] discuss solution algorithms. The steady-state mean of the d-state Markov chain is

$$\mu \ = \ EX \ = \ \sum_{k=1}^{d} x_k \pi_k \ = \ \sum_{k=1}^{d} (S - (k-1)\Delta)\, \pi_k \ = \ S - \Delta \sum_{k=1}^{d} (k-1)\pi_k. \qquad [3.1]$$

Theorem 3.1 provides a formula to calculate the autocovariance function. Since the steady-state lag-h autocovariance $R(h)$ is an even function, we will assume, without loss of generality, that the lag h is positive, unless otherwise specified.

THEOREM 3.1 The d-state Markov process with transition matrix P has steady-state lag-h autocovariance

$$R(h) = (x^*)^t \operatorname{diag}\{\pi_1, \pi_2, \ldots, \pi_d\}\ P^h\ x^*,$$

where x^* is the vector of mean-adjusted states

$$x^* = [x_1 - \mu,\ x_2 - \mu,\ \ldots\ ,\ x_d - \mu]^t.$$

Proof: In steady-state

$$\begin{aligned} \mathrm{P}(X^{(t)} = x_i,\ X^{(t+h)} = x_k) &= \mathrm{P}(X^{(t)} = x_i)\ \mathrm{P}(X^{(t+h)} = x_k | X^{(t)} = x_i) \\ &= \pi_i\ p_{ik}^{(h)}\ , \end{aligned}$$

where $p_{ik}^{(h)}$ is element (i, k) of the matrix P^h. By definition

$$R(h) = \operatorname{cov}(X^{(t)},\ X^{(t+h)}) = E(X^{(t)} - \mu^{(t)})(X^{(t+h)} - \mu^{(t+h)}).$$

But in steady-state $\mu^{(t)} = \mu^{(t+h)} = \mu$. Then

$$\begin{aligned} R(h) &= E(X^{(t)} - \mu)(X^{(t+h)} - \mu) \\ &= \sum_{i=1}^{d} \sum_{k=1}^{d} (x_i - \mu)(x_k - \mu)\ \mathrm{P}(X^{(t)} = x_i,\ X^{(t+h)} = x_k) \\ &= \sum_{i=1}^{d} \sum_{k=1}^{d} x_i^*\ \pi_i\ p_{ik}^{(h)}\ x_k^* \\ &= (x^*)^t \operatorname{diag}\{\pi_1, \pi_2, \ldots, \pi_d\}\ P^h\ x^*. \end{aligned}$$

3.3 Specializing to Bernoulli Demands

In this section we review the symmetric two-state Markov chain and analyze the d-state Bernoulli-demand (s, S)-inventory Markov process, i.e., we now assume the demand probabilities $\mathrm{P}(D^{(t)} = 0) = p$ and $\mathrm{P}(D^{(t)} = \Delta) = 1 - p$. That is, p is the probability of zero demand. This is a special case of the process described earlier, where $p_0 = p, \quad p_1 = 1 - p, \quad p_k = 0, \ \forall k \in \{2, 3, ..., d\}$. Now the state space is still $\{S, S - \Delta, \ ... \ , s\}$ and the transition matrix is

$$P = \begin{bmatrix} p & 1-p & 0 & \dots & 0 & 0 \\ 0 & p & 1-p & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & p & 1-p \\ 1-p & 0 & 0 & \dots & 0 & p \end{bmatrix}.$$

If $p = 0$ or $p = 1$, the process, once started, is deterministic, and the analysis is trivial. We therefore consider only $0 < p < 1$.

3.3.1 The 2-State Bernoulli-Demand (s, S) Inventory Markov Chain

We quickly review the more specialized case of the symmetric two-state Markov chain, S2MC [Song, 1988, and Bhat and Lal, 1990]. The state space is $\{S, s\}$ and the transition probability matrix is

$$P = \begin{bmatrix} p & 1-p \\ 1-p & p \end{bmatrix}.$$

Since the matrix is doubly stochastic, the unique stationary distribution of the chain is uniform with mean

$$\mu \ = \ \frac{S \ + \ s}{2},$$

and variance

$$R(0) \ = \ \frac{(S \ - \ s)^2}{4}. \qquad [3.2]$$

The lag-h autocorrelation is

$$\rho(h) \ = \ (2p \ - \ 1)^h. \qquad [3.3]$$

Therefore,

$$\begin{aligned}\sum_{h=-n}^{n} \rho(h) &= \frac{p - (2p - 1)^{n+1}}{1 - p}, \\ \sum_{h=-n}^{n} |h|\,\rho(h) &= \frac{(2p-1)\ \{1 - (2p-1)^n\,[2n(1-p)+1]\}}{2\ (1-p)^2}\end{aligned}$$

and

$$\sum_{h=-n}^{n} \left(1 - \frac{|h|}{n}\right) \rho(h) = \frac{p}{1 - p} - \frac{(2p - 1)\ [1 - (2p - 1)^n]}{2n\ (1 - p)^2}.$$

This last equation together with Equation 2.3 yields the variance of the sample mean V of n consecutive observations. The asymptotic V (Equation 2.5) is a function of

$$\gamma_0 = \sum_{h=-\infty}^{\infty} \rho(h) = \frac{p}{1 - p}. \qquad [3.4]$$

The weighted sum of all the correlations is

$$\gamma_1 = \sum_{h=-\infty}^{\infty} |h|\,\rho(h) = \frac{2p - 1}{2\ (1 - p)^2} = \frac{\gamma_0^2 - 1}{2}. \qquad [3.5]$$

Like the AR(1) and EAR(1) processes, the S2MC has a geometrically decreasing correlation structure and the relationships among γ_0, γ_1 and $\rho(1)$ are the same.

3.3.2 The Diagonalization of the Bernoulli-Demand Transition Matrix

Theorem 3.1 shows that the steady-state lag-h autocovariance is a function of $P^h = \left[p_{ij}^{(h)}\right]_{i,j=1}^{d}$. For Bernoulli demand, the elements $p_{ij}^{(h)}$ of P^h can be calculated using the binomial probabilities, as follows. Consider that there is a success at time t if the Markov chain changes state. Let $b(y, h, 1-p)$ denote the (binomial) probability of y successes in h independent Bernoulli trials with probability $1 - p$ of success:

$$b(y, h, 1 - p) = \begin{cases} \binom{h}{y}\ p^{h-y}\ (1 - p)^y, & y \in \{0, 1, ..., h\}, \\ 0, & \text{otherwise,} \end{cases}$$

where

$$\begin{pmatrix} h \\ y \end{pmatrix} = \frac{h!}{y!\,(h-y)!}.$$

Let

$$u = \begin{cases} j - i, & j \geq i, \\ j - i + d, & j < i. \end{cases}$$

Then $p_{ij}^{(h)} = P\left(X^{(t+h)} = x_j | X^{(t)} = x_i\right)$ is the probability of $u, u+d, u+2d, \ldots,$ or $\lfloor (h-u)/d \rfloor$ demands in h trials; that is,

$$p_{ij}^{(h)} = \sum_{k=0}^{\lfloor (h-u)/d \rfloor} b(u + kd, h, 1-p).$$

For any h, $p_{ij}^{(h)} = p_{i-1,j-1}^{(h)}$ $\forall i, j \geq 2$ and $p_{i,1}^{(h)} = p_{i-1,d}^{(h)}$ $\forall i \geq 2$. That is P^h is a circulant matrix [Brockwell and Davis, 1991], and therefore, all rows can be obtained by shifting the 1st row. Nevertheless we are interested on the diagonalization of the transition probability matrix, P, so that we can compute lag-h autocovariances in $O(d)$ rather than $O(h)$ operations.

Brockwell and Davis [1991, p. 133 and 134], letting

$$r_k = \exp\left(i\frac{2(k-1)\pi}{d}\right), \qquad \forall\, k \in \{1, 2, ..., d\},$$

state that the eigenvalues of P are

$$\begin{aligned} \lambda_k &= \sum_{h=1}^{d} p_{1,1+h}\; r_k^{-h} \\ &= p + \frac{1-p}{r_k}, \qquad \forall\, k \in \{1, 2, ..., d\}, \end{aligned}$$

with corresponding orthonormal left eigenvectors

$$v_k = \frac{1}{\sqrt{d}}\,[1, r_k, r_k^2, \;\ldots\;, r_k^{d-1}],$$

and therefore

$$P^h = W^{-1}\, D^h\, W, \qquad \forall\, h, \tag{3.6}$$

where $W = [\ v_k\ ]_{k=1}^{d}$, $W^{-1} = \overline{W}^t$ (conjugate transpose), and $D = \mathrm{diag}\{\lambda_1,\ \lambda_2, \ldots,\ \lambda_d\}$.

Some algebra yields

$$|\lambda_k| \;=\; \sqrt{1 \;-\; 2p(1-p)\left[1 - \cos\left(\frac{2(k-1)\pi}{d}\right)\right]}, \qquad [3.7]$$

and

$$\arg(\lambda_k) \;=\; -\arctan\left[\frac{(1-p)\sin\frac{2(k-1)\pi}{d}}{p \;+\; (1-p)\cos\frac{2(k-1)\pi}{d}}\right], \quad \forall k \in \{1, 2, \ldots, d\}. \qquad [3.8]$$

Thus $\lambda_1 = 1$ and $|\lambda_k| < 1, \ \forall\, k \in \{2, 3, \ldots, d\}$.

3.3.3 The Steady-State Distribution

When the state transition probabilities are defined by a Bernoulli distribution, the steady-state distribution is uniform over the state space $\{S, S - \Delta, \ldots, s\}$.

THEOREM 3.2 The unique steady-state distribution of the d-state DPSS process is

$$[\pi_k]_{k=1}^{d} \;=\; \left[\frac{1}{d}\right]_{k=1}^{d}.$$

Proof: Since the process is ergodic and doubly stochastic the result follows.

The next proposition, which follows from Equation 3.1 and Theorem 3.2, relates the mean to the process parameters.

THEOREM 3.3 The steady-state mean is

$$\mu \;=\; \frac{S+s}{2} \;=\; S \;-\; \frac{d-1}{2}\Delta.$$

Lemma 3.1 provides a formula to compute the lag-h autocovariance in $O(d^2)$ operations. By exchanging the order of the summations and developing the cosine of a sum as the difference between the product of cosines and the product of sines, an $O(d)$ formula is derived in Theorem 3.4.

LEMMA 3.1 Let x^* be the vector, and W, W^{-1} and D the matrices, defined previously. Then the steady-state lag-h autocovariance is

$$R(h) \;=\; \frac{1}{d}\,(x^*)^t\, W^{-1} D^h W x^* \;=\; \frac{\Delta^2}{d} \sum_{j=1}^{d} a_j(h)\; c_j,$$

where

$$a_j(h) \;=\; \sum_{k=1}^{d} |\lambda_k|^h \;\cos\left(h \;\arg\lambda_k \;+\; \frac{2(j-1)(k-1)}{d}\;\pi\right),$$

and

$$c_j \;=\; \frac{(d-1)(d+1)}{12} \;-\; \frac{(d+1-j)(j-1)}{2}.$$

Proof: By Theorem 3.1 and Equation 3.6

$$R(h) \;=\; \frac{1}{d}\,(x^*)^t\;\; W^{-1}\;\, D^h\;\, W\;\, x^*.$$

Let $\theta_k(h) = h \arg\lambda_k$ and $A_k(h) = |\lambda_k|^h$. Then

$$D^h \;=\; \mathrm{diag}\{A_1(h), A_2(h)e^{i\theta_2(h)}, \ldots, A_d e^{i\theta_d(h)}\}$$

and $A_k(h) = A_{d+2-k}(h)$, and $\theta_k(h) = -\theta_{d+2-k}(h)$, $\forall\; k = 2, 3, ..., d$. By using simple algebraic manipulation

$$W^{-1}D^hW \;=\; \frac{1}{d}\begin{bmatrix} a_1(h) & a_2(h) & a_3(h) & \ldots & a_{d-1}(h) & a_d(h) \\ a_d(h) & a_1(h) & a_2(h) & \ldots & a_{d-2}(h) & a_{d-1}(h) \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ a_2(h) & a_3(h) & a_4(h) & \ldots & a_d(h) & a_1(h) \end{bmatrix},$$

where

$$a_j(h) \;=\; \sum_{k=1}^{d} A_k(h)\; e^{i\left(\theta_k(h) \;+\; \frac{2(j-1)(k-1)}{d}\;\pi\right)}.$$

Since $A_k(h) \;=\; A_{d+2-k}(h)$, and $\theta_k(h) \;=\; -\theta_{d+2-k}(h)$, the imaginary part of this summation is cancelled and the $a_j(h)$ formula follows.

Let $y = (d+1)/2$. Then $x^* = (y-1, y-2, \ldots, y-d)\,\Delta$ and

$$(x^*)^t\, W^{-1} D^h W x^* \;=\; \Delta^2 \sum_{j=1}^{d} a_j(h)\; c_j,$$

where

$$\begin{aligned}
c_j &= \frac{1}{d}\left[\sum_{k=1}^{d+1-j}(y-k)(y+1-j-k) + \sum_{k=1}^{j-1}(y-d+j-1-k)(y-k)\right] \\
&= \frac{1}{d}\left[\sum_{k=1}^{d}(y-k)(y+1-j-k) + d\sum_{k=d+2-j}^{d}(y-k)\right] \\
&= \frac{1}{d}\left[\sum_{k=1}^{d}\left[y(y+1-j) - (2y+1-j)k + k^2\right] + d\sum_{k=d+2-j}^{d}(y-k)\right].
\end{aligned}$$

Finally, each of these sums is calculated and y is replaced by $(d+1)/2$, thus obtaining the c_j result and proving this proposition.

We now present Lemma 3.2, which is used in the proof of Theorem 3.4.

LEMMA 3.2

$$\sum_{j=1}^{d} c_j \;=\; 0.$$

Proof:

$$\begin{aligned}
\sum_{j=1}^{d} c_j &= \sum_{j=1}^{d}\left[\frac{d^2-1}{12} - \frac{d(j-1)}{2} + \frac{(j-1)^2}{2}\right] \\
&= \frac{d(d-1)(d+1)}{12} - \frac{d^2(d-1)}{4} + \frac{d(d-1)(2d-1)}{12} \\
&= 0.
\end{aligned}$$

Theorem 3.4 provides a formula to calculate the steady-state lag-h autocovariance in $O(d)$ arithmetic operations.

THEOREM 3.4 The steady-state lag-h autocovariance is

$$R(h) \;=\; \frac{\Delta^2}{d}\sum_{k=2}^{d}|\lambda_k|^h\left\{c_k^*\cos\left(h\;\arg\lambda_k\right) \;-\; c_k^{**}\sin\left(h\;\arg\lambda_k\right)\right\},$$

where, $\forall k = 2, \ldots, d$,

$$c_k^* = \sum_{j=1}^{d} c_j \cos\left(\frac{2(j-1)(k-1)}{d}\,\pi\right) = \sum_{j=1}^{d} \frac{j^2-(d+2)j}{2} \cos\left(\frac{2(j-1)(k-1)}{d}\,\pi\right),$$

and

$$c_k^{**} = \sum_{j=1}^{d} c_j \sin\left(\frac{2(j-1)(k-1)}{d}\,\pi\right) = \sum_{j=1}^{d} \frac{j^2-(d+2)j}{2} \sin\left(\frac{2(j-1)(k-1)}{d}\,\pi\right).$$

Proof: By Lemma 3.1,

$$\begin{aligned} R(h) &= \frac{\Delta^2}{d} \sum_{j=1}^{d}\sum_{k=1}^{d} c_j\, |\lambda_k|^h \cos\left(h \arg\lambda_k + \frac{2(j-1)(k-1)}{d}\,\pi\right) \\ &= \frac{\Delta^2}{d} \sum_{k=1}^{d} |\lambda_k|^h \sum_{j=1}^{d} c_j \left\{\cos(h \arg\lambda_k) \cos\left(\frac{2(j-1)(k-1)}{d}\,\pi\right)\right. \\ &\qquad \left. - \sin(h \arg\lambda_k) \sin\left(\frac{2(j-1)(k-1)}{d}\,\pi\right)\right\}. \end{aligned}$$

By Lemma 3.1 (definition of c_j) and Lemma 3.2 the result follows.

Theorems 3.5, 3.6 and 3.7 provide formulas to calculate the variance, autocovariance and autocorrelation functions.

THEOREM 3.5 The steady-state variance is

$$R(0) = \frac{d^2-1}{12}\,\Delta^2.$$

Proof: Applying Lemma 3.1 for $h = 0$ we get $a_1(0) = d$ and $a_j(0) = 0$, $\forall j \neq 1$. Then

$$\begin{aligned} R(0) &= \frac{1}{d}(x^*)^t\, W^{-1}\, D^0\, W\, x^* \\ &= \frac{\Delta^2}{d} \sum_{j=1}^{d} a_j(0) c_j \\ &= \Delta^2\, c_1. \end{aligned}$$

Again, by Lemma 3.1 the result follows. Alternately, the variance (and other moments) can be derived from the uniform marginal distribution.

THEOREM 3.6 The steady-state lag-h autocorrelation is

$$\rho(h) = \frac{12}{d\,(d^2-1)} \sum_{k=2}^{d} |\lambda_k|^h \,\{c_k^* \cos(h \;\arg \lambda_k) \;-\; c_k^{**} \sin(h \;\arg \lambda_k)\}\,,$$

where c_k^* and c_k^{**} are defined in Theorem 3.4.

Proof: Using the definition of correlation, Theorem 3.4 and Theorem 3.5 the result follows.

THEOREM 3.7 The steady-state lag-1 autocovariance is

$$R(1) = \frac{(d-1)(d-5+6p)}{12}\,\Delta^2,$$

and the steady-state lag-1 autocorrelation is

$$\rho(1) = \frac{d-5+6p}{d+1}.$$

Proof: Applying Lemma 3.1 for $h = 1$, we get $a_1(1) = d\,p$, $a_2(1) = d(1-p)$ and $a_j(1) = 0$, $\forall j \neq 1, 2$. Then

$$\begin{aligned} R(1) &= \frac{\Delta^2}{d} \sum_{j=1}^{d} a_j(1)\; c_j \\ &= \frac{\Delta^2}{d} \left(a_1(1)\; c_1 \;+\; a_2(1)\; c_2\right) \\ &= \Delta^2 \left(p\,\frac{d^2-1}{12} \;+\; (1-p)\frac{d^2-1}{12} \;-\; (1-p)\frac{d-1}{2}\right). \end{aligned}$$

The lag-1 autocovariance result is now obtained from simple algebraic manipulation. The lag-1 autocorrelation is derived using the definition. Notice that $\rho(1) = p$ when $d = 5$.

We now present Lemmas 3.3 through 3.7. They are needed to prove Theorem 3.8, a simple formula relating γ_0 to the parameter p of the Bernoulli demands.

LEMMA 3.3

$$\sum_{j=1}^{d} j\ c_j \;=\; -\frac{d(d-1)(d+1)}{24}.$$

Proof:

$$\begin{aligned}
\sum_{j=1}^{d} j\ c_j &= \sum_{j=1}^{d} \left[\frac{(d-1)(d+1)}{12} j \;-\; \frac{(d+1-j)(j-1)}{2} j \right] \\
&= \frac{d(d-1)(d+1)^2}{24} \;-\; \sum_{j=1}^{d-1} \frac{-j^3 + (d-1)j^2 + dj}{2} \\
&= \frac{d(d-1)(d+1)^2}{24} \;-\; \frac{d(d-1)(d^2+3d+2)}{24} \\
&= -\frac{d(d-1)(d+1)}{24}.
\end{aligned}$$

LEMMA 3.4

$$\sum_{h=0}^{N} |\lambda_k|^h \cos(h \arg \lambda_k) \;=\; \begin{cases} N, & k = 1, \\ \frac{1}{2(1-p)} \;+\; \varepsilon_1(N), & k = 2, 3, \ldots, d, \end{cases}$$

where

$$\varepsilon_1(N) \;=\; \frac{-\,|\lambda_k|^{N+1} \cos((N+1) \arg \lambda_k) \;+\; |\lambda_k|^{N+2} \cos(N \arg \lambda_k)}{2\,(1-p)^2\;\;[1 - \cos(\arg \lambda_k)]}.$$

Proof: For $k = 1$, the result is trivial since $|\lambda_1| = 1$ and $\arg \lambda_1 = 0$. For $k \in \{2, 3, \ldots, d\}$,

$$\sum_{h=0}^{N} |\lambda_k|^h \cos(h \arg \lambda_k) \;=\; \frac{1 \;-\; |\lambda_k| \cos(\arg \lambda_k)}{1 \;-\; 2|\lambda_k| \cos(\arg \lambda_k) \;+\; |\lambda_k|^2} \;+$$
$$\frac{-\,|\lambda_k|^{N+1} \cos((N+1) \arg \lambda_k) \;+\; |\lambda_k|^{N+2} \cos(N \arg \lambda_k)}{1 \;-\; 2|\lambda_k| \cos(\arg \lambda_k) \;+\; |\lambda_k|^2}.$$

Using Equations 3.7 and 3.8 the result follows.

LEMMA 3.5

$$\sum_{h=0}^{N} |\lambda_k|^h \sin(h \arg \lambda_k) \;=\; \begin{cases} 0, & k = 1, \\ -\frac{\cos\left(\frac{k-1}{d}\pi\right)}{2(1-p)\sin\left(\frac{k-1}{d}\pi\right)} \;+\; \varepsilon_2(N), & k = 2, 3, \ldots, d, \end{cases}$$

where

$$\varepsilon_2(N) \;=\; \frac{-|\lambda_k|^{N+1}\sin\left((N+1)\arg\lambda_k\right) \;+\; |\lambda_k|^{N+2}\sin\left(N\arg\lambda_k\right)}{2\;(1-p)^2\;\left[1-\cos\left(\arg\lambda_k\right)\right]}.$$

Proof: For $k=1$, the result is trivial since $\arg\lambda_1 = 0$. For $k \in \{2,3,\ldots,d\}$,

$$\begin{aligned}
\sum_{h=0}^{N}|\lambda_k|^h\sin\left(h\arg\lambda_k\right) \;&=\; \frac{|\lambda_k|\sin\left(\arg\lambda_k\right)}{1 \;-\; 2|\lambda_k|\cos\left(\arg\lambda_k\right) \;+\; |\lambda_k|^2} \\
&\qquad -\frac{|\lambda_k|^{N+1}\sin\left((N+1)\arg\lambda_k\right) \;+\; |\lambda_k|^{N+2}\sin\left(N\arg\lambda_k\right)}{1 \;-\; 2|\lambda_k|\cos\left(\arg\lambda_k\right) \;+\; |\lambda_k|^2} \\
&=\; \frac{-(1-p)\sin\left(\frac{2(k-1)}{d}\pi\right)}{2\;(1-p)^2\;\left[1-\cos\left(\frac{2(k-1)}{d}\pi\right)\right]} \;+\; \varepsilon_2(N) \\
&=\; -\frac{\sin\left(\frac{k-1}{d}\pi\right)\;\cos\left(\frac{k-1}{d}\pi\right)}{(1-p)\left[1-\cos^2\left(\frac{k-1}{d}\pi\right)+\sin^2\left(\frac{k-1}{d}\pi\right)\right]} \;+\; \varepsilon_2(N) \\
&=\; -\frac{\cos\left(\frac{k-1}{d}\pi\right)}{2(1-p)\sin\left(\frac{k-1}{d}\pi\right)} \;+\; \varepsilon_2(N).
\end{aligned}$$

LEMMA 3.6

$$\sum_{k=1}^{d}\sum_{j=1}^{d}c_j\cos\left(\frac{2(j-1)(k-1)}{d}\pi\right) \;=\; \sum_{k=2}^{d}\sum_{j=1}^{d}c_j\cos\left(\frac{2(j-1)(k-1)}{d}\pi\right) \;=\; \frac{d\;(d^2-1)}{12}.$$

Proof: For $j=1$,

$$\sum_{k=1}^{d}\cos\left(\frac{2(j-1)(k-1)}{d}\pi\right) = \sum_{k=1}^{d}\cos(0) = d.$$

For any $j>1$,

$$\sum_{k=1}^{d}\cos\left(\frac{2(j-1)(k-1)}{d}\pi\right) = 0.$$

Then

$$\sum_{k=1}^{d}\sum_{j=1}^{d}c_j\cos\left(\frac{2(j-1)(k-1)}{d}\pi\right) \;=\; d\;c_1 \;=\; \frac{d\;(d^2-1)}{12}.$$

LEMMA 3.7

$$\sum_{k=2}^{d}\sum_{j=1}^{d}c_j\frac{\cos\left(\frac{k-1}{d}\pi\right)\;\sin\left(\frac{2(j-1)(k-1)}{d}\pi\right)}{\sin\left(\frac{k-1}{d}\pi\right)} \;=\; 0.$$

Proof: Define

$$Q_{kj} = \frac{\cos\left(\frac{k-1}{d}\pi\right)\ \sin\left(\frac{2(j-1)(k-1)}{d}\pi\right)}{\sin\left(\frac{k-1}{d}\pi\right)}.$$

$Q_{k1} = 0$. For $j > 1$, by induction, yields

$$Q_{kj} = 1+2\cos\left(\frac{2(k-1)}{d}\pi\right)+\cdots+2\cos\left(\frac{2(j-2)(k-1)}{d}\pi\right)+\cos\left(\frac{2(j-1)(k-1)}{d}\pi\right).$$

Then,

$$\begin{aligned}\sum_{k=2}^{d}\sum_{j=1}^{d} c_j \frac{\cos\left(\frac{k-1}{d}\pi\right)\ \sin\left(\frac{2(j-1)(k-1)}{d}\pi\right)}{\sin\left(\frac{k-1}{d}\pi\right)} &= \sum_{k=2}^{d}\sum_{j=1}^{d} c_j\ Q_{kj} \\ &= \sum_{j=2}^{d} c_j \sum_{k=2}^{d} Q_{kj} \\ &= \sum_{j=2}^{d} c_j \sum_{k=2}^{d}\left[1+2\cos\left(\frac{2(k-1)}{d}\pi\right)+\cdots\right. \\ &\qquad + 2\cos\left(\frac{2(j-2)(k-1)}{d}\pi\right) \\ &\qquad \left. + \cos\left(\frac{2(j-1)(k-1)}{d}\pi\right)\right].\end{aligned}$$

For any $j > 1$,

$$\sum_{k=1}^{d} \cos\left(\frac{2(j-1)(k-1)}{d}\pi\right) = 0,$$

then,

$$\begin{aligned}\sum_{k=2}^{d}\sum_{j=1}^{d} c_j \frac{\cos\left(\frac{k-1}{d}\pi\right)\ \sin\left(\frac{2(j-1)(k-1)}{d}\pi\right)}{\sin\left(\frac{k-1}{d}\pi\right)} &= \sum_{j=2}^{d} c_j\ (d+2-2j) \\ &= -c_1\ d\ +\ \sum_{j=1}^{d} c_j\ (d+2-2j).\end{aligned}$$

Now, after some algebraic manipulation and by Lemmas 3.2 and 3.3 the result follows.

THEOREM 3.8 The sum of the autocorrelations for the d-state DPSS process is

$$\gamma_0 = \sum_{h=-\infty}^{\infty} \rho(h) = \frac{p}{1-p}.$$

Proof: By Theorem 3.6

$$\begin{aligned}\sum_{h=0}^{N}\rho(h) &= \frac{12}{d(d^2-1)}\sum_{h=0}^{N}\sum_{k=2}^{d}|\lambda_k|^h\sum_{j=1}^{d}c_j\left[\cos\left(\frac{2(j-1)(k-1)\pi}{d}\right)\cos\left(h\arg\lambda_k\right)\right.\\ &\quad\left.-\sin\left(\frac{2(j-1)(k-1)\pi}{d}\right)\sin\left(h\arg\lambda_k\right)\right]\\ &= \frac{12}{d(d^2-1)}\sum_{j=1}^{d}c_j\sum_{k=2}^{d}\left[\cos\left(\frac{2(j-1)(k-1)\pi}{d}\right)\sum_{h=0}^{N}|\lambda_k|^h\cos\left(h\arg\lambda_k\right)\right.\\ &\quad\left.-\sin\left(\frac{2(j-1)(k-1)\pi}{d}\right)\sum_{h=0}^{N}|\lambda_k|^h\sin\left(h\arg\lambda_k\right)\right].\end{aligned}$$

By Lemmas 3.4 and 3.5, and since $\varepsilon_1(N)$ and $\varepsilon_2(N)$ are asymptotically zero,

$$\begin{aligned}\sum_{h=0}^{\infty}\rho(h) &= \frac{12}{d(d^2-1)}\sum_{j=1}^{d}c_j\left[\sum_{k=2}^{d}\frac{1}{2(1-p)}\cos\left(\frac{2(j-1)(k-1)\pi}{d}\right)\right.\\ &\quad\left.+\sum_{k=2}^{d}\frac{\cos\left(\frac{(k-1)\pi}{d}\right)}{2(1-p)\sin\left(\frac{(k-1)\pi}{d}\right)}\sin\left(\frac{2(j-1)(k-1)\pi}{d}\right)\right].\end{aligned}$$

By Lemma 3.7

$$\begin{aligned}\sum_{h=0}^{\infty}\rho(h) &= \frac{12}{d(d^2-1)}\sum_{j=1}^{d}c_j\sum_{k=2}^{d}\frac{1}{2(1-p)}\cos\left(\frac{2(j-1)(k-1)\pi}{d}\right)\\ &= \frac{6}{d(d^2-1)(1-p)}\left[c_1(d-1)-(c_2+c_3+\cdots+c_d)\right]\\ &= \frac{1}{2(1-p)}.\end{aligned}$$

The result follows from this equation since

$$\sum_{h=-\infty}^{\infty}\rho(h) = -1 + 2\sum_{h=0}^{\infty}\rho(h).$$

Theorem 3.9 provides a formula to calculate γ_1 as a function of γ_0 (or p) and d.

THEOREM 3.9 The weighted sum of the autocorrelations for the d-state DPSS process is

$$\begin{aligned}\gamma_1 = \sum_{h=-\infty}^{\infty}|h|\,\rho(h) &= \frac{p-\phi_1(d)}{(1-p)^2}\\ &= \left[\phi_2(d)\,(\gamma_0+1)-1\right](\gamma_0+1),\end{aligned}$$

where

$$\phi_1(d) = \frac{11 + d^2}{30},$$

and

$$\phi_2(d) = \frac{19 - d^2}{30}.$$

From these equations it can be easily proven that $\gamma_1 < 0$ whenever $d \geq 5$. For $d = 4$ and $p < 0.9$, $d = 3$ and $p < 2/3$ and $d = 2$ and $p < 0.5$, γ_1 is also negative.

Technically, Theorem 3.9 is a conjecture. It was derived empirically. The subroutine SDPSS, given in Appendix B, was used to compute γ_1. The software package Mathematica was used to find the exact fit to the data. An exhaustive test was performed for $d = 2, 3, \ldots, 50$. For any d in this range and for any $10^{-3} \leq \gamma_0 \leq 10^3$, the values of γ_1 obtained computationally and those obtained using Theorem 3.9 coincide to, at least, 6 decimal places. For γ_0 outside that range the two γ_1-values also coincide whenever the subroutine SDPSS converged. In addition, Equation 3.5, which was derived analytically for d=2, is a special case of Theorem 3.9. Therefore, in spite of not providing an analytical proof, these two facts indicate that Theorem 3.9 is true.

Figures 1 and 2 show the correlation structure for a particular DPSS process.

The correlation structure, as expected from Theorem 3.6, has an oscillating behavior. The amplitude of the oscillations decrease with h, approaching zero. The damped frequency and the damping rate are determined by the eigenvalues of P (one-step transition probability matrix) which are functions of p (probability of zero demand) and d (number of states) — see Equations 3.7 and 3.8. The selection of these two parameters is, therefore, critical to obtain the desired properties (e.g.. γ_0

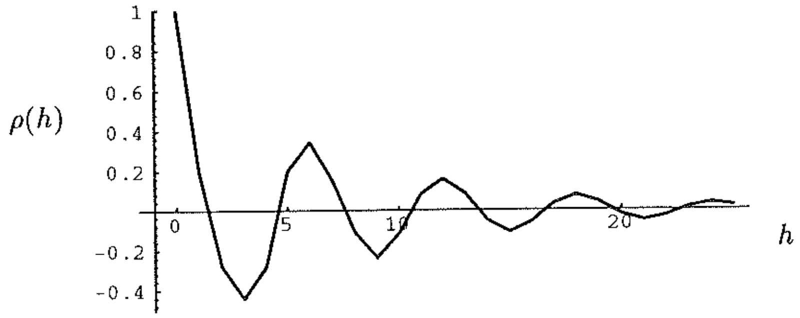

Figure 3.1 Correlation Structure for a DPSS(5, 0.2, s, S) for which $\gamma_0 = 0.25$ and $\gamma_1 = -1.5625$

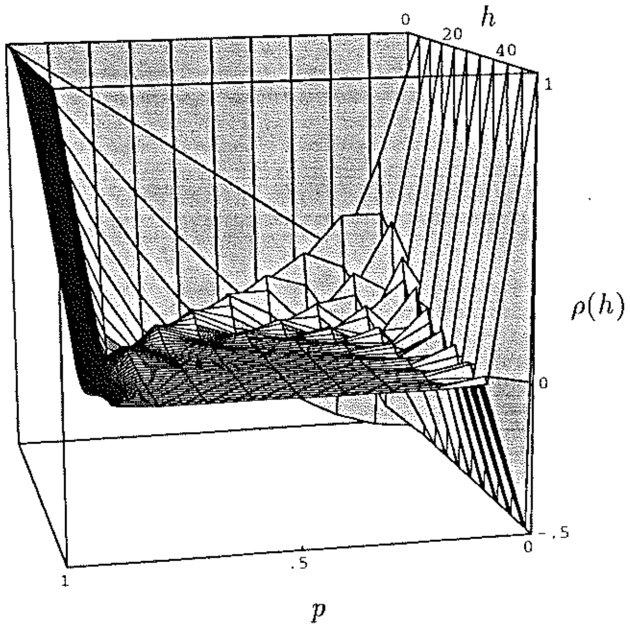

Figure 3.2 Correlation Structure for a DPSS(5, p, s, S)

and γ_1) and provides flexibility to generate processes with quite different characteristics. For example, the autocorrelation function converges to zero fastest when $p = 1/2$ ($\gamma_0 = 1$), which corresponds to smaller amplitudes of the eigenvalues (Equation 3.7). For small and large values of p (e.g., 0.001 and 0.999, respectively) the convergence is very slow. On the other hand, the damped frequency increases when p decreases (Equation 3.8). In the limit, for $p = 0$, the system changes state every time and the frequency is maximum.

The number of states, d, also plays an important role. When d increases the autocorrelation function approaches zero slower. This conclusion can be drawn from Equation 3.7. Another way to see this is that as d increases the relative distance between two adjacent states decreases and therefore the system changes relatively less when it jumps from one state to the next. The damped frequency also decreases when d increases (Equation 3.8). The recurrence time is negative binomial with d successes and probability of success $1 - p$. Therefore, the expected recurrence time (the reversal of the damped frequency) is $d/(1 - p)$, an increasing function of d.

3.4 Fitting and Generating the DPSS Process

The DPSS(d,p,s,S) process can be easily constructed and simulated. The parameters d, s and S specify the marginal distribution. The parameters d and p specify the correlation structure. For example, given the mean μ, variance $R(0)$, and lag-1 autocorrelation $\rho(1)$, we can calculate the parameters of the d-state Bernoulli-demand (s, S)-inventory Markov chain by means of

$$\Delta = \sqrt{\frac{12\ R(0)}{d^2 - 1}}, \qquad [3.9]$$

$$s = \mu - \frac{d-1}{2}\Delta, \qquad [3.10]$$

$$S = \mu + \frac{d-1}{2}\Delta, \qquad [3.11]$$

and

$$p = \frac{(d+1)\,\rho(1) - (d-5)}{6} = \frac{\gamma_0}{1+\gamma_0}. \qquad [3.12]$$

Appendix B contains subroutine RDPSS, which generates pseudo-random realizations of the DPSS(d,p,s,S) process.

Another possibility is to specify γ_1 instead of d, since by Theorem 3.9, and given γ_0, d is determined by γ_1. Because d is an integer some adjustments on the values of γ_1 and/or γ_0 may be needed. Also, the number of observations, n, and the variance of the sample mean, V, can be specified instead of the data variance, $R(0)$. By Equation 2.7 and Theorem 3.9

$$\frac{n^2\ V}{n\gamma_0 - [\phi_2(d)\ (\gamma_0+1) - 1]\,(\gamma_0+1)} \xrightarrow{n\to\infty} R(0). \qquad [3.13]$$

If n is not large enough, Equation 2.3 can be used instead:

$$R(0) = \frac{n\ V}{\sum_{h=-n}^{n}\left(1-\frac{|h|}{n}\right)\rho(h)}. \qquad [3.14]$$

To compute $R(0)$, employing this exact formula, a slight variation of subroutine SDPSS (see Appendix B) can be used.

3.5 Summary

The DPSS stochastic process, a four-parameter Markov chain based on a simple (s,S) inventory model, is developed. The process has a discrete uniform stationary distribution (which depends on d, s, and S), and a relatively complex correlation structure with a damped oscillating behavior (which depends on d and p). The damping frequency increases when p or d decreases. The damping rate decreases when d increases, and given d it is maximum for $p = 1/2$ ($\gamma_0 = 1$), and decreases when $p \longrightarrow 0$ or $p \longrightarrow 1$. The parameter p completely determines γ_0 while γ_1 is a function of d and p. The selection of parameters p and d provides flexibility to generate quite different autocorrelations. Nevertheless, the process parameters are simple functions of desired process properties, and can be calculated using the formulas given in Section 3.5.

4. COVARIANCES AND CORRELATIONS OF OVERLAPPING BATCH MEANS AND BARTLETT ESTIMATORS

4.1 Introduction

Classical issues in output analysis of a simulation study include (1) how to obtain good estimates of some measure of performance, (2) how to evaluate the quality of these estimates, and (3) how to determine the goodness of the quality measure. Often in simulation the measure of performance is a population mean, the point estimator is the sample mean, and the goodness of the point estimator is measured by its standard error.

Several procedures for estimating V, the variance of the sample mean, from stationary autocorrelated data have been proposed. For example, direct [Hannan, 1957; Moran, 1975], regenerative [Crane and Iglehart, 1975; Crane and Lemoine, 1977; Glynn and Iglehart, 1986], non-overlapping-batch-means (NBM) [Schmeiser, 1982], overlapping-batch-means (OBM) [Meketon and Schmeiser, 1984], standardized-time-series [Schruben, 1983; Glynn and Iglehart, 1990] and its variations [Foley and Goldsman, 1988; Goldsman, Kang, and Seila, 1993], and spectral [Heidelberger and Welch, 1981; Priestley, 1992]. No type of estimator dominates the others in terms of computational and statistical properties across all types of time-series data.

Our main objective is to develop robust and computationally efficient methodology to estimate the variance of the sample mean. The results of this chapter (summarized earlier in Pedrosa and Schmeiser [1993b]) are motivated by previous studies [Politis and Romano, 1994, Song and Schmeiser, 1988a] suggesting that linear combinations of estimators of the variance of the sample mean lead to better estimators; i.e., with smaller mean squared error (MSE) than the component estimators. The key to using linear combinations, as discussed in Section 4.2, is to derive the asymptotic covariance/correlation between component estimators. We study OBM estimators because they are conceptually simple methods, they are easy to compute, and they can be written as quadratic forms, which leads to tractable analysis. We study Bartlett estimators because they are widely accepted and quite similar to OBM estimators.

In Section 4.2 we show that any method to select optimal linear-combination parameters must consider the estimator covariances. In Section 4.3 we develop the theoretical structure needed to derive the asymptotic results of Section 4.4. In Section 4.5 we show empirically that the asymptotic correlation formula provides a good approximation to the finite-sample correlation, except when the batch size is quite small or larger than half the sample size. Here "small" is with respect to the sum of autocorrelations.

4.2 Linear-Combination Estimators of V

There are several examples in the literature of using linear combinations of estimators to obtain a better estimator in terms of statistical properties: small variance, bias or mean squared error. The OBM estimator can be viewed as a linear combination of NBM estimators [Meketon and Schmeiser, 1984]. Schruben [1983] considered a linear combination of the STS.A and the NBM estimators, which are asymptotically independent. Politis and Romano [1994] propose a linear combination of two Bartlett estimators of the spectral density with different bandwidths for the reduction of the bias.

The linear combination of OBM's is

$$\widehat{V}^{(LC)} = \sum_{i=1}^{p} \alpha_i \, \widehat{V}_i, \qquad [4.1]$$

where p is the number of components, the α_i's are the L.C. coefficients and the $\widehat{V}_i$'s are the OBM component estimators applied to the same data but each with a different batch size m_i. By definition, the bias of $\widehat{V}^{(LC)}$ is

$$\begin{aligned} \text{bias}\left(\widehat{V}^{(LC)}\right) &= E\left(\widehat{V}^{(LC)}\right) - V \\ &= \sum_{i=1}^{p} \alpha_i \, \text{bias}\left(\widehat{V}_i\right) + \left(-1 + \sum_{i=1}^{p} \alpha_i\right) V, \end{aligned} \qquad [4.2]$$

and the variance is

$$\text{var}\left(\widehat{V}^{(LC)}\right) = \sum_{i=1}^{p}\sum_{j=1}^{p} \alpha_i \, \alpha_j \, \text{cov}\left(\widehat{V}_i, \widehat{V}_j\right). \qquad [4.3]$$

Since the linear-combination variance depends upon the various estimator covariances, any method to select optimal linear-combination parameters must consider the estimator covariances. This application motivates our interest in asymptotic properties of $\text{cov}\left(\widehat{V}_i, \widehat{V}_j\right)$ and $\text{corr}\left(\widehat{V}_i, \widehat{V}_j\right)$.

4.3 Intermediate Results

In this section we develop the theoretical structure needed to derive the asymptotic results of Section 4.4.

4.3.1 Assumptions

To derive the asymptotic results we assume that the data satisfy a set of sufficient conditions, denoted by SC: $\{X_i\}$ is a general linear process of the form $X_i = \sum_{h=-\infty}^{+\infty} b_h \, \varepsilon_{i-h}$, where (1) b_h's are constants, (2) $\{\varepsilon_i\}$ is a sequence of iid random variables with finite variance, σ_ε^2, and finite fourth cumulant $k_{\varepsilon,4}$, and (3) $\sum_{h=-\infty}^{+\infty} |h||b_h| <$

∞. Condition $\sum_{h=-\infty}^{+\infty} |h|\,|b_h| < \infty$ is equivalent to $b_h = O\left(|h|^{-2-\delta}\right)$, $\delta > 0$, which implies that the lag-h autocovariance

$$R(h) \;=\; {\sigma_\varepsilon}^2 \sum_{i=-\infty}^{+\infty} b_i\, b_{i+|h|}, \qquad [4.4]$$

is $O\left(|h|^{-2-\delta}\right)$.

4.3.2 Derivation of Intermediate Results

The first lemma yields an equation for the covariance between two OBM estimators, two Bartlett estimators, or one OBM and one Bartlett estimator of the variance of the sample mean.

LEMMA 4.1 Suppose that the observations $\{X_i; i = 1, \ldots, n\}$ satisfy conditions SC. Assume without loss of generality that the process mean is zero. Then

$$\operatorname{cov}\left(\widehat{V}^{(1)}, \widehat{V}^{(2)}\right) \;=\; 2\Phi_A \;+\; \Phi_B,$$

where $\widehat{V}^{(i)}$ is either an OBM or a Bartlett estimator with batch/window size m_i, for $i = 1, 2$, and

$$\Phi_A \;=\; R(0)^2 \sum_{r,s,t,u=1}^{n} p_{rs}^{(1)}\, p_{tu}^{(2)}\, \rho(t-r)\, \rho(u-s),$$

$$\Phi_B \;=\; \sum_{r,s,t,u=1}^{n} p_{rs}^{(1)}\, p_{tu}^{(2)}\, K_4(X_i, X_{i+r}, X_{i+s}, X_{i+t}).$$

Proof: By definition of covariance

$$\operatorname{cov}(\widehat{V}^{(1)}, \widehat{V}^{(2)}) \;=\; \mathrm{E}(\widehat{V}^{(1)}\widehat{V}^{(2)}) \;-\; \mathrm{E}(\widehat{V}^{(1)})\,\mathrm{E}(\widehat{V}^{(2)}).$$

Taking expectations and using Equations 2.10 and 2.18

$$\operatorname{cov}(\widehat{V}^{(1)}, \widehat{V}^{(2)}) \;=\; \sum_{r,s,t,u=1}^{n} p_{rs}^{(1)}\, p_{tu}^{(2)} \left[\mathrm{E}(X_r X_s X_t X_u) - \mathrm{E}(X_r X_s)\,\mathrm{E}(X_t X_u)\right].$$

By Equation 2.30

$$\begin{aligned}\operatorname{cov}(\widehat{V}^{(1)}, \widehat{V}^{(2)}) &= \sum_{r,s,t,u=1}^{n} p_{rs}^{(1)} p_{tu}^{(2)} [\mathrm{E}(X_r X_t)\,\mathrm{E}(X_s X_u) \\ &\qquad + \mathrm{E}(X_r X_u)\,\mathrm{E}(X_s X_t) \;+\; K_4(X_r, X_s, X_t, X_u)] \\ &= 2 \sum_{r,s,t,u=1}^{n} p_{rs}^{(1)} p_{tu}^{(2)} \mathrm{E}(X_r X_t)\,\mathrm{E}(X_s X_u) \\ &\qquad + \sum_{r,s,t,u=1}^{n} p_{rs}^{(1)} p_{tu}^{(2)} K_4(X_r, X_s, X_t, X_u).\end{aligned}$$

Now applying the definition of autocorrelation, the result follows.

Lemmas 4.2, 4.3, and 4.4 yield some asymptotic results that we use later in the proof of Theorem 4.1 and its corollaries.

LEMMA 4.2 Suppose that the observations $\{X_i; i = 1, \ldots, n\}$ satisfy conditions SC. Assume that $n \to \infty$, $m_1 \to \infty$, and $m_2 \to \infty$, while simultaneously $m_1/n \to 0$ and $m_1/m_2 \to c$, where c is a non-negative constant. Assume, without loss of generality, that $m_2 \geq m_1$. Then

$$\Phi_A \;=\; R(0)^2\, {\gamma_0}^2 \sum_{r,s=1}^{n} p_{rs}^{(1)} p_{rs}^{(2)} \;+\; O\left(\frac{m_2^2}{n^4}\right) \;+\; O\left(\frac{1}{n^3}\right).$$

Proof: The proof consists of showing that the second-order effects can be ignored. By Lemma 4.1 (definition of Φ_A) and replacing $r - t$ by α and $u - s$ by β

$$\Phi_A \;=\; R(0)^2 \sum_{r,s=1}^{n} \sum_{\alpha=r-n}^{r-1} \sum_{\beta=1-s}^{n-s} p_{rs}^{(1)}\; p_{r-\alpha,s+\beta}^{(2)}\, \rho(\alpha)\, \rho(\beta).$$

First we show that the effect of α and β through the quadratic-form coefficient, say Δ_1, is asymptotically negligible. Rewrite Φ_A to obtain

$$\Phi_A \;=\; R(0)^2 \sum_{r,s=1}^{n} \sum_{\alpha=r-n}^{r-1} \sum_{\beta=1-s}^{n-s} p_{rs}^{(1)}\; p_{rs}^{(2)}\, \rho(\alpha)\, \rho(\beta) \;+\; \Delta_1,$$

where, by Equations 2.13 and 2.19,

$$|\Delta_1| \;=\; O\left(\Big| \sum_{r,s=1}^{n} \sum_{\alpha=r-n}^{r-1} \sum_{\beta=1-s}^{n-s} p_{rs}^{(1)}\; \frac{\alpha+\beta}{n^2 m_2} \rho(\alpha)\, \rho(\beta) \Big| \right).$$

Assumption $\sum_{h=-\infty}^{+\infty} |h|\,|b_h| = O(1)$ implies that

$$|\Delta_1| \;=\; O\left(|\sum_{r,s=1}^{n} p_{rs}^{(1)} \frac{1}{m_2 n^2}|\right).$$

Again, by Equations 2.13 and 2.19,

$$\begin{aligned}|\Delta_1| &= O\left(\sum_{r=1}^{n}\sum_{s=r+1}^{r+m_1} \frac{1}{m_2 n^4} + \sum_{r=1}^{n}\sum_{r+m_1+1}^{n} \frac{1}{n^5}\right)\\ &= O\left(\frac{1}{n^3}\right).\end{aligned}$$

Next we show that the effect of r through the lower and upper bounds of α, say Δ_2, is also asymptotically negligible. Rewrite Φ_A to obtain

$$\Phi_A \;=\; R(0)^2 \sum_{r=1}^{n}\sum_{s=1}^{n} p_{rs}^{(1)}\, p_{rs}^{(2)} \sum_{\alpha=-m_2}^{m_2} \rho(\alpha) \sum_{\beta=1-s}^{n-s} \rho(\beta) + O\left(\frac{1}{n^3}\right) + \Delta_2.$$

By assumption $\sum_{\beta=1-s}^{n-s} \rho(\beta) = O(1)$, which implies that the error term Δ_2 is

$$\begin{aligned}|\Delta_2| &= O\left(\sum_{r=1}^{n}\sum_{s=1}^{n}\sum_{\alpha=m_2+1}^{r-1} p_{rs}^{(1)} p_{rs}^{(2)} \rho(\alpha) + \sum_{r=1}^{n}\sum_{s=1}^{n}\sum_{\alpha=r-n}^{-m2-1} p_{rs}^{(1)} p_{rs}^{(2)} \rho(\alpha)\right)\\ &\quad - O\left(\sum_{r=1}^{m_2}\sum_{s=1}^{n}\sum_{\alpha=r-1}^{m_2-1} p_{rs}^{(1)} p_{rs}^{(2)} \rho(\alpha) + \sum_{r=n-m_2+1}^{n}\sum_{s=1}^{n}\sum_{\alpha=-m_2-1}^{r-n} p_{rs}^{(1)} p_{rs}^{(2)} \rho(\alpha)\right).\end{aligned}$$

By assumption, $\rho(h) = O\left(|h|^{-2-\delta}\right)$. This implies that $\sum_{\alpha>m_2} \rho(\alpha) = O(1/m_2)$, $\sum_{\alpha<-m_2} \rho(\alpha) = O(1/m_2)$, $\sum_{\alpha=r-1}^{m_2-1} \rho(\alpha) = O(1)$, and $\sum_{\alpha=-m_2-1}^{r-n} \rho(\alpha) = O(1)$. Therefore,

$$|\Delta_2| \;=\; O\left(\frac{1}{m_2}\sum_{r=1}^{n}\sum_{s=1}^{n} p_{rs}^{(1)}\, p_{rs}^{(2)}\right) + O\left(\sum_{r=1}^{m_2}\sum_{s=1}^{n} p_{rs}^{(1)}\, p_{rs}^{(2)}\right).$$

By Equations 2.13 and 2.19, (see also Lemma 4.4) $\sum_{r=1}^{n}\sum_{s=1}^{n} p_{rs}^{(1)}\, p_{rs}^{(2)} = O(m_2/n^3)$. Then

$$|\Delta_2| \;=\; O\left(\frac{1}{n^3}\right) + O\left(\frac{m_2^2}{n^4}\right).$$

Similarly, the effect of s through the lower and upper bounds of β is

$$|\Delta_3| = O\left(\frac{1}{n^3}\right) + O\left(\frac{m_2^2}{n^4}\right).$$

Therefore,

$$\Phi_A = R(0)^2 \sum_{r=1}^{n}\sum_{s=1}^{n} p_{rs}^{(1)}\, p_{rs}^{(2)} \sum_{\alpha=-m_2}^{m_2} \rho(\alpha) \sum_{\beta=-m_2}^{m_2} \rho(\beta) + O\left(\frac{1}{n^3}\right) + O\left(\frac{m_2^2}{n^4}\right).$$

By Equation 2.4 (definition of γ_0) the result follows.

Now we present the corresponding lemma for Φ_B.

LEMMA 4.3 Suppose that the observations $\{X_i; i = 1, \ldots, n\}$ satisfy conditions SC. Assume that $n \to \infty$, $m_1 \to \infty$, and $m_2 \to \infty$, while simultaneously $m_1/n \to 0$ and $m_1/m_2 \to c$, where c is a non-negative constant. Assume, without loss of generality, that $m_2 \geq m_1$. Then

$$\Phi_B = O\left(\frac{1}{n^3}\right).$$

Proof: Again, the proof proceeds by showing that the error caused by ignoring the second-order effects is asymptotically zero. By Lemma 4.1 (definition of Φ_B) and Equation 2.31,

$$\Phi_B = \sum_{r,s,t,u=1}^{n} p_{rs}^{(1)}\, p_{tu}^{(2)}\, K_{\varepsilon,4} \sum_{v=\infty}^{\infty} b_v\, b_{v+s-r}\, b_{v+t-r}\, b_{v+u-r}.$$

By Equations 2.13 and 2.19, $p_{rs}^{(1)}\, p_{tu}^{(2)} = O(1/n^4)$. Let $f = v+s-r$, $g = v+t-r$ and $h = v+u-r$. Rewrite Φ_B to obtain

$$\Phi_B = O\left(\frac{1}{n^4}\sum_{r=1}^{n}\sum_{f=-\infty}^{\infty} b_f \sum_{g=-\infty}^{\infty} b_g \sum_{h=-\infty}^{\infty} b_h \sum_{v=-\infty}^{\infty} b_v\right).$$

By assumption $\sum_{h=-\infty}^{\infty} b_h = O(1)$. Therefore,

$$\Phi_B = O\left(\frac{1}{n^3}\right).$$

The next lemma refers to the estimation of the asymptotic $\sum_{r,s} p_{rs}^{(1)} p_{rs}^{(2)}$.

LEMMA 4.4 Assume that $n \to \infty$, $m_1 \to \infty$, and $m_2 \to \infty$, while simultaneously $m_1/n \to 0$ and $m_1/m_2 \to c$, where c is a non-negative constant. Assume, without loss of generality, that $m_2 \geq m_1$. Then

$$\sum_{r,s=1}^{n} p_{rs}^{(1)} p_{rs}^{(2)} \;=\; \frac{2m_1}{3n^3}\left(1 + \frac{m_2 - m_1}{2m_2}\right) + O\left(\frac{1}{m_2 n^3}\right) + O\left(\frac{m_2^2}{n^4}\right).$$

Proof: Again, the proof shows that the end-effects are asymptotically negligible. First we decompose the summation into two summations.

$$\sum_{r,s=1}^{n} p_{rs}^{(1)} p_{rs}^{(2)} \;=\; \sum_{r=1}^{n} p_{rr}^{(1)} p_{rr}^{(2)} + 2\sum_{r=1}^{n}\sum_{s=r+1}^{n} p_{rs}^{(1)} p_{rs}^{(2)}.$$

In the case of the Bartlett estimator (Equation 2.19) $p_{rr}^{(i)} = 1/n^2 + O(m_i/n^3)$. In the case of the OBM estimator (Equation 2.13) $p_{rr}^{(i)} = O\left(1/n^2\right)$ for $r \leq m_i$ or $r \geq n - m_i + 1$, and $p_{rr}^{(i)} = 1/n^2 + O(m_i/n^3)$ for $m_i + 1 \leq r \leq n - m_i$. Then,

$$\sum_{r,s=1}^{n} p_{rs}^{(1)} p_{rs}^{(2)} \;=\; \frac{1}{n^3} + 2\sum_{r=1}^{n}\sum_{s=r+1}^{n} p_{rs}^{(1)} p_{rs}^{(2)} + \Delta_1 ,$$

where Δ_1, the error caused by replacing $\sum_{r=1}^{n} p_{rr}^{(1)} p_{rr}^{(2)}$ by $1/n^3$, is at most

$$|\Delta_1| \;=\; O\left(\sum_{r=1}^{m_2}\frac{1}{n^4} + \sum_{r=m_2+1}^{n-m_2}\frac{m_2}{n^5}\right) \;=\; O\left(\frac{m_2}{n^4}\right).$$

Therefore,

$$\sum_{r,s=1}^{n} p_{rs}^{(1)} p_{rs}^{(2)} \;=\; \frac{1}{n^3} + 2\sum_{r=1}^{n}\sum_{s=r+1}^{n} p_{rs}^{(1)} p_{rs}^{(2)} + O\left(\frac{m_2}{n^4}\right).$$

Let s vary between $r+1$ and $\min(r+m_1, n)$ instead of $r+1$ and n. Notice that for $s > r + m_1$, $p_{rs}^{(1)}$ is $O(m_1/n^3)$. Then

$$\sum_{r,s=1}^{n} p_{rs}^{(1)} p_{rs}^{(2)} \;=\; \frac{1}{n^3} + 2\sum_{r=1}^{n}\sum_{s=r+1}^{\min(r+m_1,n)} p_{rs}^{(1)} p_{rs}^{(2)} + O\left(\frac{m_2}{n^4}\right) + \Delta_2,$$

where, by Equations 2.13 and 2.19,

$$|\Delta_2| \;=\; O\left(\sum_{r=1}^{n}\sum_{s=r+m_1+1}^{n}\frac{m_1}{n^3}\, p_{rs}^{(2)}\right)$$

$$= O\left(\frac{m_1}{n^3}\sum_{r=1}^{n}\sum_{s=r+m_1+1}^{r+m_2}\frac{1}{n^2} + \frac{m_1}{n^3}\sum_{r=1}^{n}\sum_{s=r+m_2+1}^{n}\frac{m_2}{n^3}\right)$$
$$= O\left(\frac{m_1\, m_2}{n^4}\right),$$

yielding, because $m_1 \leq m_2$,

$$\sum_{r,s=1}^{n} p_{rs}^{(1)} p_{rs}^{(2)} = \frac{1}{n^3} + 2\sum_{r=1}^{n}\sum_{s=r+1}^{\min(r+m_1,n)} p_{rs}^{(1)} p_{rs}^{(2)} + O\left(\frac{m_2^2}{n^4}\right).$$

Replacing $\min(r+m_1,n)$, the upper bound of s, by $r+m_1$

$$\sum_{r,s=1}^{n} p_{rs}^{(1)} p_{rs}^{(2)} = \frac{1}{n^3} + 2\sum_{r=1}^{n}\sum_{s=r+1}^{r+m_1} p_{rs}^{(1)} p_{rs}^{(2)} + O\left(\frac{m_2^2}{n^4}\right) + \Delta_3,$$

where, by Equations 2.13 and 2.19,

$$|\Delta_3| = O\left(\sum_{r=n-m_1+1}^{n}\sum_{s=n+1}^{r+m_1}\frac{1}{n^4}\right) = O\left(\frac{{m_1}^2}{n^4}\right).$$

The next step is to replace $p_{rs}^{(1)}$ and $p_{rs}^{(2)}$ by respectively $(1/n^2)[1-(s-r)/m_1]$ and $(1/n^2)[1-(s-r)/m_2]$. There is an error Δ_4 caused by these replacements. Then,

$$\sum_{r,s=1}^{n} p_{rs}^{(1)} p_{rs}^{(2)} = \frac{1}{n^3} + 2\sum_{r=1}^{n}\sum_{s=r+1}^{r+m_1}\frac{1}{n^4}\left(1-\frac{s-r}{m_1}\right)\left(1-\frac{s-r}{m_2}\right) + O\left(\frac{{m_1}^2}{n^4}\right) + \Delta_4,$$

where, by Equations 2.13 and 2.19,

$$|\Delta_4| = O\left(\sum_{r=1}^{n}\sum_{s=r+1}^{r+m_1}\left(\frac{1}{n^2}\right)\left(\frac{m_2}{n^3}\right)\right) = O\left(\frac{m_1 m_2}{n^4}\right).$$

Therefore, because $m_1 \leq m_2$,

$$\sum_{r,s=1}^{n} p_{rs}^{(1)} p_{rs}^{(2)} = \frac{1}{n^3} + 2\sum_{r=1}^{n}\sum_{s=r+1}^{r+m_1}\frac{1}{n^4}\left(1-\frac{s-r}{m_1}\right)\left(1-\frac{s-r}{m_2}\right) + O\left(\frac{{m_2}^2}{n^4}\right).$$

Let $h = s - r$. Then

$$\sum_{r,s=1}^{n} p_{rs}^{(1)} p_{rs}^{(2)} = \frac{1}{n^3} + \frac{2}{n^3}\sum_{h=1}^{m_1}\left(1-\frac{m_1+m_2}{m_1\, m_2}h + \frac{1}{m_1 m_2}h^2\right) + O\left(\frac{{m_2}^2}{n^4}\right)$$
$$= \frac{m_1}{n^3} - \frac{m_1^2}{3\, m_2 n^3} + \frac{1}{3\, m_2 n^3} + O\left(\frac{{m_2}^2}{n^4}\right).$$

The result follows after simple algebra.

4.4 Covariance among OBM and Bartlett Estimators

Theorem 4.1 states the asymptotic covariance between two OBM, two Bartlett or one OBM and one Bartlett estimator of the variance of the sample mean, $\widehat{V}^{(1)}$ and $\widehat{V}^{(2)}$, with two different truncation points m_1 and m_2 from n observations. Theorem 4.1 results directly from Lemmas 4.1, 4.2, 4.3 and 4.4.

THEOREM 4.1 Suppose that the observations $\{X_i; i = 1, \ldots, n\}$ satisfy conditions SC. Assume that $n \to \infty$, $m_1 \to \infty$, and $m_2 \to \infty$, while simultaneously $m_1/n \to 0$ and $m_1/m_2 \to c$, where c is a non-negative constant. Assume, without loss of generality, that $m_2 \geq m_1$. Then

$$\operatorname{cov}\left(\widehat{V}^{(1)}, \widehat{V}^{(2)}\right) = \frac{4\, m_1\, \gamma_0^2\, R(0)^2}{3\, n^3}\left(1 + \frac{m_2 - m_1}{2\, m_2}\right) + O\left(\frac{m_2^2}{n^4}\right) + O\left(\frac{1}{n^3}\right),$$

where $\widehat{V}^{(i)}$ is either an OBM or a Bartlett estimator with batch/window size m_i, for $i = 1, 2$.

Corollary 4.1 provides a formula to estimate the asymptotic variance of an OBM or a Bartlett estimator of V and is a direct consequence of Theorem 4.1 for $m_1 = m_2 = m$.

COROLLARY 4.1 Suppose that the observations $\{X_i; i = 1, \ldots, n\}$ satisfy conditions SC. Assume that $n \to \infty$, $m \to \infty$, while simultaneously $m/n \to 0$. Then

$$\operatorname{var}\left(\widehat{V}\right) = \frac{4\, m\, \gamma_0^2\, R(0)^2}{3\, n^3} + O\left(\frac{m^2}{n^4}\right) + O\left(\frac{1}{n^3}\right),$$

where $\widehat{V}$ is either an OBM or a Bartlett estimator of $\operatorname{var}(\bar{X})$ with batch/window size m.

This variance result was used and assumed as a conjecture in Song [1988].

Corollary 4.2 yields an equation for the asymptotic correlation of OBM and Bartlett estimators of the variance of the sample mean. It is obtained using the definition of correlation, Theorem 4.1 and Corollary 4.1.

COROLLARY 4.2 Suppose that the observations $\{X_i; i = 1, \ldots, n\}$ satisfy conditions SC. Assume that $n \to \infty$, $m_1 \to \infty$, and $m_2 \to \infty$, while simultaneously $m_1/n \to 0$ and $m_1/m_2 \to c$, where c is a non-negative constant. Assume, without loss of generality, that $m_2 \geq m_1$. Then

$$\operatorname{corr}\left(\widehat{V}^{(1)}, \widehat{V}^{(2)}\right) = \left(\frac{m_1}{m_2}\right)^{\frac{1}{2}} \left(1 + \frac{m_2 - m_1}{2\,m_2}\right) + O\left(\frac{m_2}{n}\right) + O\left(\frac{1}{m_1}\right),$$

where $\widehat{V}^{(i)}$ is either an OBM or a Bartlett estimator with batch/window size m_i, for $i = 1, 2$.

Using Equation 2.7, Corollary 4.1 can be rewritten as

$$\operatorname{var}\left(\widehat{V}\right) = \frac{4\,m}{3\,n} V^2 + O\left(\frac{m^2}{n^4}\right) + O\left(\frac{1}{n^3}\right). \qquad [4.5]$$

This result shows that, for a large fixed value of n, the variance of OBM and Bartlett estimators is directly proportional to m and the square of the variance of the sample mean.

Corollary 4.2 can also be rewritten as

$$\operatorname{cov}\left(\widehat{V}^{(1)}, \widehat{V}^{(2)}\right) = \frac{4\,m_1}{3\,n} V^2 \left(1 + \frac{m_2 - m_1}{2\,m_2}\right) + O\left(\frac{m_2^2}{n^4}\right) + O\left(\frac{1}{n^3}\right). \qquad [4.6]$$

Therefore, the asymptotic covariance of OBM and Bartlett estimators of V can be viewed as the product of the asymptotic variance of the estimator with smaller batch/window size, $\operatorname{var}(\widehat{V}^{(1)}) \simeq 4\,m_1\,V^2/(3\,n)$, by the correction factor $1 + (m_2 - m_1)/(2m_2)$. This indicates that given the data type the asymptotic covariance depends only on the batch/window sizes m_1 and m_2.

The asymptotic correlation of OBM and Bartlett estimators does not depend upon the data type. Rather it depends only upon the two relative batch/window sizes through $(m_1/m_2)^{1/2}$ and $1 + (m_2 - m_1)/(2m_2)$.

4.5 Finite-Sample Results

We now describe Monte Carlo experiments that estimate the correlation between two OBM estimators of V and compare these finite-sample results to the asymptotic results of the previous section. The purpose is to study the applicability of the asymptotic formulas for finite samples and different data types. Corollary 4.2 and Equation 4.6 indicate that the quality of the covariance approximation is similar to that provided by the correlation approximation. Therefore, we consider only correlation here.

4.5.1 The Monte Carlo Experiment

The Monte Carlo experiment estimates corr $\left(\widehat{V}^{(i)}, \widehat{V}^{(j)}\right)$ for six cases: two sample sizes and three steady-state data processes. The sample sizes are $n = 100$ and $n = 1000$; the three data processes are iid normal ($\gamma_0 = 1$, $\gamma_1 = 0$), AR(1) normal with $\gamma_0 = 10$ and $\gamma_1 = 49.5$, and 5-state DPSS [Pedrosa and Schmeiser, 1993a] with $\gamma_0 = 10$ and $\gamma_1 = -35.2$ (and uniform marginal distribution). For each sample $\{X_1, X_2, ..., X_n\}$ from the data process, OBM estimates $\widehat{V}^{(m)}$ are computed for $m = 1, 2, ..., 99$ if $n = 100$ and for $m = 1, 10, 20, ..., 990$ if $n = 1000$. From 10,000 such samples the correlations corr $\left(\widehat{V}^{(i)}, \widehat{V}^{(j)}\right)$ are estimated to negligible sampling error. We review the three data processes in Appendix.

The values of the process mean μ and the process variance $R(0)$ are irrelevant in this set of experiments, because (1) OBM estimators are invariant in respect to the process mean, and (2) the correlation is unscaled. For simplicity, the mean is always $\mu = 0$, and the marginal variance is always set so that the variance of the sample mean is $V = 1$.

4.5.2 Discussion of Experimental Results

The experimental results indicate that the quality of the approximation provided by the asymptotic correlation in Corollary 4.2 is good if both batch sizes are between γ_0 and $n/2$. The quality is relatively insensitive to the marginal distribution and to the weighted sum of correlations γ_1.

To aid the discussion, we introduce two figures. Figure 4.1, for normal data, and Figure 4.2, for DPSS data, illustrate the results for $n = 100$, the smallest sample size considered. Each figure contains six charts, each corresponding to a batch size $m_i = 5, 10, 30, 50, 60, 90$. The horizontal axis is the other batch size m_j, ranging from 1 to $n - 1$. For each chart, two curves are shown: the asymptotic correlation from Corollary 4.2 and the true finite-sample correlation; the true correlation goes to zero for large values of m_j. The closer are the two curves the better is the approximation. The approximation is exact at $m_i = m_j$, since both the approximation and the true correlation are then one.

A relatively good approximation is shown in Figure 4.1. The independent normal data have $\gamma_0 = 1$, so $n = 100$ is a relatively large sample size. Whenever both batch sizes are less than $n/2$, the approximation quality is good. The quality typically degenerates as either batch size increases beyond $n/2$. The $n = 1000$ graphs (not shown) are similar to Figure 4.1 for all three process. In both the AR(1) and DPSS cases, $\gamma_0 = 10$, so the equivalent number of independent observations $n/\gamma_0 = 100$.

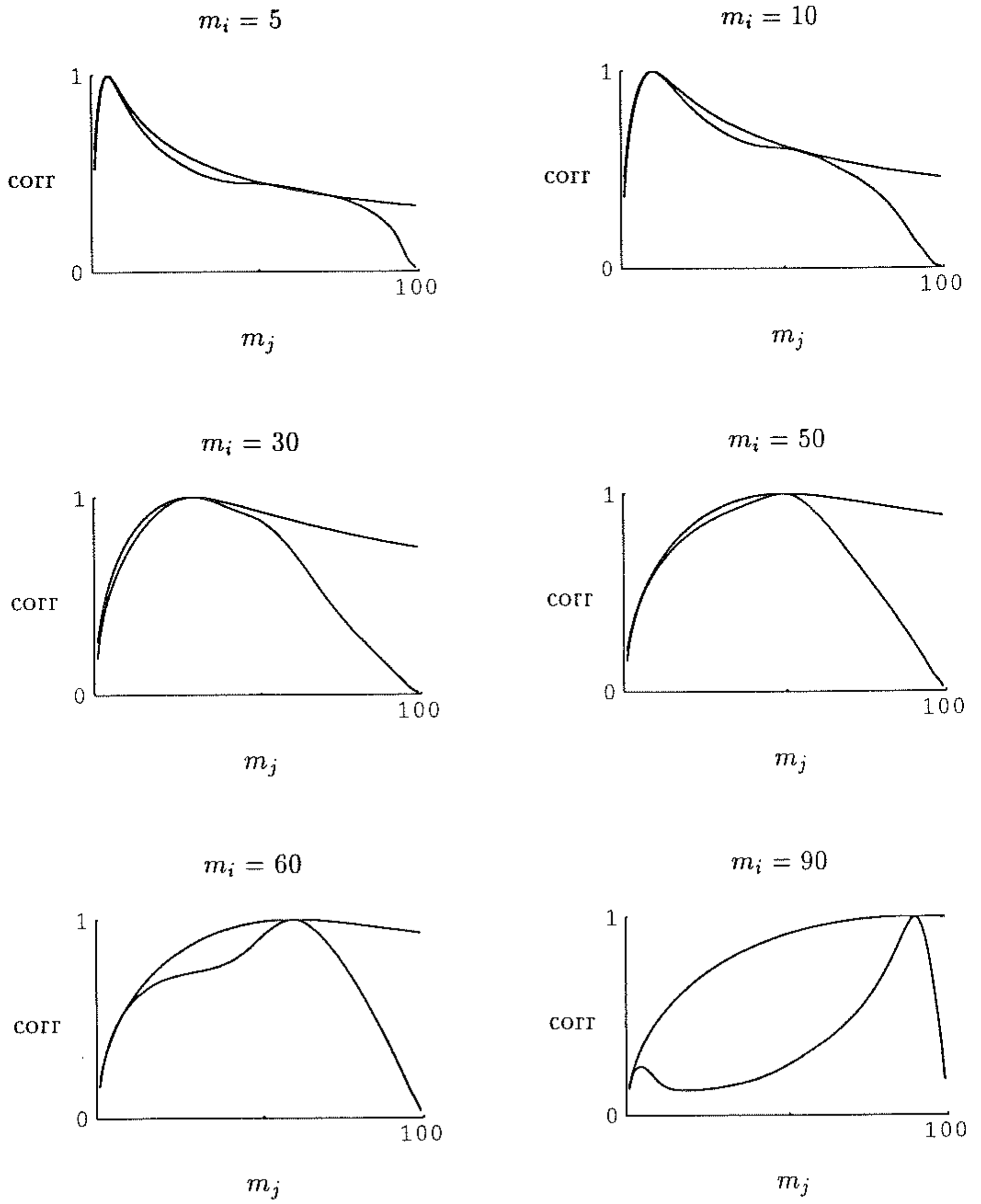

Figure 4.1 Correlation corr $\left(\widehat{V}^{(i)}, \widehat{V}^{(j)}\right)$ as a function of m_j for an iid-normal and a sample size $n = 100$.

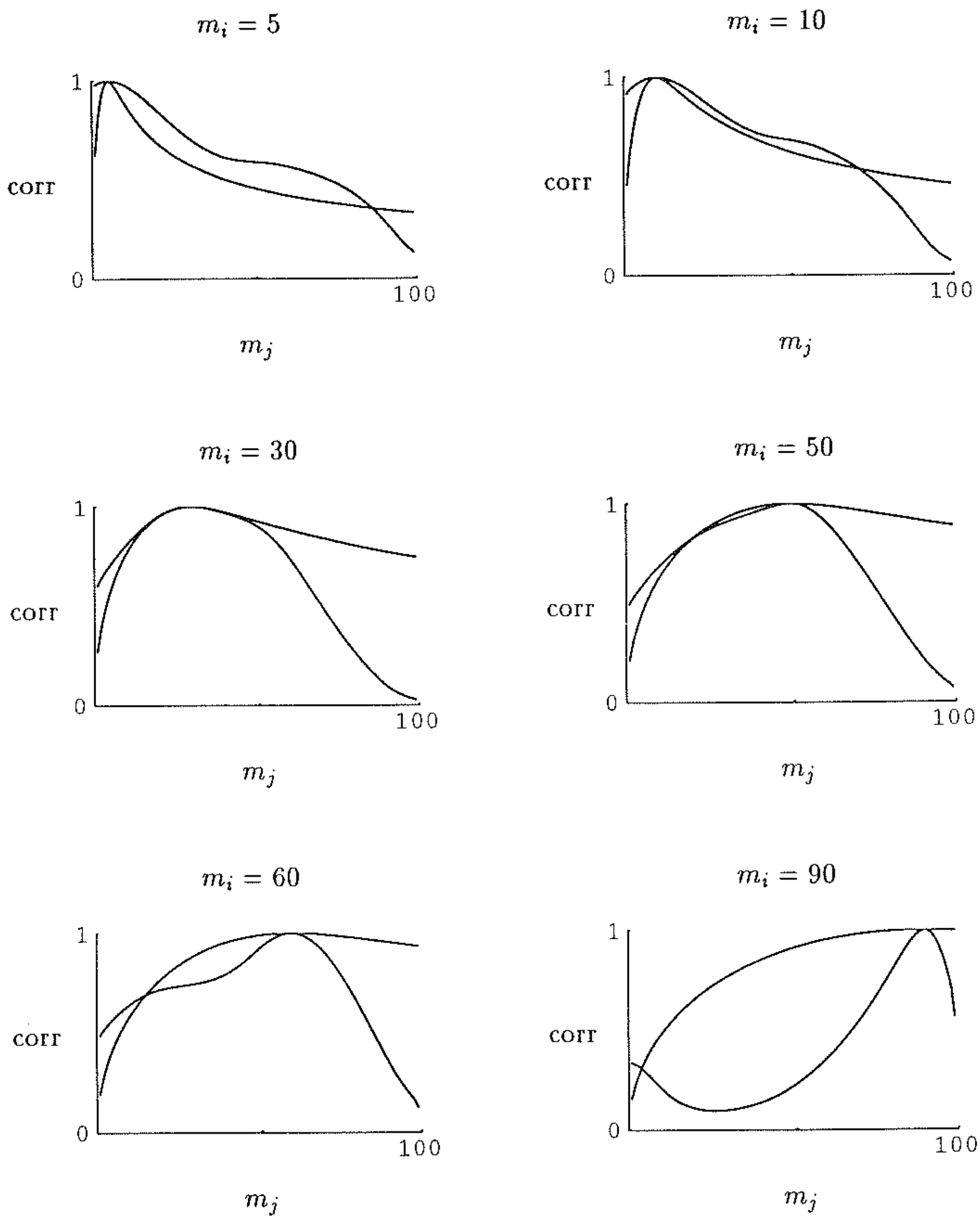

Figure 4.2 Correlation corr $\left(\widehat{V}^{(i)}, \widehat{V}^{(j)}\right)$ as a function of m_j for a DPSS($d = 5, p = 0.91, s = -2, S = 2$) and a sample size $n = 100$.

A less good approximation is shown in Figure 4.2. The DPSS dependent data have $\gamma_0 = 10$, so the equivalent number of independent observations is quite small, $n/\gamma_0 = 10$. Here the approximation quality degenerates when either m_j is too large or too small, as suggested by Corollary 4.2. Roughly, the quality of the approximation is good whenever both batch sizes are between γ_0 and $n/2$. Similar graphs result for the AR(1) process with $\gamma_0 = 10$ and $n = 100$.

For sample sizes n at least a few multiples of γ_0, these experimental results suggest these four qualitative conclusions:

1. The graphs in Figure 4.1 are representative of the asymptotic correlations, regardless of marginal distribution and autocorrelation structure.
2. The quality of the approximation is insensitive to the marginal distribution and to γ_1.
3. The equivalent sample size n/γ_0 is sufficient information to characterize the quality of the approximation.
4. The approximation is good if both batch sizes are between γ_0 and $n/2$.

This robustness to batch size is encouraging since various guidelines for choosing batch size fall within the range $(\gamma_0, n/2)$, unless the run length is quite small. For example, (1) Schmeiser [1982] suggests batch sizes between $n/30$ and $n/10$ for good confidence-interval performance and (2) the optimal batch size [Song and Schmeiser, 1994] is $O(\sqrt[3]{n})$.

5. COVARIANCE BETWEEN BARTLETT ESTIMATORS OF THE SPECTRAL DENSITY

5.1 Introduction

Song and Schmeiser [1988a], for frequency $\omega = 0$, and Politis and Romano [1994], for general frequencies ω, have suggested using linear combinations of classical estimators to estimate the spectral density. But using linear combinations requires choosing p estimators and p corresponding weights. A special case is to choose p bandwidths for Bartlett estimators, a family of estimators that is both classical and relatively simple. Both for choosing bandwidths and for determining appropriate linear-combination weights, we are interested in the asymptotic covariances among Bartlett estimators.

The results of this chapter generalize to non-zero frequencies the results of Chapter 4, which considers zero frequency (i.e., the variance of the sample mean) for the Bartlett estimator and the asymptotically equivalent overlapping-batch-means estimator [Pedrosa and Schmeiser, 1993c]. Chapter 4 also empirically suggests that the zero-frequency asymptotic covariances provide a good approximation to finite-sample covariances.

In Section 5.2 we prove intermediate results, in the form of six lemmas. In Section 5.3 we derive our main result: the asymptotic covariance between Bartlett estimators of the spectral density function at frequency ω.

5.2 Intermediate Results

Here we develop the theoretical structure needed to derive the asymptotic results of Section 5.3.

5.2.1 Assumptions and Notation

To derive the asymptotic results we assume that the data satisfy a set of sufficient conditions, denoted by SC: $\{X_i\}$ is a general linear process of the form $X_i = \sum_{h=-\infty}^{+\infty} b_h\, \varepsilon_{i-h}$, where (1) b_h's are constants, (2) $\{\varepsilon_i\}$ is a sequence of iid random variables with finite variance, σ_ε^2, and finite fourth cumulant $k_{\varepsilon,4}$, and (3) $\sum_{h=-\infty}^{+\infty} |h||b_h| < \infty$. Condition $\sum_{h=-\infty}^{+\infty} |h|\,|b_h| < \infty$ is equivalent to $b_h = O\left(|h|^{-2-\delta}\right)$, $\delta > 0$, which implies that the lag-h autocovariance (Equation 2.23) is $R(h) = O\left(|h|^{-2-\delta}\right)$.

We denote the lag-h autocorrelation by $\rho(h)$ and the quadratic-form coefficients of $\widehat{f}^{(1)}(\omega)$ and $\widehat{f}^{(2)}(\omega)$ by, respectively, $p_{rs}^{(1)}$ and $p_{rs}^{(2)}$. These quadratic-form coefficients are defined by Equation 2.19.

5.2.2 Derivation of Intermediate Results

The first lemma yields an equation for the covariance between two Bartlett estimators of the spectral density.

LEMMA 5.1 Suppose that the observations $\{X_i; i = 1, \ldots, n\}$ satisfy conditions SC. Assume without loss of generality that the process mean is zero. Then the covariance between $\widehat{f}^{(1)}(\omega)$ and $\widehat{f}^{(2)}(\omega)$ is

$$\operatorname{cov}\left(\widehat{f}^{(1)}(w), \widehat{f}^{(2)}(w)\right) = \left(\frac{n\,R(0)}{2\pi}\right)^2 (\Phi_A + \Phi_B + \Phi_C),$$

where

$$\Phi_A = \sum_{r,s,t,u=1}^{n} p_{rs}^{(1)}\, p_{tu}^{(2)}\, e^{-j(r-s)\omega}\, e^{-j(t-u)\omega}\, \rho(t-r)\,\rho(u-s),$$

$$\Phi_B = \sum_{r,s,t,u=1}^{n} p_{rs}^{(1)}\, p_{tu}^{(2)}\, e^{-j(r-s)\omega}\, e^{-j(t-u)\omega}\, \rho(u-r)\, \rho(t-s),$$

$$\Phi_C = \frac{1}{R(0)^2} \sum_{r,s,t,u=1}^{n} p_{rs}^{(1)}\, p_{tu}^{(2)}\, e^{-j(r-s)\omega}\, e^{-j(t-u)\omega}\, K_4(X_i, X_{i+r}, X_{i+s}, X_{i+t}).$$

Proof: By definition of covariance

$$\operatorname{cov}(\hat{f}^{(1)}, \hat{f}^{(2)}) = \mathrm{E}(\hat{f}^{(1)} \hat{f}^{(2)}) - \mathrm{E}(\hat{f}^{(1)})\, \mathrm{E}(\hat{f}^{(2)}).$$

By Equation 2.27

$$\begin{aligned}\operatorname{cov}(\hat{f}^{(1)}, \hat{f}^{(2)}) = \left(\frac{n}{2\pi}\right)^2 \sum_{r,s,t,u=1}^{n} p_{rs}^{(1)}\, p_{tu}^{(2)}\, e^{-j(r-s)\omega}\, e^{-j(t-u)\omega}\, &[\mathrm{E}(X_r X_s X_t X_u) \\ &- \mathrm{E}(X_r X_s)\, \mathrm{E}(X_t X_u)]\,.\end{aligned}$$

By Equation 2.30

$$\begin{aligned}\operatorname{cov}(\hat{f}^{(1)}, \hat{f}^{(2)}) = \left(\frac{n}{2\pi}\right)^2 \sum_{r,s,t,u=1}^{n} p_{rs}^{(1)}\, p_{tu}^{(2)}\, e^{-j(r-s)\omega}\, e^{-j(t-u)\omega}\, &[\mathrm{E}(X_r X_t)\, \mathrm{E}(X_s X_u) \\ &+ \mathrm{E}(X_r X_u)\, \mathrm{E}(X_s X_t) + K_4(X_r, X_s, X_t, X_u)]\,.\end{aligned}$$

Now applying the definition of autocorrelation, the result follows.

Lemmas 5.2 through 5.6 yield some asymptotic results that we use in the proof of Theorem 5.1 and its corollaries.

LEMMA 5.2 Suppose that the observations $\{X_i; i = 1, \ldots, n\}$ satisfy conditions SC. Assume that $n \to \infty$, $m_1 \to \infty$, and $m_2 \to \infty$, while simultaneously $m_1/n \to 0$ and $m_1/m_2 \to c$, where c is a non-negative constant. Assume, without loss of generality, that $m_2 \geq m_1$. Then

$$\Phi_B = \frac{(2\pi f(\omega))^2}{R(0)^2} \sum_{r,s=1}^{n} p_{rs}^{(1)}\, p_{rs}^{(2)} + O\left(\frac{1}{n^3}\right) + O\left(\frac{m_2^2}{n^4}\right).$$

Proof: The proof consists of showing that the second-order effects can be ignored. Interchanging t and u in the definition of Φ_B in Lemma 5.1 yields

$$\Phi_B = \sum_{r,s,t,u=1}^{n} p_{rs}^{(1)} p_{tu}^{(2)} e^{-j(r-t)\omega} e^{-j(u-s)\omega} \rho(t-r)\rho(u-s).$$

Replacing $r-t$ by α and $u-s$ by β

$$\Phi_B = \sum_{r=1}^{n}\sum_{s=1}^{n}\sum_{\alpha=r-n}^{r-1}\sum_{\beta=1-s}^{n-s} p_{rs}^{(1)} p_{r-\alpha,s+\beta}^{(2)} e^{-j(\alpha+\beta)\omega} \rho(\alpha)\rho(\beta).$$

First we show that the effect of α and β through the quadratic-form coefficient, say Δ_1, is asymptotically negligible. Rewrite Φ_B to obtain

$$\Phi_B = \sum_{r=1}^{n}\sum_{s=1}^{n}\sum_{\alpha=r-n}^{r-1}\sum_{\beta=1-s}^{n-s} p_{rs}^{(1)} p_{rs}^{(2)} e^{-j(\alpha+\beta)\omega} \rho(\alpha)\rho(\beta) + \Delta_1,$$

where by Equation 2.19

$$|\Delta_1| = O\left(\left|\sum_{r,s=1}^{n}\sum_{\alpha=r-n}^{r-1}\sum_{\beta=1-s}^{n-s} p_{rs}^{(1)} \frac{\alpha+\beta}{m_2 n^2}\rho(\alpha)\rho(\beta)\right|\right).$$

Assumption $\sum_{h=-\infty}^{+\infty} |h|\,|b_h| = O(1)$ implies that

$$|\Delta_1| = O\left(\left|\sum_{r,s=1}^{n} p_{rs}^{(1)} \frac{1}{m_2 n^2}\right|\right).$$

Again, by Equation 2.19

$$|\Delta_1| = O\left(\sum_{r=1}^{n}\sum_{s=r+1}^{r+m_1} \frac{1}{m_2 n^4}\right) = O\left(\frac{1}{n^3}\right).$$

Therefore,

$$\Phi_B = \sum_{r=1}^{n}\sum_{s=1}^{n}\sum_{\alpha=r-n}^{r-1}\sum_{\beta=1-s}^{n-s} p_{rs}^{(1)} p_{rs}^{(2)} e^{-j(\alpha+\beta)\omega} \rho(\alpha)\rho(\beta) + O\left(\frac{1}{n^3}\right).$$

Next we show that the effect of r through the lower and upper bounds of α, say Δ_2, is also asymptotically negligible. Rewrite Φ_B to obtain

$$\Phi_B = \sum_{r=1}^{n}\sum_{s=1}^{n} p_{rs}^{(1)} p_{rs}^{(2)} \sum_{\alpha=-m_2}^{m_2} \rho(\alpha) e^{-j\alpha\omega} \sum_{\beta=1-s}^{n-s} \rho(\beta) e^{-j\beta\omega} + O\left(\frac{1}{n^3}\right) + \Delta_2.$$

By assumption $\sum_{\beta=1-s}^{n-s} \rho(\beta) = O(1)$, which implies that the error term Δ_2 is

$$|\Delta_2| = O\left(\sum_{r=1}^{n}\sum_{s=1}^{n}\sum_{\alpha=m_2+1}^{r-1} p_{rs}^{(1)} p_{rs}^{(2)} \rho(\alpha) + \sum_{r=1}^{n}\sum_{s=1}^{n}\sum_{\alpha=r-n}^{-m2-1} p_{rs}^{(1)} p_{rs}^{(2)} \rho(\alpha)\right)$$

$$- O\left(\sum_{r=1}^{m_2}\sum_{s=1}^{n}\sum_{\alpha=r-1}^{m_2-1} p_{rs}^{(1)} p_{rs}^{(2)} \rho(\alpha) + \sum_{r=n-m_2+1}^{n}\sum_{s=1}^{n}\sum_{\alpha=-m_2-1}^{r-n} p_{rs}^{(1)} p_{rs}^{(2)} \rho(\alpha)\right).$$

By assumption, $\rho(h) = O\left(|h|^{-2-\delta}\right)$. This implies that (i) $\sum_{\alpha>m_2} \rho(\alpha) = O(1/m_2)$, (ii) $\sum_{\alpha<-m_2} \rho(\alpha) = O(1/m_2)$, (iii) $\sum_{\alpha=r-1}^{m_2-1} \rho(\alpha) = O(1)$, and (iv) $\sum_{\alpha=-m_2-1}^{r-n} \rho(\alpha) = O(1)$. Therefore,

$$|\Delta_2| = O\left(\frac{1}{m_2}\sum_{r=1}^{n}\sum_{s=1}^{n} p_{rs}^{(1)} p_{rs}^{(2)}\right) + O\left(\sum_{r=1}^{m_2}\sum_{s=1}^{n} p_{rs}^{(1)} p_{rs}^{(2)}\right).$$

By Equation 2.19, (see also Lemma 5.5) $\sum_{r=1}^{n}\sum_{s=1}^{n} p_{rs}^{(1)} p_{rs}^{(2)} = O(m_1/n^3)$.Then

$$|\Delta_2| = O\left(\frac{1}{n^3}\right) + O\left(\frac{m_2^2}{n^4}\right).$$

Similarly, the effect of s through the lower and upper bounds of β is

$$|\Delta_3| = O\left(\frac{1}{n^3}\right) + O\left(\frac{m_2^2}{n^4}\right).$$

Therefore,

$$\Phi_B = \sum_{r,s=1}^{n} p_{rs}^{(1)} p_{rs}^{(2)} \left(\sum_{\alpha=-m_2}^{m_2} \rho(\alpha)\, e^{-j\alpha\omega}\right)^2 + O\left(\frac{1}{n^3}\right) + O\left(\frac{m_2^2}{n^4}\right).$$

Again, by Equation 2.22, and using the relationship between autocorrelation and autocovariance, yields the result.

Similarly, there is an equivalent lemma for Φ_A.

LEMMA 5.3 Suppose that the observations $\{X_i; i = 1, \ldots, n\}$ satisfy conditions SC. Assume that $n \to \infty$, $m_1 \to \infty$, and $m_2 \to \infty$, while simultaneously $m_1/n \to 0$ and $m_1/m_2 \to c$, where c is a non-negative constant. Assume, without loss of generality, that $m_2 \geq m_1$. Then

$$\Phi_A = \frac{(2\pi f(\omega))^2}{R(0)^2} \sum_{r,s=1}^{n} p_{rs}^{(1)} p_{rs}^{(2)} e^{-j2(r-s)\omega} + O\left(\frac{1}{n^3}\right) + O\left(\frac{m_2^2}{n^4}\right).$$

Proof: By Lemma 5.1 (definition of Φ_A)

$$\Phi_A = \sum_{r,s,t,u=1}^{n} p_{rs}^{(1)} p_{tu}^{(2)} e^{-j(r-s)\omega} e^{-j(t-u)\omega} \rho(t-r)\rho(u-s).$$

Let $\alpha = t - r$ and $\beta = s - u$. Then $t - u = r - s + \alpha + \beta$ and

$$\Phi_A = \sum_{r=1}^{n}\sum_{s=1}^{n}\sum_{\alpha=1-r}^{n-r}\sum_{\beta=s-n}^{s-1} p_{rs}^{(1)} p_{r+\alpha,s-\beta}^{(2)} e^{-j2(r-s)\omega} e^{-j(\alpha+\beta)\omega} \rho(\alpha)\rho(\beta).$$

Since $|e^{-j2(r-s)\omega}| = 1$, we can use the same arguments of the previous lemma to obtain

$$\Phi_A = \sum_{r,s=1}^{n} p_{rs}^{(1)} p_{r,s}^{(2)} e^{-j2(r-s)\omega} \left(\sum_{\alpha=-m_2}^{m_2} \rho(\alpha) e^{-j\alpha\omega}\right)^2 + O\left(\frac{1}{n^3}\right) + O\left(\frac{m_2^2}{n^4}\right).$$

By Equation 2.22, and using the relationship between autocorrelation and autocovariance, yields the result.

Now we present the corresponding lemma for Φ_C.

LEMMA 5.4 Suppose that the observations $\{X_i; i = 1, \ldots, n\}$ satisfy conditions SC. Assume that $n \to \infty$, $m_1 \to \infty$, and $m_2 \to \infty$, while simultaneously $m_1/n \to 0$ and $m_1/m_2 \to c$, where c is a non-negative constant. Assume, without loss of generality, that $m_2 \geq m_1$. Then

$$\Phi_C = O\left(\frac{1}{n^3}\right).$$

Proof: Again, the proof proceeds by showing that the error caused by ignoring the second-order effects is asymptotically zero. By Lemma 5.1 (definition of Φ_C) and Equation 2.31

$$\Phi_C = \frac{1}{R(0)^2} \sum_{r,s,t,u=1}^{n} p_{rs}^{(1)} p_{tu}^{(2)} e^{-j(r-s)\omega} e^{-j(t-u)\omega} K_{\varepsilon,4} \sum_{v=-\infty}^{\infty} b_v\, b_{v+s-r}\, b_{v+t-r}\, b_{v+u-r}.$$

By Equation 2.19, $p_{rs}^{(1)} p_{tu}^{(2)} = O(1/n^4)$. Let $f = v+s-r$, $g = v+t-r$ and $h = v+u-r$. Rewrite Φ_C to obtain

$$\Phi_C = O\left(\frac{1}{n^4}\sum_{r=1}^{n}\sum_{f=-\infty}^{\infty} b_f \sum_{g=-\infty}^{\infty} b_g \sum_{h=-\infty}^{\infty} b_h \sum_{v=-\infty}^{\infty} b_v\right).$$

By assumption $\sum_{h=-\infty}^{\infty} b_h = O(1)$. Therefore,

$$\Phi_B \;=\; O\left(\frac{1}{n^3}\right).$$

The next lemma refers to the estimation of the asymptotic $\sum_{r,s} p_{rs}^{(1)} p_{rs}^{(2)}$.

LEMMA 5.5 Assume that $n \to \infty$, $m_1 \to \infty$, and $m_2 \to \infty$, while simultaneously $m_1/n \to 0$ and $m_1/m_2 \to c$, where c is a non-negative constant. Assume, without loss of generality, that $m_2 \geq m_1$. Then

$$\sum_{r,s=1}^{n} p_{rs}^{(1)} p_{rs}^{(2)} \;=\; \frac{2\,m_1}{3\,n^3}\left(1 + \frac{m_2 - m_1}{2\,m_2}\right) + O\left(\frac{1}{m_2\,n^3}\right) + O\left(\frac{m_2^2}{n^4}\right).$$

Proof: Again, the proof shows that the end-effects are asymptotically negligible. First we decompose the summation into two summations.

$$\sum_{r,s=1}^{n} p_{rs}^{(1)} p_{rs}^{(2)} \;=\; \sum_{r=1}^{n} p_{rr}^{(1)} p_{rr}^{(2)} + 2\sum_{r=1}^{n}\sum_{s=r+1}^{n} p_{rs}^{(1)} p_{rs}^{(2)}.$$

By Equation 2.19 $p_{rr}^{(i)} = 1/n^2 + O(m_i/n^3)$. Therefore,

$$\sum_{r,s=1}^{n} p_{rs}^{(1)} p_{rs}^{(2)} \;=\; \frac{1}{n^3} + 2\sum_{r=1}^{n}\sum_{s=r+1}^{n} p_{rs}^{(1)} p_{rs}^{(2)} + O\left(\frac{m_2}{n^4}\right).$$

Let s vary between $r+1$ and $\min(r+m_1, n)$ instead of $r+1$ and n. Notice that for $s > r + m_1$, $p_{rs}^{(1)}$ is at most $O(m_1/n^3)$. Then

$$\sum_{r,s=1}^{n} p_{rs}^{(1)} p_{rs}^{(2)} \;=\; \frac{1}{n^3} + 2\sum_{r=1}^{n}\sum_{s=r+1}^{\min(r+m_1,n)} p_{rs}^{(1)} p_{rs}^{(2)} + O\left(\frac{m_2}{n^4}\right) + \Delta_1,$$

where, by Equation 2.19,

$$\begin{aligned}
|\Delta_1| &= O\left(\sum_{r=1}^{n}\sum_{s=r+m_1+1}^{n} \frac{m_1}{n^3}\, p_{rs}^{(2)}\right) \\
&= O\left(\frac{m_1}{n^3}\sum_{r=1}^{n}\sum_{s=r+m_1+1}^{r+m_2} \frac{1}{n^2} + \frac{m_1}{n^3}\sum_{r=1}^{n}\sum_{s=r+m_2+1}^{n} \frac{m_2}{n^3}\right) \\
&= O\left(\frac{m_1\,m_2}{n^4}\right),
\end{aligned}$$

yielding, because $m1 \leq m2$,

$$\sum_{r,s=1}^{n} p_{rs}^{(1)} p_{rs}^{(2)} = \frac{1}{n^3} + 2 \sum_{r=1}^{n} \sum_{s=r+1}^{\min(r+m_1,n)} p_{rs}^{(1)} p_{rs}^{(2)} + O\left(\frac{m_2^2}{n^4}\right).$$

Replacing $\min(r + m_1, n)$, the upper bound of s, by $r + m_1$,

$$\sum_{r,s=1}^{n} p_{rs}^{(1)} p_{rs}^{(2)} = \frac{1}{n^3} + 2\sum_{r=1}^{n} \sum_{s=r+1}^{r+m_1} p_{rs}^{(1)} p_{rs}^{(2)} + O\left(\frac{m_2^2}{n^4}\right) + \Delta_2,$$

where by Equation 2.19

$$|\Delta_2| = O\left(\sum_{r=n-m_1+1}^{n} \sum_{s=n+1}^{r+m_1} \frac{1}{n^4}\right) = O\left(\frac{{m_1}^2}{n^4}\right).$$

The next step is to replace $p_{rs}^{(1)}$ and $p_{rs}^{(2)}$ by respectively $(1/n^2)[1 - (s-r)/m_1]$ and $(1/n^2)[1 - (s - r)/m_2]$. There is an error Δ_3 caused by these replacements. Then

$$\sum_{r,s=1}^{n} p_{rs}^{(1)} p_{rs}^{(2)} = \frac{1}{n^3} + 2 \sum_{r=1}^{n} \sum_{s=r+1}^{r+m_1} \frac{1}{n^4} \left(1 - \frac{s-r}{m_1}\right) \left(1 - \frac{s-r}{m_2}\right) + O\left(\frac{m_2^2}{n^4}\right) + \Delta_3,$$

where, by Equation 2.19,

$$|\Delta_3| = O\left(\sum_{r=1}^{n} \sum_{s=r+1}^{r+m_1} \left(\frac{1}{n^2}\right) \left(\frac{m_2}{n^3}\right)\right) = O\left(\frac{m_1 m_2}{n^4}\right).$$

Therefore, because $m_1 \leq m_2$,

$$\sum_{r,s=1}^{n} p_{rs}^{(1)} p_{rs}^{(2)} = \frac{1}{n^3} + 2 \sum_{r=1}^{n} \sum_{s=r+1}^{r+m_1} \frac{1}{n^4} \left(1 - \frac{s-r}{m_1}\right) \left(1 - \frac{s-r}{m_2}\right) + O\left(\frac{{m_2}^2}{n^4}\right).$$

Let $h = s - r$. Then

$$\begin{aligned}\sum_{r,s=1}^{n} p_{rs}^{(1)} p_{rs}^{(2)} &= \frac{1}{n^3} + \frac{2}{n^3} \sum_{h=1}^{m_1} \left(1 - \frac{m_1 + m_2}{m_1 m_2} h + \frac{1}{m_1 m_2} h^2\right) + O\left(\frac{{m_2}^2}{n^4}\right) \\ &= \frac{m_1}{n^3} - \frac{m_1^2}{3 m_2 n^3} + \frac{1}{3 m_2 n^3} + O\left(\frac{{m_2}^2}{n^4}\right).\end{aligned}$$

The result follows after simple algebra.

Similarly, $\sum_{r,s} p_{rs}^{(1)} p_{rs}^{(2)} e^{-j2(r-s)\omega}$ for non-zero frequency is given in Lemma 5.6.

LEMMA 5.6 Assume that $n \to \infty$, $m_1 \to \infty$, and $m_2 \to \infty$, while simultaneously $m_1/n \to 0$ and $m_1/m_2 \to c$, where c is a non-negative constant. Assume, without loss of generality, that $m_2 \geq m_1$. Then for any ω such that $0 < |\omega| < \pi$

$$\sum_{r,s=1}^{n} p_{rs}^{(1)}\, p_{rs}^{(2)}\, e^{-j2(r-s)\omega} \;=\; O\left(\frac{1}{m_1\, n^3}\right) \;+\; O\left(\frac{m_2^2}{n^4}\right).$$

Proof: First we decompose the summation into two summations.

$$\sum_{r,s=1}^{n} p_{rs}^{(1)}\, p_{rs}^{(2)}\, e^{-j2(r-s)\omega} \;=\; 2\sum_{r=1}^{n}\sum_{s=r+1}^{n} p_{rs}^{(1)}\, p_{rs}^{(2)}\, \cos\left[2(r-s)\omega\right] \;+\; \sum_{r=1}^{n} p_{rr}^{(1)}\, p_{rr}^{(2)}.$$

We can use the same arguments of Lemma 5.5 to obtain

$$\sum_{r,s=1}^{n} p_{rs}^{(1)}\, p_{rs}^{(2)}\, e^{-j2(r-s)\omega} \;=\; \frac{1}{n^3} + 2\sum_{r=1}^{n}\sum_{s=r+1}^{r+m_1} p_{rs}^{(1)}\, p_{rs}^{(2)}\, \cos\left[2(r-s)\omega\right] \;+\; O\left(\frac{m_2^2}{n^4}\right).$$

By Equation 2.19 and letting $h = s - r$

$$\sum_{r,s=1}^{n} p_{rs}^{(1)} p_{rs}^{(2)} e^{-j2(r-s)\omega} = \frac{1}{n^3} + \frac{2}{n^3}\sum_{h=1}^{m_1}\left(1 - \frac{m_1+m_2}{m_1\, m_2}h + \frac{1}{m_1\, m_2}h^2\right)\cos(2\omega h) + O\left(\frac{m_2^2}{n^4}\right).$$

Notice that for $0 < |\omega| < \pi$

$$\sum_{h=1}^{m_1} \sin(2\omega h) \;=\; \frac{\sin\left[(m_1+1)\omega\right]\, \sin(m_1\omega)}{\sin\,\omega},$$

and

$$\sum_{h=1}^{m_1} \cos(2\omega h) \;=\; \frac{1}{2}\left(-1 + \frac{\sin\left[(2m_1+1)\omega\right]}{\sin\,\omega}\right).$$

On the other hand,

$$\sum_{h=1}^{m_1} h\, \cos(2\omega h) \;=\; \frac{1}{2}\frac{d}{d\omega}\left(\sum_{h=1}^{m_1} \sin(2\omega h)\right),$$

and

$$\sum_{h=1}^{m_1} h^2\, \cos(2\omega h) \;=\; -\frac{1}{4}\frac{d^2}{d\omega^2}\left(\sum_{h=1}^{m_1} \cos(2\omega h)\right).$$

Therefore

$$\sum_{r,s=1}^{n} p_{rs}^{(1)}\, p_{rs}^{(2)}\, e^{-j2(r-s)\omega} \;=\; \frac{1}{n^3}\left[\frac{\sin\left[(2\,m_1\,+\,1)\,\omega\right]}{\sin\omega} - \frac{m_1+m_2}{m_1\,m_2}\right.$$

$$\left.\frac{d}{d\omega}\left(\frac{\sin\left[(m_1\,+\,1)\,\omega\right]\,\sin(m_1\,\omega)}{\sin\omega}\right) - \frac{1}{4\,m_1\,m_2}\,\frac{d^2}{d\omega^2}\left(\frac{\sin\left[(2\,m_1\,+\,1)\,\omega\right]}{\sin\omega}\right)\right].$$

These derivatives are given by

$$\frac{d}{d\omega}\left(\frac{\sin\left[(m_1+1)\omega\right]\,\sin(m_1\omega)}{\sin\omega}\right) \;=\; m_1\,\frac{\sin\left[(2m_1+1)\omega\right]}{\sin\omega} - \left(\frac{\sin(m_1\omega)}{\sin\omega}\right)^2,$$

and

$$\frac{d^2}{d\omega^2}\left(\frac{\sin\left[(2m_1+1)\omega\right]}{\sin\omega}\right) = -4{m_1}^2\,\frac{\sin\left[(2m_1+1)\omega\right]}{\sin\omega} + 4(2m_1+1)\left(\frac{\sin(m_1\omega)}{\sin\omega}\right)^2$$

$$+\frac{2}{(\sin\omega)^2}\left(\frac{\sin\left[(2m_1+1)\omega\right]}{\sin\omega} - (2m_1+1)\right).$$

After some algebraic manipulation

$$\sum_{r,s=1}^{n} p_{rs}^{(1)}\, p_{rs}^{(2)}\, e^{-j2(r-s)\omega} \;=\; \frac{1}{n^3}\left[\left(\frac{m_2-m_1-1}{m_1\,m_2}\right)\left(\frac{\sin(m_1\omega)}{\sin w}\right)^2 - \left(\frac{1}{2\,m_1\,m_2\,(\sin\omega)^2}\right)\right.$$

$$\left.\left(\frac{\sin\left[(2m_1+1)\omega\right]}{\sin\omega} - (2m_1+1)\right)\right] + O\left(\frac{m_2^2}{n^4}\right)$$

$$= O\left(\frac{1}{m_1\,n^3}\right) + O\left(\frac{m_2^2}{n^4}\right).$$

5.3 Covariance between Bartlett Estimators of $f(\omega)$

In this section we first develop the asymptotic covariance between two Bartlett estimators of the spectral density, $\widehat{f}^{(1)}(\omega)$ and $\widehat{f}^{(2)}(\omega)$, with two different truncation points m_1 and m_2, from n common observations. We then discuss two corollaries.

Theorem 5.1 follows directly from Lemmas 5.1 through 5.6.

THEOREM 5.1 Suppose that the observations $\{X_i; i = 1,\ldots,n\}$ satisfy conditions SC. Assume that $n \to \infty$, $m_1 \to \infty$, and $m_2 \to \infty$, while simultaneously $m_1/n \to$

0 and $m_1/m_2 \rightarrow c$, where c is a non-negative constant. Assume, without loss of generality, that $m_2 \geq m_1$. Then for any $-\pi < \omega < \pi$

$$\operatorname{cov}\left(\hat{f}^{(1)}(w),\, \hat{f}^{(2)}(w)\right) \;=\; \frac{2\,(1+\delta_\omega)\; m_1\; f(\omega)^2}{3\,n}\left(1+\frac{m_2-m_1}{2\,m_2}\right)+O\left(\frac{1}{n}\right)+O\left(\frac{m_2^2}{n^2}\right),$$

where $\delta_\omega = 1$ for $\omega = 0$ and $\delta_\omega = 0$ for $\omega \neq 0$.

Corollary 5.1 gives the asymptotic variance of a Bartlett estimator of the spectral density and follows from Theorem 5.1 for $m_1 = m_2 = m$.

COROLLARY 5.1 Suppose that the observations $\{X_i; i = 1, \ldots, n\}$ satisfy conditions SC. Assume that $n \rightarrow \infty$, $m_1 \rightarrow \infty$, and $m_2 \rightarrow \infty$, while simultaneously $m_1/n \rightarrow 0$ and $m_1/m_2 \rightarrow c$, where c is a non-negative constant. Assume, without loss of generality, that $m_2 \geq m_1$. Then for any $-\pi < \omega < \pi$

$$\operatorname{var}\left(\hat{f}(w)\right) \;=\; \frac{2\;(1+\delta_\omega)\; m\; f(\omega)^2}{3\,n} \;+\; O\left(\frac{1}{n}\right) \;+\; O\left(\frac{m^2}{n^2}\right),$$

where $\delta_\omega = 1$ for $\omega = 0$ and $\delta_\omega = 0$ for $\omega \neq 0$.

Priestley [1992] derives a similar variance result, but he does not provide indication of the magnitude of the second-order error terms.

Corollary 5.2 yields an equation for the asymptotic correlation between two Bartlett estimators and is obtained using the definition of correlation, Theorem 5.1 and Corollary 5.1.

COROLLARY 5.2 Suppose that the observations $\{X_i; i = 1, \ldots, n\}$ satisfy conditions SC. Assume that $n \rightarrow \infty$, $m_1 \rightarrow \infty$, and $m_2 \rightarrow \infty$, while simultaneously $m_1/n \rightarrow 0$ and $m_1/m_2 \rightarrow c$, where c is a non-negative constant. Assume, without loss of generality, that $m_2 \geq m_1$. Then for any $-\pi < \omega < \pi$

$$\operatorname{corr}\left(\hat{f}^{(1)}(\omega),\, \hat{f}^{(2)}(\omega)\right) \;=\; \left(\frac{m_1}{m_2}\right)^{\frac{1}{2}}\left(1 \;+\; \frac{m_2-m_1}{2\,m_2}\right) \;+\; O\left(\frac{m_2}{n}\right) \;+\; O\left(\frac{1}{m_1}\right),$$

The asymptotic covariance depends upon the application in a simple way: it is the product of the asymptotic variance of the Bartlett estimator with the smaller truncation point, $\mathrm{var}(\widehat{f}^{(1)}(\omega))$, and of the correction factor $1 + (m_2 - m_1)/(2m_2)$. Because the spectral density $f(\omega)$ is fixed for a given application, the asymptotic covariance (and therefore the asymptotic variances) depends only on the analyst-specified truncation points m_1 and m_2. The asymptotic correlation and the ratios of the asymptotic variances are independent of the spectral density, and therefore independent of the application. This independence is central to our motivation, which is to use linear combinations of Bartlett estimators to estimate the spectral density, without having to estimate the various covariances.

6. EQUIVALENCE OF OVERLAPPING BATCH MEANS AND BARTLETT ESTIMATORS

6.1 Introduction

Among the many procedures for estimating the standard error of the sample mean from stationary autocorrelated data are Bartlett's spectral estimator at zero frequency and overlapping batch means (OBM). In proposing the latter, Meketon and Schmeiser [1984] provided a heuristic argument that these two estimators are essentially equivalent when the Bartlett window size equals the OBM batch size. Welch [1987], in proposing and studying partial overlapping batch means, states that overlapping batch means "are related to spectral estimation via the time averaging of subsequence periodograms." That the estimators are quite similar is also clear from the three-dimensional graphs of the estimators' quadratic-form coefficients in Song and Schmeiser [1993].

We show in two ways that Bartlett and OBM estimators are equivalent: (1) their correlation is asymptotically one, and (2) their difference converges to zero in mean squared error (MSE), even after rescaling to estimate $n\,\mathrm{var}(\bar{X})$. Section 6.2 consists of a graphical analysis of the similarities between OBM and Bartlett estimators. In Section 6.3, the expected values and biases of each of these two estimators are derived. In Section 6.4, we prove the asymptotic correlation and MSE equivalence of OBM and Bartlett estimators.

6.2 Comparison of OBM and Simplified Bartlett Estimators

We first compare the OBM and the simplified Bartlett quadratic-form coefficients. Figures 6.1 and 6.2 show three-dimensional plots of $p_{ij}^{(O)}$ and $p_{ij}^{(SB)}$, respectively, as functions of i and j, for $n = 50$ observations and batch/window size $m = 10$. Both show a triangular ridge of width $2m$ along the main diagonal $i = j$. These graphs help understanding the following analysis, which is based on Equations 2.11 through 2.13 for the OBM estimator and Equation 2.17 for the simplified Bartlett spectral estimator.

There are three effects to consider when comparing the two quadratic-forms.

1. The *end effect*. For $1 \leq i, j \leq m$ or $n - m + 1 \leq i, j \leq n$

$$|p_{ij}^{(O)} - p_{ij}^{(SB)}| = O\left(\frac{1}{n^2}\right). \qquad [6.1]$$

   This effect is visible on the plots at the ends of the ridges.

2. The "scaling effect". This effect, not visible in the graphs, is due to the different denominators: n^2 for the simplified Bartlett and $(n - m + 1)(n - m)$ for the OBM. Then outside the "end effect" region and for $|i - j| \leq m$

$$p_{ij}^{(O)} = \left(1 + O\left(\frac{m}{n}\right)\right) p_{ij}^{(SB)},$$

   which implies that

$$|p_{ij}^{(O)} - p_{ij}^{(SB)}| = O\left(\frac{m}{n^3}\right). \qquad [6.2]$$

3. The "second-order effect". This effect is for the small terms off the ridges. It is most visible at the four edges of the plots. For $|i - j| > m$, $p_{ij}^{(SB)} = 0$ and $|p_{ij}^{(O)}| = O\left(m/n^3\right)$, and therefore

$$|p_{ij}^{(O)} - p_{ij}^{(SB)}| = O\left(\frac{m}{n^3}\right). \qquad [6.3]$$

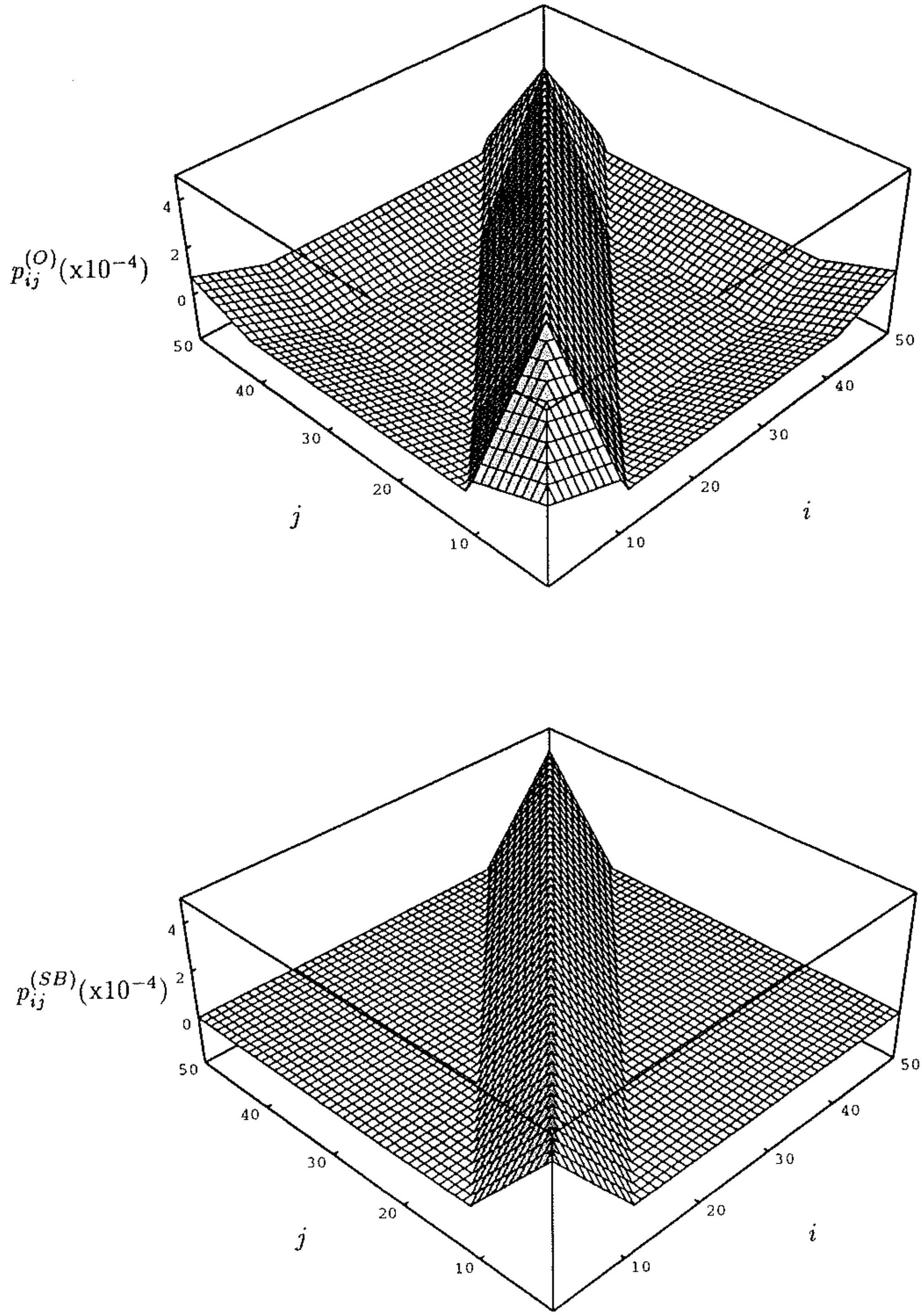

Figure 6.1 Comparison of OBM and Simplified Bartlett quadratic-form coefficients for $n = 50$ and $m = 10$.

6.3 Expected Values of OBM and Bartlett Estimators

To derive the asymptotic results of this Section, we assume that the data satisfy a set of sufficient conditions denoted by SC: $\{X_i\}$ is a general linear process of the form $X_i = \sum_{h=-\infty}^{+\infty} b_h \, \varepsilon_{i-h}$, where (1) b_h's are constants, (2) $\{\varepsilon_i\}$ is a sequence of iid random variables with finite variance, σ_ε^2, and finite fourth cumulant $k_{\varepsilon,4}$, and (3) $\sum_{h=-\infty}^{+\infty} |h|\,|b_h| < \infty$. Condition $\sum_{h=-\infty}^{+\infty} |h|\,|b_h| < \infty$ is equivalent to $b_h = O\left(|h|^{-2-\delta}\right)$, $\delta > 0$, which implies that the lag-h autocovariance, $R(h) = \sigma_\varepsilon{}^2 \sum_{i=\infty}^{+\infty} b_i \, b_{i+|h|}$, is $O\left(|h|^{-2-\delta}\right)$. We also assume, without loss of generality (see Section 2.5) that the process mean is zero. The first result is Theorem 6.1, which provides a formula to compute the expected value of an OBM estimator of V. Song and Schmeiser [1994] and Goldsman and Meketon [1986] discuss this result but do not provide a rigorous proof.

THEOREM 6.1 Suppose that the observations $\{X_i; i = 1, \ldots, n\}$ satisfy conditions SC. Assume, without loss of generality that the process mean is zero. Let $\widehat{V}^{(O)}$ be an OBM estimator of V, the variance of the sample mean, with batch size m. Assume that $n \to \infty$ and m is a function of n satisfying $m \to \infty$ and $m/n \to 0$ as $n \to \infty$. Then

$$\mathrm{E}(\widehat{V}^{(O)}) \quad = \quad \frac{R(0)\,\gamma_0}{n} - \frac{R(0)\,\gamma_1}{m\,n} + O\left(\frac{m}{n^2}\right).$$

Proof: By Equation 2.10

$$\widehat{V}^{(O)} \;=\; \sum_{r=1}^{n} \sum_{s=1}^{n} p_{rs}^{(O)} \, X_r X_s .$$

Taking expectations and using Equation 2.13,

$$\mathrm{E}(\widehat{V}^{(O)}) \;=\; R(0) \sum_{r=1}^{n} \sum_{s=\max(1,r-m)}^{\min(n,r+m)} p_{rs}^{(O)} \, \rho(r-s) \;+\; \Delta_1,$$

where,

$$|\Delta_1| \;=\; O\left(\sum_{r=1}^{n}\sum_{s=1}^{r-m-1}\frac{m}{n^3}\,\rho(r-s) \;+\; \sum_{r=1}^{n}\sum_{s=r+m+1}^{n}\frac{m}{n^3}\,\rho(r-s)\right)$$

$$\;=\; O\left(\frac{m}{n^2}\sum_{h=-n}^{-m-1}\rho(h) \;+\; \frac{m}{n^2}\sum_{h=m+1}^{n}\rho(h)\right).$$

As discussed earlier, assumption $\sum_{h=-\infty}^{+\infty}|h|\,|b_h|<\infty$ implies that $\rho(h)=O(|h|^{-2-\delta})$ and therefore $\sum_{h>m}\rho(h)=O(m^{-1})$ and $\sum_{h<-m}\rho(h)=O(m^{-1})$, yielding

$$|\Delta_1| \;=\; O\left(\frac{1}{n^2}\right).$$

Then

$$\mathrm{E}(\widehat{V}^{(O)}) \;=\; R(0)\sum_{r=1}^{n}\sum_{s=\max(1,r-m)}^{\min(n,r+m)} p_{rs}^{(O)}\,\rho(r-s) \;+\; O\left(\frac{1}{n^2}\right).$$

Replacing the lower and upper bound of s, respectively $\max(1,r-m)$ and $\min(n,r+m)$, by $r-m$ and $r+m$,

$$\mathrm{E}(\widehat{V}^{(O)}) \;=\; R(0)\sum_{r=1}^{n}\sum_{s=r-m}^{r+m} p_{rs}^{(O)}\,\rho(r-s) \;+\; O\left(\frac{1}{n^2}\right) \;-\; \Delta_2,$$

where, by Equation 2.13,

$$|\Delta_2| \;=\; O\left(\sum_{r=1}^{m}\sum_{s=r-m}^{0}\frac{1}{n^2}\,\rho(r-s) + \sum_{r=n-m+1}^{n}\sum_{s=n}^{r+m}\frac{1}{n^2}\,\rho(r-s)\right).$$

By assumption, $\rho(h)=O(|h|^{-2-\delta})$, yielding

$$|\Delta_2| \;=\; O\left(\frac{1}{n^2}\sum_{r=1}^{m}\sum_{h=-m}^{-r}\frac{1}{|h|^{2+\delta}}\right) \;=\; O\left(\frac{1}{n^2}\sum_{h=1}^{m}\frac{1}{|h|^{1+\delta}}\right) \;=\; O\left(\frac{1}{n^2}\right).$$

Therefore, by Equation 2.13,

$$\mathrm{E}(\widehat{V}^{(O)}) \;=\; R(0)\sum_{r=1}^{n}\sum_{s=r-m}^{r+m}\frac{1}{n^2}\left(1-\frac{|r-s|}{m}\right)\left[1+O\left(\frac{m}{n}\right)\right]\rho(r-s)+O\left(\frac{1}{n^2}\right)+\Delta_3,$$

where,

$$|\Delta_3| \;=\; O\left(\sum_{r=1}^{m}\sum_{s=1}^{m}\frac{1}{n^2}\,\rho(r-s) + \sum_{r=n-m+1}^{n}\sum_{s=n-m+1}^{n}\frac{1}{n^2}\,\rho(r-s)\right).$$

Assumption $\sum_{h=-\infty}^{+\infty} |h|\,|b_h| < \infty$ implies that $\sum_h \rho(h) = O(1)$. Then

$$|\Delta_3| \;=\; O\left(\frac{m}{n^2}\right).$$

Rewriting $\mathrm{E}(\widehat{V}^{(O)})$ and letting $h = r - s$ yields

$$\mathrm{E}(\widehat{V}^{(O)}) \;=\; \frac{R(0)}{n}\left[1 + O\left(\frac{m}{n}\right)\right]\sum_{h=-m}^{m}\left(1 - \frac{|h|}{m}\right)\rho(r-s) \;+\; O\left(\frac{m}{n^2}\right).$$

By Equations 2.4 and 2.6 the result follows.

Corollary 6.1 states that the bias of an OBM estimator goes to zero as n increases.

COROLLARY 6.1 Suppose that the observations $\{X_i; i = 1, \ldots, n\}$ satisfy conditions SC. Let $\widehat{V}^{(O)}$ be an OBM estimator of V, the variance of the sample mean, with batch size m. Assume that $n \to \infty$ and m is a function of n satisfying $m \to \infty$ and $m/n \to 0$ as $n \to \infty$. Then

$$\mathrm{bias}\left(\widehat{V}^{(O)}\right) \;=\; -\,\frac{R(0)\,\gamma_1}{m\,n} + O\left(\frac{m}{n^2}\right).$$

Proof: By definition of bias

$$\mathrm{bias}\left(\widehat{V}^{(O)}\right) \;=\; \mathrm{E}(\widehat{V}^{(O)}) - V.$$

By Equation 2.7,

$$V \;=\; \frac{\gamma_0\,R(0)}{n} + O\left(\frac{1}{n^2}\right).$$

Then, by Theorem 6.1, the result follows.

Notice that $\mathrm{bias}\left(\widehat{V}^{(O)}\right) \simeq -\,(R(0)\,\gamma_1)\,/\,(m\,n)$ if $m = o(\sqrt{n})$ and $\mathrm{bias}\left(\widehat{V}^{(O)}\right) = O\,(m/n^2)$ if $m \geq O(\sqrt{n})$.

Theorem 6.2 provides a formula for the expectation of a Bartlett estimator of V.

THEOREM 6.2 Suppose that the observations $\{X_i; i = 1, \ldots, n\}$ satisfy conditions SC. Assume, without loss of generality that the process mean is zero. Let $\widehat{V}^{(B)}$ be a Bartlett estimator of V, the variance of the sample mean, with window size m. Assume that $n \to \infty$ and m is a function of n satisfying $m \to \infty$ and $m/n \to 0$ as $n \to \infty$. Then

$$\mathrm{E}(\widehat{V}^{(B)}) = \frac{R(0)\,\gamma_0}{n} - \frac{R(0)\,\gamma_1}{m\,n} + O\left(\frac{m}{n^2}\right).$$

Theorem 6.2 is proved making the same arguments as in Theorem 6.1, but basing the arguments in Equations 2.18 and 2.19 rather than in Equations 2.10 and 2.13. We omit the details here.

In Corollary 6.2, the bias of Bartlett estimators of V is stated. Politis and Romano [1993] derive a similar result, in the context of spectral density estimation. The proof of Corollary 6.2 is similar to the proof of Corollary 6.1.

COROLLARY 6.2 Suppose that the observations $\{X_i; i = 1, \ldots, n\}$ satisfy conditions SC. Let $\widehat{V}^{(B)}$ be a Bartlett estimator of V, the variance of the sample mean, with window size m. Assume that $n \to \infty$ and m is a function of n satisfying $m \to \infty$ and $m/n \to 0$ as $n \to \infty$. Then

$$\text{bias}\left(\widehat{V}^{(B)}\right) = -\frac{R(0)\,\gamma_1}{m\,n} + O\left(\frac{m}{n^2}\right).$$

6.4 Asymptotic Equivalence of OBM and Bartlett Estimators

Here we characterize the asymptotic relationship between OBM with batch size m and Bartlett spectral estimator with window size m from n observations as a MSE convergence. Theorem 6.3 states that the asymptotic correlation between OBM with batch size m and the Bartlett estimator with window size m is 1. This theorem results directly from Corollary 4.2 for $m = m_1 = m_2$.

THEOREM 6.3 Suppose that the observations $\{X_i; i = 1, \ldots, n\}$ satisfy conditions SC. Let $\widehat{V}^{(O)}$ and $\widehat{V}^{(B)}$ be, respectively, an OBM and a Bartlett estimator of V with batch/window size m. Assume that $n \to \infty$ and m is a function of n satisfying $m \to \infty$ and $m/n \to 0$ as $n \to \infty$. Then

$$\operatorname{corr}\left(\widehat{V}^{(O)}, \widehat{V}^{(B)}\right) = 1 + O\left(\frac{m}{n}\right) + O\left(\frac{1}{m}\right).$$

Theorem 6.4 states that the OBM estimator converges in MSE to the Bartlett estimator.

THEOREM 6.4 Suppose that the observations $\{X_i; i = 1, \ldots, n\}$ satisfy conditions SC. Let $\widehat{V}^{(O)}$ and $\widehat{V}^{(B)}$ be, respectively, an OBM and a Bartlett estimator of V with batch/window size m. Assume that $n \to \infty$ and m is a function of n satisfying $m \to \infty$ and $m/n \to 0$ as $n \to \infty$. Then

$$\frac{n^3}{m} \mathrm{E}\left[\left(\widehat{V}^{(O)} - \widehat{V}^{(B)}\right)^2\right] \xrightarrow{n \to \infty} 0.$$

Proof: Expressing the squared difference in terms of variance and covariance

$$\begin{aligned}
\mathrm{E}\left(\widehat{V}^{(O)} - \widehat{V}^{(B)}\right)^2 &= \mathrm{E}\left[\left(\widehat{V}^{(O)} - \mathrm{E}\widehat{V}^{(O)}\right) - \left(\widehat{V}^{(B)} - \mathrm{E}\widehat{V}^{(B)}\right) + \left(\mathrm{E}\widehat{V}^{(O)} - \mathrm{E}\widehat{V}^{(B)}\right)\right]^2 \\
&= \mathrm{E}\left[\left(\widehat{V}^{(O)} - \mathrm{E}\widehat{V}^{(O)}\right)^2 + \left(\widehat{V}^{(B)} - \mathrm{E}\widehat{V}^{(B)}\right)^2\right. \\
&\qquad \left. -2\left(\widehat{V}^{(O)} - \mathrm{E}\widehat{V}^{(O)}\right)\left(\widehat{V}^{(B)} - \mathrm{E}\widehat{V}^{(B)}\right)\right] + \left[\mathrm{E}\widehat{V}^{(O)} - \mathrm{E}\widehat{V}^{(B)}\right]^2 \\
&= \operatorname{var}\left(\widehat{V}^{(O)}\right) + \operatorname{var}\left(\widehat{V}^{(B)}\right) - 2\operatorname{cov}\left(\widehat{V}^{(O)}, \widehat{V}^{(B)}\right) \\
&\qquad + \left[\mathrm{E}\widehat{V}^{(O)} - \mathrm{E}\widehat{V}^{(B)}\right]^2.
\end{aligned}$$

By Theorems 6.1 and 6.2

$$\left(\mathrm{E}\widehat{V}^{(O)} - \mathrm{E}\widehat{V}^{(B)}\right)^2 = O\left(\frac{m^2}{n^4}\right).$$

Rewriting $\mathrm{E}\left[\left(\widehat{V}^{(O)} - \widehat{V}^{(B)}\right)^2\right]$ and using Theorem 4.1 and Corollary 4.1

$$\mathrm{E}\left[\left(\widehat{V}^{(O)} - \widehat{V}^{(B)}\right)^2\right] = O\left(\frac{m^2}{n^4}\right) + O\left(\frac{1}{n^3}\right).$$

The result follows, since

$$\frac{n^3}{m}\,\mathrm{E}\left[\left(\widehat{V}^{(O)}-\widehat{V}^{(B)}\right)^2\right] \;=\; O\left(\frac{m}{n}\right)+O\left(\frac{1}{m}\right).$$

Theorem 6.4 characterizes the asymptotic convergence as being in MSE which is a stronger condition than convergence in probability, in the sense that the former implies the latter [Priestley, 1992, p.151].

Theorems 6.2, 6.3, and 6.4 are also valid for the simplified Bartlett estimator, because the simplified Bartlett estimator coefficients (Equation 2.17) differ from the corresponding Bartlett estimator coefficients (Equation 2.19) by at most $O(m/n^3)$.

7. ESTIMATING VARIANCE OF THE SAMPLE MEAN: OPTIMAL BATCH SIZE ESTIMATION AND 1-2-1 OVERLAPPING BATCH MEANS

7.1 Introduction

One approach for developing robust and computationally efficient methodology to estimate V is minimizing the mean squared error (MSE). For estimators parameterized by batch size (for example, NBM, OBM, and standardized-time-series) this approach requires determining the optimal batch size for a given data sample of size n. The optimal batch size is a function of the sum of the autocorrelations, $\gamma_0 = \sum_{h=-\infty}^{+\infty} \rho(h)$, and of the weighted sum of autocorrelations, $\gamma_1 = \sum_{h=-\infty}^{+\infty} |h|\rho(h)$. It follows that the natural way to estimate γ_0 and γ_1 is estimating all $n-1$, or at least $m(n, \gamma_0, \gamma_1)$ of the autocorrelations, leading to methods that are $O(n^2)$ (i.e., computation time is proportional to n^2) and $O(m\ n)$, respectively. We propose a more efficient way of estimating the optimal batch size. Our method is $O(n)$ and has good MSE performance even when compared to the idealistic performance that could be obtained if γ_0 and γ_1 were known.

This chapter is organized as follows. In Section 7.2 we define the "approximate optimal batch size". In Section 7.3 we develop methodology for estimating γ_1/γ_0, the center of gravity of the autocorrelations. Because $V \overset{n\to\infty}{\longrightarrow} 0$, we redefine MSE-consistency in the context of V estimation as $\text{mse}(n\ \widehat{V}_m) \overset{n\to\infty}{\longrightarrow} 0$ instead of $\text{mse}(\widehat{V}_m) \overset{n\to\infty}{\longrightarrow} 0$. We state conditions for OBM estimators to be MSE-consistent. We also derive $O(n)$ estimators $\widehat{\gamma}_0$ and $\widehat{\gamma}_1$ of γ_0 and γ_1, respectively, and state conditions for $\widehat{\gamma}_0$ and $\widehat{\gamma}_1$ to be MSE-consistent. In Section 7.4 we propose the 1–2–1 OBM algorithm, a computationally efficient — $O(n)$ — and relatively simple algorithm for estimating V using four OBM estimates. We evaluate the performance of this algorithm in Section 7.5, showing that the 1–2–1 OBM has good properties. Theoretical results suggest that (1) the 1–2–1 OBM is MSE-consistent; and (2) the 1–2–1 OBM performance approaches, as the sample size increases, the idealistic performance that could be obtained if the data properties γ_0 and γ_1 were known. The results of an extensive set of Monte Carlo experiments are also discussed in Section 7.5 and support the theoretical indications. The three processes used in the Monte Carlo experiments are reviewed in Chapters 2 and 3. Appendix C contains a FORTRAN implementation of the 1–2–1 OBM algorithm.

The methodology that we develop is based on OBM estimation. Analogous development can be considered for other estimators parameterized by batch size. Also, algorithms structurally different from the 1–2–1 OBM (for example, with a different number of estimation stages) could be based on the theoretical results derived in this chapter.

7.2 Approximate Optimal Batch Size

The major difficulty for OBM is to select the batch size. One approach for selecting batch size is minimizing the mean squared error. Pedrosa and Schmeiser [1993c and 1993d] (see also Chapters 4 and 6) prove asymptotic results discussed by Song and Schmeiser [1994] and Goldsman and Meketon [1986] regarding the bias, variance, and MSE of OBM estimators. These asymptotic ($n \to \infty$, $m \to \infty$, and $m/n \to 0$)

properties are:

$$\text{bias}(\widehat{V}_m) \;=\; -\frac{\gamma_1\, R(0)}{m\, n} \;+\; O\left(\frac{m}{n^2}\right), \qquad [7.1]$$

$$\text{var}(\widehat{V}_m) \;=\; \frac{4\, m\, \gamma_0^2\, R(0)^2}{3\, n^3} \;+\; O\left(\frac{1}{n^3}\right) \;+\; O\left(\frac{m^2}{n^4}\right), \qquad [7.2]$$

and

$$\text{mse}(\widehat{V}_m) \;=\; \frac{\gamma_1^2\, R(0)^2}{m^2\, n^2} \;+\; \frac{4\, m\, \gamma_0^2\, R(0)^2}{3\, n^3} \;+\; O\left(\frac{1}{n^3}\right) \;+\; O\left(\frac{m^2}{n^4}\right). \qquad [7.3]$$

Pedrosa and Schmeiser [1993c] (see Chapter 4) also derived the asymptotic covariance ($n \to \infty$, $m_1 \to \infty$, $m_1/n \to 0$, $m_1/m_2 \to$ constant, and $m_1 \leq m_2$) between two OBM estimators with batch sizes m_1 and m_2,

$$\text{cov}\left(\widehat{V}_{m_1}, \widehat{V}_{m_2}\right) \;=\; \frac{4\, m_1\, \gamma_0^2\, R(0)^2}{3\, n^3}\left(1 \;+\; \frac{m_2 - m_1}{2\, m_2}\right) \;+\; O\left(\frac{1}{n^3}\right) \;+\; O\left(\frac{m_2^2}{n^4}\right). \qquad [7.4]$$

One set of sufficient conditions that imply Equations 7.1 through 7.4 is [Pedrosa and Schmeiser, 1993c] $\{X_i\}$ is a general linear process of the form $X_i = \sum_{h=-\infty}^{+\infty} b_h\, \varepsilon_{i-h}$, where (1) b_h's are constants, (2) $\{\varepsilon_i\}$ is a sequence of iid random variables with finite variance, σ_ε^2, and finite fourth cumulant $k_{\varepsilon,4}$, and (3) $\sum_{h=-\infty}^{+\infty} |h|\, |b_h| < \infty$.

Define "approximate MSE" as

$$\vec{\text{mse}}(\widehat{V}_m) \;=\; \frac{\gamma_1^2\, R(0)^2}{m^2\, n^2} \;+\; \frac{4\, m\, \gamma_0^2\, R(0)^2}{3\, n^3}. \qquad [7.5]$$

Define "approximate optimal batch size" $\vec{m}^*$ as the batch size that minimizes $\vec{\text{mse}}(\widehat{V}_m)$ plus 1, i.e.,

$$\vec{m}^* \;=\; \left[1.5\, n \left(\frac{\gamma_1}{\gamma_0}\right)^2\right]^{\frac{1}{3}} \;+\; 1. \qquad [7.6]$$

By adding 1, $\vec{m}^*$ is the optimal batch size for $\gamma_1 = 0$ (uncorrelated observations). Finite-sample results suggest that batch size $\vec{m}^*$ does not differ much from the real optimal batch size, even for small sample sizes [Song and Schmeiser, 1994]. Notice that the vector symbol denotes approximation and the star denotes optimality.

7.3 OBM Methodology to Estimate γ_0 and γ_1

In this section we develop methodology for estimating the sum of autocorrelations γ_0 and the weighted sum of autocorrelations γ_1.

7.3.1 MSE Consistency of OBM Estimators

The MSE is not adequate to measure the performance of estimators of the variance of the sample mean, V, because (1) V goes to zero when the sample size increases (Equation 2.7), and (2) the mean squared error of an OBM estimator also goes to zero when the sample size increases and the batch size is well chosen, i.e., $m \rightarrow \infty$ and $m/n \rightarrow 0$ (Equation 7.3). We define "standardized MSE" as the ratio of the MSE of the variance estimator to the square of the variance of the point estimator. Therefore, in the context of V estimation we say that $\widehat{V}$ is MSE-consistent if $\text{mse}(\widehat{V})/V^2 \rightarrow 0$ (or equivalently, if $\text{mse}(n\,\widehat{V}) \rightarrow 0$) when the sample size n increases.

Theorem 7.1 provides some insight on how to select the batch size if no information about the correlation structure is available.

THEOREM 7.1 Let V denote the variance of the sample mean of stationary data $\{X_i; i = 1, \ldots, n\}$. Let $\widehat{V}_m$ be an OBM estimator with batch size m. Assume that Equations 7.3 holds. Let $m \rightarrow \infty$ and $m/n \rightarrow 0$ as $n \rightarrow \infty$. Then

$$\frac{\text{mse}(\widehat{V}_m)}{V^2} \xrightarrow{n \rightarrow \infty} 0.$$

Proof: From Equation 2.7

$$V = \frac{\gamma_0 \, R(0)}{n} + O\left(\frac{1}{n^2}\right).$$

Then

$$\frac{\text{mse}(\widehat{V}_m)}{V^2} = \text{mse}(\widehat{V}_m) \left[\frac{\gamma_0^2 \, R(0)^2}{n^2} + O\left(\frac{1}{n^3}\right)\right]^{-1}$$

$$= \frac{n^2}{\gamma_0^2\, R(0)^2}\, \mathrm{mse}(\widehat{V}_m)\, \left[1 + O\left(\frac{1}{n}\right)\right].$$

By Equation 7.3

$$\frac{\mathrm{mse}(\widehat{V}_m)}{V^2} = \left[\frac{1}{m^2}\left(\frac{\gamma_1}{\gamma_0}\right)^2 + \frac{4\,m}{3\,n} + \frac{b_{m,n}}{n^2\,\gamma_0^2\,R(0)^2} + \frac{m\,c_{m,n}}{n\,\gamma_0^2\,R(0)^2}\right]\left[1 + O\left(\frac{1}{n}\right)\right] \xrightarrow{n \to \infty} 0.$$

Corollary 7.1 follows directly from Theorem 7.1 and suggests a formula to select batch size.

COROLLARY 7.1 Let V denote the variance of the sample mean of stationary data $\{X_i; i = 1, \ldots, n\}$. Let $\widehat{V}_m$ be an OBM estimator with batch size m defined by $m = c\,n^{\alpha} + d$, where c, d, and α are finite constants, $c > 0$, and $0 < \alpha < 1$. Assume that Equations 7.3 holds. Then

$$\frac{\mathrm{mse}(\widehat{V}_m)}{V^2} \xrightarrow{n \to \infty} 0.$$

7.3.2 OBM Methodology to Estimate γ_0

Theorem 7.2 provides an MSE-consistent estimator of the sum of autocorrelations, γ_0, using OBM, as suggested by Equation 2.5.

THEOREM 7.2 Let V denote the variance of the sample mean of stationary data $\{X_i; i = 1, \ldots, n\}$ with sum of autocorrelations $\gamma_0 = \sum_{h=-\infty}^{+\infty} \rho(h)$. Assume that Equations 7.1, 7.2 and 7.3 hold. Let $m \to \infty$ and $m/n \to 0$ as $n \to \infty$. Define the estimator

$$\widehat{\gamma}_0 = \frac{n\,\widehat{V}_m}{R(0)},$$

where R(0) is the marginal variance and $\widehat{V}_m$ is the OBM estimator with batch size m. Then

(1) $\quad \mathrm{E}\,(\widehat{\gamma}_0) = \gamma_0 - \dfrac{\gamma_1}{m} + O\left(\dfrac{m}{n}\right),$

$$(2) \quad \text{bias}(\widehat{\gamma}_0) = -\frac{\gamma_1}{m} + O\left(\frac{m}{n}\right),$$

$$(3) \quad \text{var}(\widehat{\gamma}_0) = \frac{4\,m\,\gamma_0^2}{3\,n} + O\left(\frac{1}{n}\right) + O\left(\frac{m^2}{n^2}\right),$$

and

$$(4) \quad \text{mse}(\widehat{\gamma}_0) = \frac{4\,m\,\gamma_0^2}{3\,n} + \frac{\gamma_1^2}{m^2} + O\left(\frac{1}{n}\right) + O\left(\frac{m^2}{n^2}\right) \xrightarrow{n\to\infty} 0.$$

Proof: We start proving the first and second results. Taking expectations and using the definition of $\widehat{\gamma}_0$,

$$\mathrm{E}(\widehat{\gamma}_0) = \frac{n}{R(0)}\,\mathrm{E}\left(\widehat{V}_m\right).$$

By Equation 7.1,

$$\mathrm{E}(\widehat{\gamma}_0) = \frac{n}{R(0)}\left[V - \frac{\gamma_1\,R(0)}{m\,n} + O\left(\frac{m}{n^2}\right)\right].$$

By Equation 2.7, $(nV/R(0)) = \gamma_0 + O(1/n)$. Therefore,

$$\mathrm{E}(\widehat{\gamma}_0) = \gamma_0 - \frac{\gamma_1}{m} + O\left(\frac{m}{n}\right),$$

and

$$\text{bias}(\widehat{\gamma}_0) = -\frac{\gamma_1}{m} + O\left(\frac{m}{n}\right).$$

Now we prove the third result. By definition of $\widehat{\gamma}_0$ and definition of variance,

$$\text{var}(\widehat{\gamma}_0) = \frac{n^2\text{var}\left(\widehat{V}_m\right)}{R(0)^2}.$$

By Equation 7.2,

$$\begin{aligned}\text{var}(\widehat{\gamma}_0) &= \frac{n^2}{R(0)^2}\left[\frac{4\,m\,\gamma_0^2\,R(0)^2}{3\,n^3} + O\left(\frac{1}{n^3}\right) + O\left(\frac{m^2}{n^4}\right)\right]\\ &= \frac{4\,m\,\gamma_0^2}{3\,n} + O\left(\frac{1}{n}\right) + O\left(\frac{m^2}{n^2}\right).\end{aligned}$$

The MSE result follows from the bias and variance results.

7.3.3 OBM Methodology to Estimate γ_1

Theorem 7.3 suggests a formula to estimate γ_1.

THEOREM 7.3 Let $\widehat{V}_{m_1}$ and $\widehat{V}_{m_2}$ be two OBM estimators of the variance of the sample mean V with, respectively, batch sizes m_1 and m_2, and $m_1 \neq m_2$. Assume that Equation 7.1 holds. Let $m_1/n^{\alpha} \overset{n\to\infty}{\longrightarrow} a$, $m_1/m_2 \overset{n\to\infty}{\longrightarrow} b$, $a > 0$, $b > 0$, and $0 < \alpha \leq 1/3$. Then

$$\gamma_1 \;=\; \frac{n\, m_1\, m_2}{R(0)\,(m_2 - m_1)} \left[\mathrm{E}(\widehat{V}_{m_2}) - \mathrm{E}(\widehat{V}_{m_1})\right] \;+\; O\left(\frac{m_1^3}{n}\right).$$

Proof: From Equation 7.1

$$\mathrm{E}(\widehat{V}_{m_1}) - V \;=\; -\frac{\gamma_1\, R(0)}{m_1\, n} \;+\; O\left(\frac{m_1}{n^2}\right),$$

and

$$\mathrm{E}(\widehat{V}_{m_2}) - V \;=\; -\frac{\gamma_1\, R(0)}{m_2\, n} \;+\; O\left(\frac{m_2}{n^2}\right).$$

Then

$$\begin{aligned}
\mathrm{E}(\widehat{V}_{m_2}) - \mathrm{E}(\widehat{V}_{m_1}) \;&=\; \frac{\gamma_1\, R(0)}{n}\left(\frac{1}{m_1} - \frac{1}{m_2}\right) \;+\; O\left(\frac{m_1}{n^2}\right) \\
&=\; \frac{\gamma_1\, R(0)\,(m_2 - m_1)}{n\, m_1\, m_2} \;+\; O\left(\frac{m_1}{n^2}\right).
\end{aligned}$$

The result follows after simple algebraic manipulation.

Theorem 7.4 provides a formula to estimate γ_1 from two OBM estimates $\widehat{V}_{m_1}$ and $\widehat{V}_{m_2}$.

THEOREM 7.4 Let $\widehat{V}_{m_1}$ and $\widehat{V}_{m_2}$ be two OBM estimators of the variance of the sample mean V with, respectively, batch sizes m_1 and m_2. Let $m_1/n^{\alpha} \overset{n\to\infty}{\longrightarrow} a$, $a > 0$, and $0 < \alpha \leq 1/3$. Let $m_2 = m_1 + \lambda_{m_1}$, where λ_{m_1} is a function of m_1 such that

$\lambda_{m_1} \geq 1$ and $\lambda_{m_1}/m_1 \stackrel{n \to \infty}{\longrightarrow}$ constant. Assume that Equations 7.1 through 7.4 hold. Assume further that the marginal variance $R(0)$ is known. Let

$$\widehat{\gamma}_1 = \frac{n\, m_1\, m_2}{R(0)\,(m_2 - m_1)} \left(\widehat{V}_{m_2} - \widehat{V}_{m_1}\right)$$

be an estimator of γ_1, the weighted sum of correlations. Then

$$\text{(1)} \quad \text{var}(\widehat{\gamma}_1) = \frac{4\,\gamma_0^2\, m_1^2\,(m_1 + \lambda_{m_1})}{3\,n} + O\left(\frac{m_2^2}{n}\right),$$

$$\text{(2)} \quad \text{bias}(\widehat{\gamma}_1) = O\left(\frac{m_2^2}{n}\right),$$

and

$$\text{(3)} \quad \text{mse}(\widehat{\gamma}_1) = \frac{4\,\gamma_0^2\, m_1^2\,(m_1 + \lambda_{m_1})}{3\,n} + O\left(\frac{m_2^2}{n}\right).$$

Proof: By definition of variance,

$$\text{var}\left(\widehat{V}_{m_2} - \widehat{V}_{m_1}\right) = \text{var}(\widehat{V}_{m_1}) + \text{var}(\widehat{V}_{m_2}) - 2\,\text{cov}(\widehat{V}_{m_1}, \widehat{V}_{m_2}).$$

By Equation 7.2 and because $m_2/m_1 \stackrel{n \to \infty}{\longrightarrow}$ constant,

$$\text{var}(\widehat{V}_{m_1}) = \frac{4\, m_1\, \gamma_0^2\, R(0)^2}{3\,n^3} \left[1 + O\left(\frac{1}{m_2}\right) + O\left(\frac{m_2}{n}\right)\right],$$

and

$$\text{var}(\widehat{V}_{m_2}) = \frac{4\, m_2\, \gamma_0^2\, R(0)^2}{3\,n^3} \left[1 + O\left(\frac{1}{m_2}\right) + O\left(\frac{m_2}{n}\right)\right].$$

Similarly, by Equation 7.4,

$$\text{cov}\left(\widehat{V}_{m_1}, \widehat{V}_{m_2}\right) = \frac{4\, m_1\, \gamma_0^2\, R(0)^2}{3\,n^3} \left(1 + \frac{m_2 - m_1}{2\,m_2}\right) \left[1 + O\left(\frac{1}{m_2}\right) + O\left(\frac{m_2}{n}\right)\right].$$

After simple algebra

$$\text{var}\left(\widehat{V}_{m_2} - \widehat{V}_{m_1}\right) = \frac{4\,\gamma_0^2\, R(0)^2\,(m_2 - m_1)^2}{3\,n^3\, m_2} \left[1 + O\left(\frac{1}{m_2}\right) + O\left(\frac{m_2}{n}\right)\right].$$

Replacing m_2 by $m_1 + \lambda_{m_1}$,

$$\mathrm{var}(\widehat{\gamma}_1) = \frac{4\,\gamma_0^2\, m_1^2\,(m_1 + \lambda_{m_1})}{3\,n}\left[1 + O\left(\frac{1}{m_2}\right) + O\left(\frac{m_2}{n}\right)\right].$$

Therefore,

$$\mathrm{var}(\widehat{\gamma}_1) = \frac{4\,\gamma_0^2\, m_1^2\,(m_1 + \lambda_{m_1})}{3\,n} + O\left(\frac{m_2^2}{n}\right).$$

Now we prove the second result. By Equation 7.1, and because $m_2/m_1 \overset{n\to\infty}{\longrightarrow}$ constant,

$$\mathrm{E}(\widehat{V}_{m_1}) - V = -\frac{\gamma_1\, R(0)}{n\, m_1}\left[1 + O\left(\frac{m_2^2}{n}\right)\right],$$

and

$$\mathrm{E}(\widehat{V}_{m_2}) - V = -\frac{\gamma_1\, R(0)}{n\, m_2}\left[1 + O\left(\frac{m_2^2}{n}\right)\right].$$

Then

$$\mathrm{E}(\widehat{V}_{m_2}) - \mathrm{E}(\widehat{V}_{m_1}) = \frac{\gamma_1\, R(0)\,(m_2 - m_1)}{n\, m_1\, m_2}\left[1 + O\left(\frac{m_2^2}{n}\right)\right].$$

By definition of $\widehat{\gamma}_1$

$$\mathrm{E}(\widehat{\gamma}_1) = \gamma_1\left[1 + O\left(\frac{m_2^2}{n}\right)\right],$$

yielding

$$\mathrm{bias}(\widehat{\gamma}_1) = O\left(\frac{m_2^2}{n}\right).$$

The third result is obtained from the two former results after simple algebra.

Corollary 7.2 follows directly from Theorem 7.4 and provides a guideline on how to select batch size m_2 to estimate γ_1.

COROLLARY 7.2 Assume the conditions of Theorem 7.4. Then batch size $m_2 = m_1 + 1$ minimizes $\mathrm{var}(\widehat{\gamma}_1)$ and $\mathrm{mse}(\widehat{\gamma}_1)$.

Theorem 7.4 indicates that estimator $\widehat{\gamma}_1$ converges only if $m_1 = o(\sqrt[3]{n})$ and that m_1 should be selected as small as possible. But m_1 must also go to infinity so that the effects of autocorrelation are included within the batch means. We need empirical guidelines to aid on the selection of m_1.

7.4 1–2–1 OBM Algorithm for Estimating Variance of the Sample Mean

We would like to use the OBM method with the optimal batch size to estimate the variance of the sample mean. To determine the asymptotic approximate optimal batch size (Equation 7.6) we need to know γ_1/γ_0. One possibility is to estimate each individual correlation and then estimate γ_0 and γ_1. We propose a computationally more efficient method that estimates the center of gravity γ_1/γ_0 without estimating individual correlations. The building blocks of our method are: (1) OBM estimators; (2) the asymptotic results from Chapters 4 and 6; and (3) the theory developed on Section 7.3 to estimate γ_0 and γ_1. To estimate the center of gravity of the autocorrelations, γ_1/γ_0, the algorithm still needs to select a batch size m_1, but now the level of decision is lower. The empirical results in Section 7.5 indicate that the performance of $\widehat{V}_{\widehat{m}^*}$ is not sensitive to minor changes in m_1.

Good values of m_1 depend on the correlation structure of the data. We assume that the user does not provide any information about the data to be analyzed. Therefore, an initial assessment of the correlation structure must be obtained through an initial estimate of γ_0.

7.4.1 Generating an Initial Estimate of γ_0

From Theorem 7.2, $\widehat{\gamma}_0$ can be obtained from an OBM estimate $\widehat{V}_{m_0}$ with batch size m_0 using

$$\widehat{\gamma}_0 = \frac{n\,\widehat{V}_{m_0}}{\widehat{R}(0)}.$$

We set $m_0 = a_0\, n^{1/r_0} + b_0$, where a_0, b_0, and r_0 are constants, $a_0 > 0$, and $r_0 \geq 2$. This choice is a bit arbitrary, but by Theorem 7.2, $m_0 = a_0\, n^{1/r_0} + b_0$ leads to $\text{mse}(\widehat{\gamma}_0) \xrightarrow{n\to\infty} 0$. In the Monte Carlo experiment of Section 7.5 we use $m_0 = \sqrt{n}$.

Other choices of m_0 may also generate good initial estimates of γ_0, but our empirical results show that this choice works well.

7.4.2 Estimating Optimal Batch Size

The derivation of the asymptotic results indicate that batch sizes neeed to be large enough to include the effects of autocorrelation within batch means. An approximate measure of these effects is max $\{\gamma_0,\ 1/\gamma_0\}$. We set

$$m_1 = \max\left\{\mathrm{a}_1\, \mathrm{n}^{1/r_1}\ +\ \mathrm{b}_1\ ,\ \mathrm{c}_1 \max\{\widehat{\gamma}_0,\ 1/\widehat{\gamma}_0\}\ +\ \mathrm{d}_1\right\},$$

where a_1, b_1, c_1, d_1, and r_1 are constants, $a_1 > 0$, $r_1 > 3$, and $c_1 > 0$. For small sample sizes, we verified experimentally (see Section 7.5) that good performance is obtained for $m_1 = 0.75 \max\{\widehat{\gamma}_0,\ 1/\widehat{\gamma}_0\}$. For large sample sizes, Theorem 7.4 indicates that $m_1 = o(\sqrt[3]{n})$. Therefore we suggest the choice we make in Section 7.5, which is $m_1 = \max\{\sqrt[6]{n},\ 0.75 \max\{\widehat{\gamma}_0,\ 1/\widehat{\gamma}_0\}\}$. Selecting m_2 is easier, since by Corollary 7.2 the best is to set $m_2 = m_1 + 1$. Then, applying Theorem 7.2 and 7.4, estimates $\widehat{\gamma}_0$ and $\widehat{\gamma}_1$ are obtained, respectively. The estimate of the center of gravity γ_1/γ_0 follows naturally. Finally, Equation 7.6 yields an estimate of the optimal batch size.

7.4.3 1–2–1 OBM Algorithm

We can combine the results of Sections 7.4.1 and 7.4.2 in many ways to estimate V. Here we state a relatively simple algorithm, 1–2–1 OBM, whose performance we study in Section 7.5. The 1–2–1 OBM algorithm uses four OBM estimates: <u>one</u> to get an initial γ_0 estimate, <u>two</u> to estimate optimal batch size, and <u>one</u> to estimate V. The time complexity of this algorithm is $O(n)$.

<u>Algorithm A1 (1–2–1 OBM)</u>

1. Estimate γ_0.

(a) Estimate marginal variance, $R(0)$. Set $m_0 = a_0 n^{1/r_0} + b_0$, where a_0, b_0, and r_0 are constants, $a_0 > 0$, and $r_0 \geq 2$. Suggestion: $m_0 = \sqrt{n}$.

(b) Compute OBM estimate $\widehat{V}_{m_0}$.

(c) Estimate γ_0 from $\widehat{\gamma}_0^{(0)} = \dfrac{n\,\widehat{V}_{m_0}}{\widehat{R}(0)}$.

2. Estimate center of gravity.

(a) Set $m_1 = \max\left\{\mathrm{a}_1\,\mathrm{n}^{1/\mathrm{r}_1} + \mathrm{b}_1\,,\ \mathrm{c}_1 \max\{\widehat{\gamma}_0,\ 1/\widehat{\gamma}_0\} + \mathrm{d}_1\right\}$, where a_1, b_1, c_1, d_1, and r_1 are constants, $a_1 > 0$, $r_1 > 3$, and $c_1 > 0$.

Suggestion: $m_1 = \max\left\{\sqrt[6]{n},\ 0.75\max\{\widehat{\gamma}_0,\ 1/\widehat{\gamma}_0\}\right\}$.

(b) Set $m_2 = m_1 + 1$.

(c) Compute OBM estimates $\widehat{V}_{m_1}$ and $\widehat{V}_{m_2}$.

(d) Compute $\widehat{\gamma}_1 = \dfrac{n\,m_1\,(m_1+1)}{\widehat{R}(0)}\left(\widehat{V}_{m_2} - \widehat{V}_{m_1}\right)$,

and $\widehat{\gamma}_0 = \dfrac{n\,\widehat{V}_{m_1}}{\widehat{R}(0)}$.

3. Estimate V.

(a) Estimate optimal batch size from

$$\widehat{m}^* = \left[1.5\,n\left(\frac{\widehat{\gamma}_1}{\widehat{\gamma}_0}\right)^2\right]^{\frac{1}{3}} + 1. \qquad [7.7]$$

(b) Estimate V using the OBM estimate $\widehat{V}_{\widehat{m}^*}$.

7.5 Evaluation of 1–2–1 OBM Algorithm

7.5.1 Benchmark and Theoretical Properties

We now evaluate the performance of the 1–2–1 OBM algorithm. We describe the Monte Carlo experiments and we discuss the experimental results. We conclude that in spite of its simplicity (in respect to conceptual design and time complexity) the performance of the 1–2–1 OBM is good.

To evaluate the algorithm A1 we need a benchmark. If the sum of correlations γ_0 and the weighted sum of correlations γ_1 were known, the following algorithm could be used to estimate the variance of the sample mean.

Algorithm A0

1. Read true γ_0 and true γ_1.

2. Compute $\vec{m}^*$ (approximate optimal batch size) from Equation 7.6.

3. Estimate V using OBM estimate $\widehat{V}_{\vec{m}^*}$.

In practice, γ_0 and γ_1 are not known. Therefore, "actual" algorithms estimating the optimal batch size are less efficient than A0, but ideally, the performance of a good algorithm is close to the performance of A0, our benchmark.

The 1–2–1 OBM algorithm has good properties, as indicated by Theorem 7.5.

THEOREM 7.5 Assume conditions of Theorem 7.4. Let $\widehat{V}_{\widehat{m}^*}$ and $\widehat{V}_{\vec{m}^*}$ be two OBM estimators of the variance of the sample mean V obtained respectively from algorithms A1 (1–2–1 OBM) and A0. Therefore, for $n \to \infty$,

$$\frac{\widehat{m}^*}{\vec{m}^*} \xrightarrow{\text{P}} 1.$$

Proof: By Theorem 7.2 (MSE-consistency of $\widehat{\gamma}_0$), Theorem 7.4 (MSE-consistency of $\widehat{\gamma}_1$), and Slutsky's Theorem

$$\frac{\widehat{\gamma}_1}{\widehat{\gamma}_0} \xrightarrow{\text{P}} \frac{\gamma_1}{\gamma_0}.$$

By definition of $\widehat{m}^*$ (Equation 7.7) and $\vec{m}^*$ (Equation 7.6) the result follows.

Theorem 7.5 suggests that for large sample sizes the performance of the 1–2–1 OBM algorithm and the A0 algorithm are the same, i.e.,

$$\frac{\text{mse}(\widehat{V}_{\widehat{m}^*})}{V^2} \simeq \frac{\text{mse}(\widehat{V}_{\tilde{m}^*})}{V^2}.$$

By Theorem 7.1, mse $\left(\widehat{V}_{\tilde{m}^*}\right)/V^2 \xrightarrow{n \to \infty} 0$. Therefore Theorem 7.5 also suggests that the 1–2–1 OBM algorithm is MSE-consistent, i.e.,

$$\frac{\text{mse}(\widehat{V}_{\widehat{m}^*})}{V^2} \xrightarrow{n \to \infty} 0.$$

7.5.2 The Monte Carlo Experiment

The Monte Carlo experiment estimates the standardized MSE (i.e., mse$(\widehat{V})/V^2$) as a function of $|\gamma_1|/\gamma_0$ for twelve cases: four sample sizes and three steady-state data processes. The sample sizes are $n = 1000, 4000, 16000, 64000$; the three data processes are AR(1) normal, 5-state DPSS [Pedrosa and Schmeiser, 1993a], and M/M/1 queue waiting time (QWT). We review these three processes in Chapters 2 and 3. For each sample $\{X_1, X_2, \ldots, X_n\}$ from the data process, the 1–2–1 OBM algorithm yields an estimate $\widehat{V}_{\widehat{m}^*}$ of the variance of the sample mean and algorithm A0 yields an estimate $\widehat{V}_{\tilde{m}^*}$. For each $|\gamma_1|/\gamma_0$, 500 such samples are generated to estimate the standardized MSE to negligible sampling error. The standardized MSE is computed for $\gamma_1/\gamma_0 = 0.375, 0.75, 1.5, \ldots, 96$ in the case of AR(1) and M/M/1, and it is computed for $|\gamma_1|/\gamma_0 = 3, 6, 12, \ldots, 96$ in the case of 5-state DPSS, since this process only yields $\gamma_1/\gamma_0 \leq -2.38$. The 1–2–1 OBM algorithm uses the values suggested in Section 7.4.3, i.e., $m_0 = \sqrt{n}$ and $m_1 = \max\{0.75\,\widehat{\gamma}_0,\ \sqrt[4]{n}\}$.

The values of the process mean μ and the process variance $R(0)$ are irrelevant in this set of experiments, because (1) OBM estimators are invariant in respect to the process mean, and (2) the standardized MSE is unscaled. For simplicity, the mean is always $\mu = 0$, and the marginal variance is always set so that the variance of the sample mean is $V = 1$.

7.5.3 Discussion of Experimental Results

The experimental results indicate that the performance of the 1–2–1 OBM algorithm (A1) is close to the performance of the idealistic A0.

To aid in the discussion, we introduce two figures, each containing twelve charts displayed in four rows and three columns. Each column corresponds to different data processes (AR(1) normal, 5-state DPSS, and M/M/1-QWT) and each row corresponds to a different sample size n (1000, 4000, 16000, and 64000). The horizontal axis is $|\gamma_1|/\gamma_0$ and the vertical axis is the standardized MSE. For each chart, two curves are shown: the thicker represents the performance of A1 and the thinner represents the performance of A0. The less is the height of the curve the better is the performance of the corresponding algorithm.

Figure 7.1 indicates that the performance of A1 is close to the performance of A0, even for small sample sizes or for large $|\gamma_1|/\gamma_0$. Other conclusions suggested by the figure also support the theoretical conclusions of Section 7.5.1 and include:

1. A1 is a consistent estimator of the variance of the sample mean since $\text{mse}(\widehat{V})/V^2 \to 0$ when $n \to \infty$.

2. The performance of A1 converges to the performance of A0 when the sample size increases or equivalently when $|\gamma_1|/\gamma_0$ decreases.

3. Of the three data processes, the absolute performance of both A1 and A0 is best for AR(1) and worst for M/M/1.

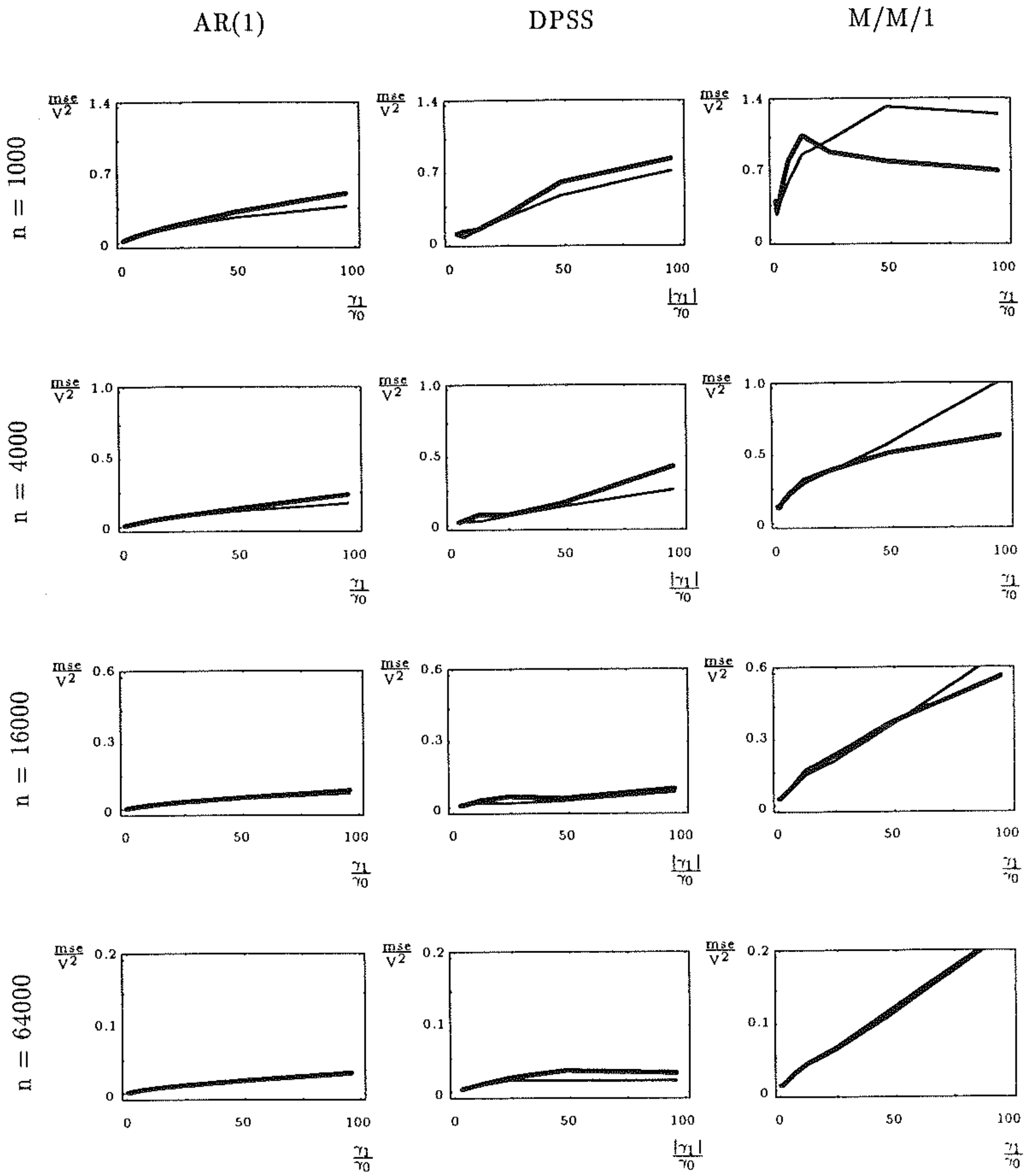

Figure 7.1 Standardized MSE curves for algorithms A0 and A1.

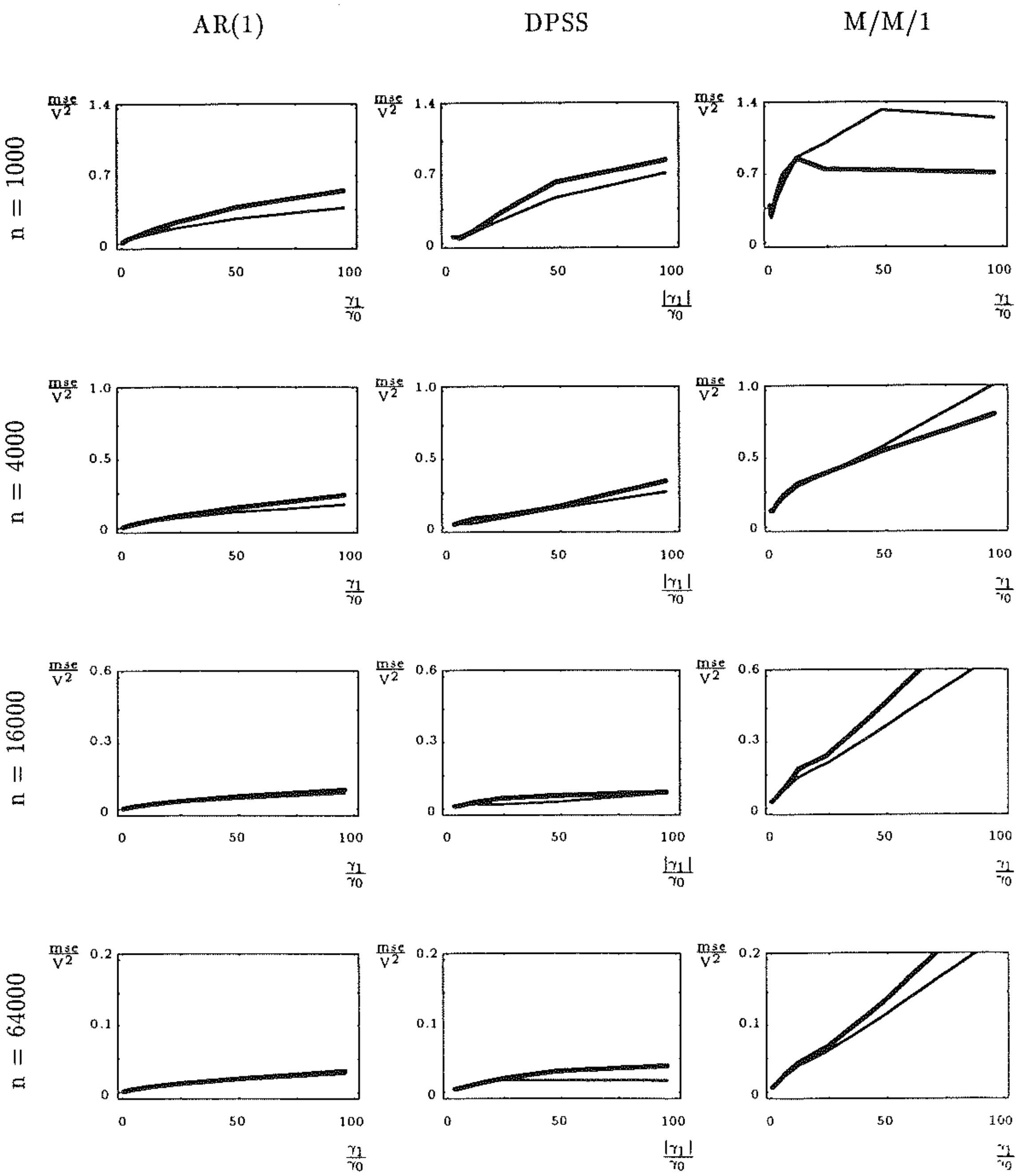

Figure 7.2 Standardized MSE curves for algorithms A0 and A2.

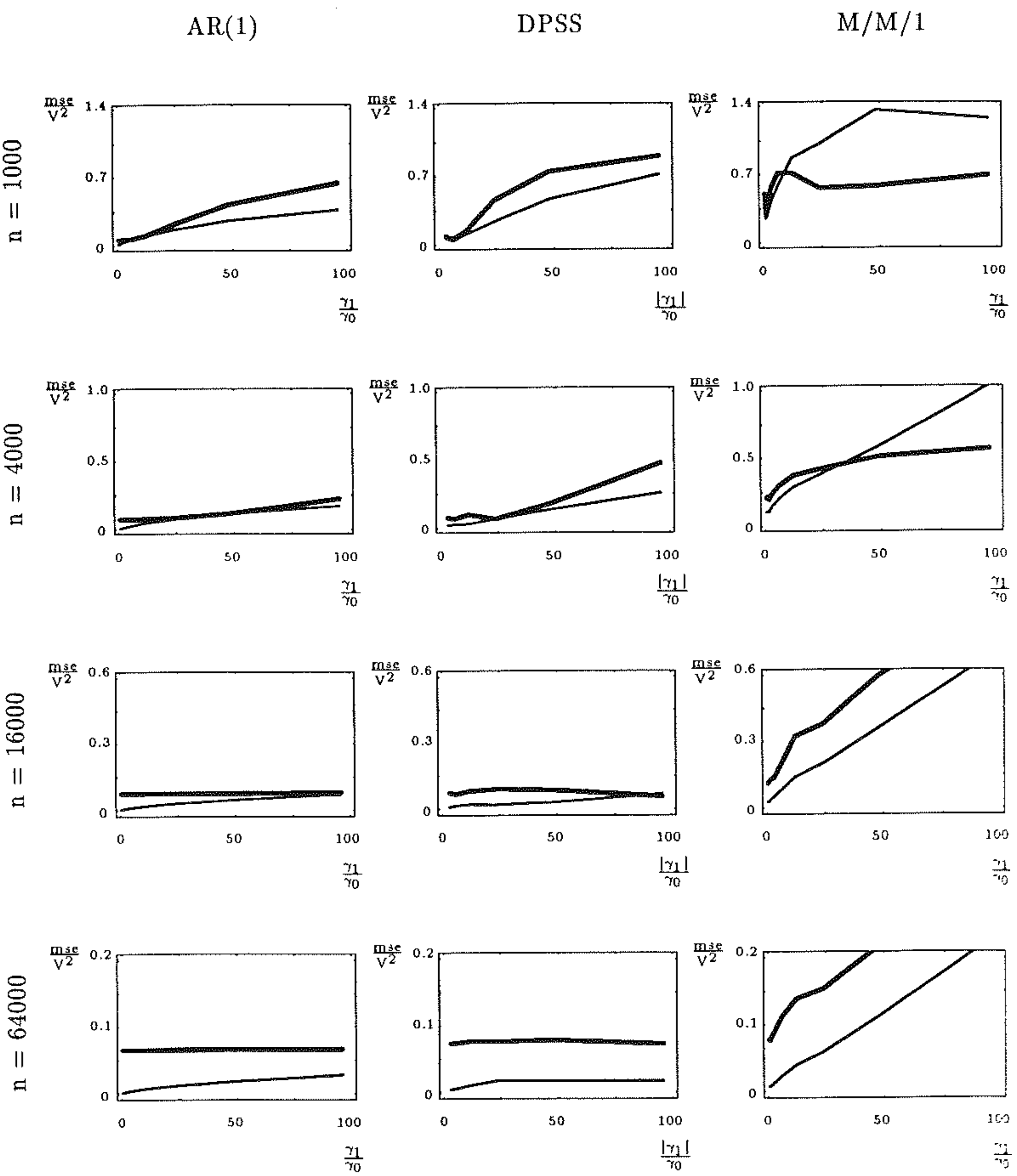

Figure 7.3 Standardized MSE curves for algorithms A0 and A3.

- DPSS has an oscillating correlation structure which adds on a non-desirable effect (increases MSE). This effect becomes more important when the sample size decreases or when $|\gamma_1|/\gamma_0$ increases.

- M/M/1 has heavy tails and therefore the sample sizes to obtain good estimates must be larger. For this type of data, a sample size $n = 1000$ is so small that neither $\bar{X}$, nor $\widehat{V}_{m^*}$, nor $\widehat{V}_{\tilde{m}^*}$ work well.

7.5.4 Comparison with Other Algorithms

We analyze the performance of two other algorithms and compare it with the performance of the 1–2–1 OBM. The Monte Carlo experimental conditions are described in Section 7.5.2.

We discuss in Section 7.4 that a natural way to estimate γ_0 and γ_1 and then the asymptotic approximate batch size is to estimate the individual autocorrelations. Song's method [1993] is based on this approach. The method, which we denote by A2, consists of (1) estimating the number of sample autocorrelations that are not negligible (m_0) and, simultaneously, calculating the sample autocorrelations; (2) estimating γ_0 and γ_1 by substituting the calculated sample autocorrelations; and (3) estimating the optimal batch size using Equation 7.6. A2 is an iterative process and m_0 is the smallest integer such that $|\widehat{\rho}(h)| < p\,\sigma(\widehat{\rho}(h))$ for $h = m_0 + 1, \ldots, m_0 + q$, where $\widehat{\rho}(h)$ is the estimated lag-h autocorrelation, p and q are constants, and $\widehat{\sigma}(\widehat{\rho}(h)) = [n^{-1}\ (1 + 2\sum_{i=1}^{m_0} \widehat{\rho}(i)^2)]^{1/2}$ estimates the standard error of $\widehat{\rho}(h)$. In this set of experiments we use $p = 3$ and $q = 6$, as suggested by Song.

The other algorithm is A3, which uses one OBM estimate with batch size $m = n/d$ where d, the number of batches, is fixed. In this set of experiments we use $d = 20$, as suggested by Schmeiser [1982] to obtain good confidence-interval behavior. A3 is not MSE-consistent, in the sense defined in Section 7.3.1, because m/n is constant.

Figures 7.2 and 7.3 show the MSE performance curves of, respectively, A2 and A3. Comparing the charts displayed in each of these figures with the corresponding charts of Figure 7.1 (1–2–1 OBM) we conclude that:

1. A1 (1–2–1 OBM) and A2 (Song's algorithm) have similar MSE performance for AR(1) and DPSS data, but A1 has better MSE performance for M/M/1 data.

2. A3 has the worst MSE performance of all algorithms analyzed in this chapter. It is not MSE-consistent, which was expected since the number of batches is fixed.

In terms of time complexity A1 and A3 are $O(n)$ while A2 is $O(m_0 n)$, where m_0 depends on the correlation structure of the data. Therefore, A1 is computationally more efficient than A2.

8. REGRESSION OF OVERLAPPING BATCH MEANS

8.1 Introduction

Another approach to develop computationally and statistically efficient methodology to estimate the variance of the sample mean is, as discussed in Chapter 4, to use linear combinations. There are several examples in the literature of using linear combinations of classical estimators to obtain a better estimator in terms of statistical properties: variance, bias or mean squared error. For example, Schruben [1983] considered a linear combination of standardized time series area (STS.A) and non-overlapping batch means (NBM) estimators, which are asymptotically independent. Schmeiser and Song [1989] discuss the problem of selecting the component estimators and determining the optimal linear combination coefficients given the covariances between component estimators and their individual bias. Politis and Romano [1994], in the context of spectral function estimation, propose a linear combination of two Bartlett spectral estimators with different bandwidths for the reduction of the bias. We discuss in this chapter a different direction: using linear regression of overlapping batch means (OBM) estimators.

Regression techniques are well known and understood. OBM estimators are simple and easy to compute. The mathematical models of the bias [Pedrosa and Schmeiser, 1993d] and covariances [Pedrosa and Schmeiser 1993c] of OBM estimators are known.

Therefore, we develop asymptotically unbiased estimators of the variance of the sample mean (V), the sum of autocorrelations (γ_0), and the weighted sum of autocorrelations (γ_1) using regression of overlapping batch means.

This chapter is organized as follows. In Section 8.2, we summarize asymptotic results regarding bias, variance, and covariance of OBM estimators. In Section 8.3, we review classical regression analysis. Finally, in Section 8.4, we develop the linear regression model of OBM estimators and the new asymptotically unbiased estimators $\widehat{V}^{\mathrm{RO}}$, $\widehat{\gamma}_0^{\mathrm{RO}}$, and $\widehat{\gamma}_1^{\mathrm{RO}}$ of V, γ_0, and γ_1, respectively. We also discuss implementation aspects of algorithms using regression of OBM estimators.

The methodology that we develop is based on OBM estimators. Analogous development can be considered for other estimators parameterized by batch size, such as non-overlapping batch means and standardized time series.

8.2 Asymptotic Results

We summarize in this section the asymptotic results of Pedrosa and Schmeiser [1993c and 1993d] that are needed to develop the linear regression model. Let $\widehat{V}_m$ be an OBM estimator with batch size m. The asymptotic ($n \to \infty$, $m \to \infty$, and $m/n \to 0$)

$$\mathrm{bias}(\widehat{V}_m) = -\frac{\gamma_1 R(0)}{m\, n} + O\left(\frac{m}{n^2}\right), \qquad [8.1]$$

and the asymptotic

$$\mathrm{var}(\widehat{V}_m) = \frac{4\, m\, \gamma_0^2\, R(0)^2}{3\, n^3} + O\left(\frac{1}{n^3}\right) + O\left(\frac{m^2}{n^4}\right). \qquad [8.2]$$

The asymptotic ($n \to \infty$, $m_1 \to \infty$, $m_1/n \to 0$, $m_1/m_2 \to$ constant, and $m_1 \le m_2$)

$$\mathrm{cov}\left(\widehat{V}_{m_1}, \widehat{V}_{m_2}\right) = \frac{4\, m_1\, \gamma_0^2\, R(0)^2}{3\, n^3}\left(1 + \frac{m_2 - m_1}{2\, m_2}\right) + O\left(\frac{1}{n^3}\right) + O\left(\frac{m_2^2}{n^4}\right). \qquad [8.3]$$

One set of sufficient conditions that imply Equations 8.1, 8.2, and 8.3 is [Pedrosa and Schmeiser, 1993c] $\{X_i\}$ is a general linear process of the form $X_i = \sum_{h=-\infty}^{+\infty} b_h \, \varepsilon_{i-h}$, where (1) b_h's are constants, (2) $\{\varepsilon_i\}$ is a sequence of iid random variables with finite variance, σ_ε^2, and finite fourth cumulant $k_{\varepsilon,4}$, and (3) $\sum_{h=-\infty}^{+\infty} |h| \, |b_h| < \infty$.

8.3 Review of Classical Linear Regression Analysis

Regression analysis is the statistical methodology for predicting values of one or more "response" (or dependent) variables from a collection of "predictor" (or independent) variables values [Johnson and Wichern, 1988]. The first step on using regression analysis is to choose the model. The linear model is the most studied and used [Searle, 1971]. It is mathematically tractable and many models that are apparently non-linear can be rearranged in a linear form. Here we summarize results for the classical linear regression model. This review is based mostly on Johnson and Wichern [1988] and on Searle [1971].

Suppose that the response Y depends linearly on r predictor variables $x_1, \ldots, x_r$ and on a random error ε (which includes the measurement error and the effects of other variables not considered). Suppose that set of p observations (on Y and associated x_j's) are independent. Assume that the random error has mean zero and variance σ^2. Assume also that the random errors of observations i and j are uncorrelated for any i and j. The classical linear regression model is

$$\begin{aligned} \underline{Y} &= \underline{X} \, \underline{b} + \underline{\varepsilon}, \\ \mathrm{E}(\underline{\varepsilon}) &= \underline{0}, \\ \mathrm{cov}(\underline{\varepsilon}) &= \sigma^2 \, \underline{I}, \end{aligned} \qquad [8.4]$$

where (1) $\underline{Y}$ and $\underline{\varepsilon}$ are order-p column vectors, $\underline{b}$ is an order-$(r+1)$ column vector, and $\underline{X}$ is an order-$(p*(r+1))$ matrix with ith row $[1, x_{i1}, \ldots, x_{ir}]$; and (2) $\underline{b}$ (regression coefficients) and σ^2 (error variance) are unknown parameters. The method of least squares involves selecting $\widehat{\underline{b}}$ (least square estimator) to minimize the sum of squares of deviations from their expected values, i.e.,

$$\sum_{i=1}^{p} [Y_i - (b_0 + b_1 x_{i1} + \ldots + b_r x_{ir})]^2 = (\underline{Y} - \underline{X}\,\underline{b})^t\,(\underline{Y} - \underline{X}\,\underline{b}) .$$

The least squares estimator of $\underline{b}$ for the classical regression model (Equations 8.4) is

$$\widehat{\underline{b}} = \left(\underline{X}^t\,\underline{X}\right)^{-1} \underline{X}^t\,\underline{Y}, \tag{8.5}$$

and has the properties

$$\mathrm{E}\left(\widehat{\underline{b}}\right) = \underline{b} \tag{8.6}$$

and

$$\mathrm{cov}\left(\widehat{\underline{b}}\right) = \sigma^2 \left(\underline{X}^t\,\underline{X}\right)^{-1} . \tag{8.7}$$

If $\underline{X}$ is not full rank $\left(\underline{X}^t\underline{X}\right)^{-1}$ is repaced by a generalized inverse of $\underline{X}^t\underline{X}$ as discussed in Johnson and Wichern [1988, p. 334, exercise 7.6].

The vector of residuals

$$\widehat{\underline{\varepsilon}} = \underline{Y} - \underline{X}\,\widehat{\underline{b}} = \left[\underline{I} - \underline{X}\left(\underline{X}^t\,\underline{X}\right)^{-1} \underline{X}^t\right] \underline{Y} \tag{8.8}$$

contains the information about the remaining unknown parameter σ^2. The estimator

$$\widehat{\sigma}^2 = \frac{\widehat{\underline{\varepsilon}}^t\,\widehat{\underline{\varepsilon}}}{n - r - 1} \tag{8.9}$$

is an unbiased estimator of the error variance since $\mathrm{E}(\widehat{\sigma}^2) = \sigma^2$.

8.4 Regression of OBM Estimators

We now develop the linear regression model of OBM estimators. By Equation 8.1, for any OBM estimator $\widehat{V}_{m_i}$ with batch size m_i

$$\mathrm{E}(\widehat{V}_{m_i}) \;=\; V \;-\; \frac{\gamma_1\, R(0)}{m_i\, n} \;+\; O\left(\frac{m_i}{n^2}\right), \qquad [8.10]$$

or equivalently

$$\widehat{V}_{m_i} \;=\; V \;-\; \frac{\gamma_1\, R(0)}{m_i\, n} \;+\; O\left(\frac{m_i}{n^2}\right) \;+\; \varepsilon_{m_i}, \qquad [8.11]$$

where ε_{m_i} is a random error with zero mean, i.e., $\mathrm{E}(\varepsilon_{m_i}) = 0$. If a number $p < \infty$ of OBM estimators are computed,

$$\begin{bmatrix} \widehat{V}_{m_1} \\ \widehat{V}_{m_2} \\ \cdots \\ \widehat{V}_{m_p} \end{bmatrix} = \begin{bmatrix} 1 & 1/m_1 & m_1 \\ 1 & 1/m_2 & m_2 \\ \cdots & \cdots & \cdots \\ 1 & 1/m_p & m_p \end{bmatrix} \begin{bmatrix} V \\ -\gamma_1 R(0)/n \\ O\left(1/n^2\right) \end{bmatrix} + \begin{bmatrix} \varepsilon_{m_1} \\ \varepsilon_{m_2} \\ \cdots \\ \varepsilon_{m_p} \end{bmatrix}. \qquad [8.12]$$

To simplify the notation let

$$\underline{\widehat{V}} = \begin{bmatrix} \widehat{V}_{m_1} \\ \widehat{V}_{m_2} \\ \cdots \\ \widehat{V}_{m_p} \end{bmatrix}, \quad \underline{M} = \begin{bmatrix} 1 & 1/m_1 & m_1 \\ 1 & 1/m_2 & m_2 \\ \cdots & \cdots & \cdots \\ 1 & 1/m_p & m_p \end{bmatrix}, \quad \underline{b} = \begin{bmatrix} V \\ -\gamma_1 R(0)/n \\ O\left(1/n^2\right) \end{bmatrix}, \text{and } \underline{\varepsilon} = \begin{bmatrix} \varepsilon_{m_1} \\ \varepsilon_{m_2} \\ \cdots \\ \varepsilon_{m_p} \end{bmatrix}. \qquad [8.13]$$

Equation 8.12 can be rewritten as

$$\underline{\widehat{V}} \;=\; \underline{M}\,\underline{b} \;+\; \underline{\varepsilon}.$$

As discussed earlier, the random errors have zero mean

$$\mathrm{E}\left(\underline{\varepsilon}\right) \;=\; \underline{0}, \qquad [8.14]$$

and because the OBM estimators are correlated the random errors are also correlated,

$$\operatorname{cov}(\underline{\varepsilon}) \quad = \quad \mathrm{E}\left(\underline{\varepsilon}\,\underline{\varepsilon}^t\right) \quad = \quad \operatorname{cov}\left(\underline{\widehat{V}}\right). \qquad [8.15]$$

Lemma 8.1 is a formula to compute $\operatorname{cov}(\underline{\varepsilon})$. It also defines matrix $\underline{\Sigma}$ which is needed in Theorem 8.1.

LEMMA 8.1 Assume that Equations 8.1, 8.2, and 8.3 hold. Let $\underline{\widehat{V}}$, $\underline{b} = [b_1\ b_2\ b_3]$, $\underline{\varepsilon}$, and $\underline{M}$ be the vectors and the matrix defined by Equation 8.13. For $i = 1, \ldots, p$, let batch size $m_i = r_i\, m_1 + s_i$, where r_i and s_i are constants such that $m_1 < m_2 < \ldots < m_p$. Then

$$\operatorname{cov}(\underline{\varepsilon}) \quad = \quad \mathrm{E}\left(\underline{\varepsilon}\,\underline{\varepsilon}^t\right) \quad = \quad \sigma^2\,\underline{\Sigma},$$

where

$$\sigma^2 \quad = \quad \frac{4\, m_1\, \gamma_0^2\, R(0)^2}{3\, n^3}, \qquad [8.16]$$

and

$$\underline{\Sigma} \quad = \quad \left[\left(1 + O\left(\frac{1}{m_1}\right) + O\left(\frac{m_1}{n}\right)\right)\left(1 + \frac{|m_i - m_j|}{2\max(m_i, m_j)}\right)\frac{\min(m_i, m_j)}{m_1}\right]_{i,j=1}^{p}. \qquad [8.17]$$

Proof: The result follows directly from Equations 8.15 and 8.3.

Now, we state Theorem 8.1, which provides a new method to estimate the variance of the sample mean.

THEOREM 8.1 Assume that Equations 8.1, 8.2, and 8.3 hold. Let $\underline{\widehat{V}}$, $\underline{b} = [b_1\ b_2\ b_3]$, $\underline{\varepsilon}$, and $\underline{M}$ be the vectors and the matrix defined by Equation 8.13. Let σ^2 and $\underline{\Sigma}$ be defined by Equations 8.16 and 8.17, respectively. For $i = 1, \ldots, p$, let batch size $m_i = r_i\, m_1 + s_i$, where r_i and s_i are constants such that $m_1 < m_2 < \ldots < m_p$. Let

$n \to \infty$ and $m_1/n^\alpha \to \beta$, where α and β are constants such that $0 < \alpha < 1$ and $\beta > 0$. Let

$$\underline{\theta}^t = \left(\underline{M}^t\,\underline{\Sigma}^{-1}\,\underline{M}\right)^{-1}\underline{M}^t\,\underline{\Sigma}^{-1}. \qquad [8.18]$$

Then estimator

$$\widehat{\underline{b}} = \underline{\theta}^t\,\widehat{\underline{V}},$$

and has the properties

$$\mathrm{E}\left(\widehat{\underline{\mathrm{b}}}\right) = \underline{\mathrm{b}}$$

and

$$\mathrm{cov}\left(\widehat{\underline{\mathrm{b}}}\right) = \sigma^2\left(\underline{\mathrm{M}}^t\,\underline{\Sigma}^{-1}\,\underline{\mathrm{M}}\right)^{-1}.$$

Proof: Define $\underline{Y}^* = \underline{\Sigma}^{-1/2}\widehat{\underline{V}}$, $\underline{X}^* = \underline{\Sigma}^{-1/2}\underline{M}$, and $\underline{\varepsilon}^* = \underline{\Sigma}^{-1/2}\underline{\varepsilon}$. It follows that $\underline{Y}^* = \underline{X}^*\,\underline{b} + \underline{\varepsilon}^*$ is the classical linear regression form given by Equations 8.4. Then the results for $\underline{Y}^* = \underline{X}^*\,\underline{b} + \underline{\varepsilon}^*$ are obtained directly from the classical model results. The least squares estimator of $\underline{b}$ is

$$\begin{aligned}\widehat{\underline{b}} &= \left(\underline{X}^{*t}\,\underline{X}^*\right)^{-1}\underline{X}^{*t}\,\underline{Y}^* \\ &= \left(\underline{M}^t\,\underline{\Sigma}^{-1}\,\underline{M}\right)^{-1}\underline{M}^t\,\underline{\Sigma}^{-1}\,\widehat{\underline{V}}.\end{aligned}$$

Because $\underline{Y}^* = \underline{X}^*\,\underline{b} + \underline{\varepsilon}^*$ is of the classical linear regression form, the properties of $\underline{b}$ follow from Equations 8.6 and 8.7.

Corollary 8.1 provides an asymptotically unbiased estimator of σ^2.

COROLLARY 8.1 Assume that Theorem 8.1 holds. Let $\widehat{\underline{\varepsilon}}$ be the vector of residuals

$$\widehat{\underline{\varepsilon}} = \widehat{\underline{V}} - \underline{M}\,\widehat{\underline{b}} = \left[\underline{I} - \underline{M}\left(\underline{M}^t\,\underline{\Sigma}^{-1}\,\underline{M}\right)^{-1}\underline{M}^t\,\underline{\Sigma}^{-1}\right]\widehat{\underline{V}}.$$

Then,

$$\widehat{\sigma}^2 = \frac{\widehat{\underline{\varepsilon}}^t\,\underline{\Sigma}^{-1}\,\widehat{\underline{\varepsilon}}}{n - 4}$$

is an asymptotically unbiased estimator of σ^2.

Proof: Use the same technique as in Theorem 8.1, i.e., define $\underline{Y}^* = \underline{\Sigma}^{-1/2}\underline{\widehat{V}}$, $\underline{X}^* = \underline{\Sigma}^{-1/2}\underline{M}$, and $\underline{\varepsilon}^* = \underline{\Sigma}^{-1/2}\underline{\varepsilon}$. Then use the results of the classical linear regression form to prove Corollary 8.1.

Corollary 8.2 provides a formula to estimate cov $\left(\underline{\widehat{b}}\right)$ and follows directly from Theorem 8.1 and Corollary 8.1.

COROLLARY 8.2 Assume that Corollary 8.1 holds. Then

$$\widehat{\text{cov}}\left(\underline{\widehat{b}}\right) = \widehat{\sigma}^2 \left(\underline{M}^t\, \underline{\Sigma}^{-1}\, \underline{M}\right)^{-1}$$

is an asymptotically unbiased estimator of cov $\left(\underline{\widehat{b}}\right)$.

Lemma 8.1, Theorem 8.1, and Corollaries 8.1 and 8.2 provide the theoretical background to develop estimators $\widehat{V}^{\text{RO}}$, $\widehat{\gamma}_0^{\text{RO}}$, and $\widehat{\gamma}_1^{\text{RO}}$ of V, γ_0, and γ_1, respectively. For example, by Equation 8.13, $\widehat{V}^{\text{RO}} = \widehat{b}_1$, is an asymptotically unbiased estimator of V. Notice that $\widehat{V}^{\text{RO}}$ is a linear combination of OBM estimators: From Theorem 8.1

$$\widehat{V}^{\text{RO}} = \sum_{i=1}^{p} \theta_{i1}\, \widehat{V}_{m_i}, \qquad [8.19]$$

where θ_{i1} is the $(i, 1)$ element of matrix $\underline{\theta}$. The variance of estimator $\widehat{V}^{\text{RO}}$ can be estimated using Corollary 8.2.

The sum of autocorrelations γ_0 can be estimated using $\widehat{\sigma}^2$ defined in Corollary 8.1. By Equation 8.16, $\text{E}(\widehat{\sigma}^2) = (4m_1\gamma_0^2 R(0)^2)/(3n^3)$. Therefore, an estimator $\widehat{\gamma}_0^{\text{RO}}$ of γ_0 is

$$\widehat{\gamma}_0^{\text{RO}} = \sqrt{\frac{3\, n^3\, \widehat{\sigma}^2}{4\, m_1\, \widehat{R}(0)^2}}, \qquad [8.20]$$

where $\widehat{R}(0)$ is the sample variance.

Similarly, the weighted sum of autocorrelations γ_1 can be estimated from $\widehat{b}_2$. By Equation 8.13 and Theorem 8.1, $\mathrm{E}(\widehat{b}_2) = -\gamma_1 R(0)/n$. Therefore an estimator $\widehat{\gamma}_1^{\mathrm{RO}}$ of γ_1 is

$$\widehat{\gamma}_1^{\mathrm{RO}} = -\frac{n\,\widehat{b}_2}{\widehat{R}(0)}, \qquad [8.21]$$

where $\widehat{R}(0)$ is, again, the sample variance.

Theorem 8.1 and Corollary 8.1 suggest good statistical properties for estimators $\widehat{V}^{\mathrm{RO}}$, $\widehat{\gamma}_0^{\mathrm{RO}}$, and $\widehat{\gamma}_1^{\mathrm{RO}}$. For example, they are asymptotically unbiased. For finite samples, and depending on the equivalent number of independent observations, the bias may not be negligible. Monte Carlo experimentation is needed to study the bias behavior for finite samples. The asymptotic variance of estimators $\widehat{V}^{\mathrm{RO}}$ and $\widehat{\gamma}_1^{\mathrm{RO}}$ can be studied from result of Corollary 8.2.

We now focus on the the analysis of the variance of $\widehat{V}^{\mathrm{RO}}$, only. We discuss the results of a numerical study of var $\left(\widehat{V}^{\mathrm{RO}}\right)$ as a function of batch-size selection policy and number of OBM estimators used. These results are shown in Table 8.1. In column m_i of Table 8.1, we define the three different batch-size selection policies analyzed. For example, the first is $m_2 = 2m_1$, $m_3 = 3m_1$, $\ldots$, $m_p = pm_1$. For each batch-size selection policy we consider seven different values of p (number of OBM estimators): 3, 4, 5, 6, 10, 20, and 50. In column var $\left(\widehat{V}^{\mathrm{RO}}\right)/\mathrm{var}\left(\widehat{V}_{m_1}\right)$ we present the relationship between the variance of new estimator $\widehat{V}^{\mathrm{RO}}$ and the variance of OBM component estimator with smaller batch size (therefore, OBM component estimator with smaller variance).

The numerical results shown in Table 8.1 indicate that

1. The reduction of variance obtained by increasing the number of OBM estimators beyond 6 may not compensate the extra computational effort required;

m_i	p	$\frac{\mathrm{var}(\widehat{V}^{RO})}{\mathrm{var}(\widehat{V}_{m_1})}$
$i * m_1$	3	7.00
	4	5.98
	5	5.31
	6	4.99
	10	4.46
	30	4.04
	50	3.97
$2^{i-1} * m_1$	3	6.00
	4	4.86
	5	4.31
	6	4.11
	10	3.92
	30	3.91
	50	3.91
$[1 + 0.5(i-1)]\, m_1$	3	8.50
	4	6.74
	5	5.72
	6	5.21
	10	4.37
	30	3.71
	50	3.60

Table 8.1 Asymptotic variance of $\widehat{V}^{RO}$ as a function of batch-size selection policy and number of OBM estimators

2. The second policy of batch-size selection (i.e., $m_i = 2^{i-1}m_1$) is the best among the three analyzed;

3. Conclusions (1) and (2) imply that var $\left(\widehat{V}^{\mathrm{RO}}\right)$ is approximately four times larger than var $\left(\widehat{V}_{m_1}\right)$;

4. Because the asymptotic variance is proportional to batch size (Equation 8.2), var $\left(\widehat{V}^{\mathrm{RO}}\right)$ is smaller than var $\left(\widehat{V}_{m^*}\right)$ (where m^* is the optimal OBM batch size) only if $m_1 < m^*/4$.

Therefore, if the asymptotic results are still valid for finite sample size n and for $m_1 < m^*/4$, the regression approach may lead to methods statistically better than those based on estimating optimal batch size. Preliminary Monte Carlo experimentation suggests that none of the two approaches dominates the other. The relative statistical performance seems to depend on the sample size n and on the autocorrelation structure of the data. Further investigation is needed. Hybrid algorithms (combining the regression and optimal batch size approaches) may be statistically the most efficient.

Finally, we discuss the computational efficiency of regression based algorithms estimating V. Conclusion (1) implies that these algorithms are $O(n)$. In addition, the coefficients θ_{i1} of Equation 8.19 depend only on the number of OBM estimators and on their batch sizes, which are defined when the algorithms are developed. Therefore, coefficients θ_{i1} can be computed before algorithms are used, avoiding extra computational effort.

9. SUMMARY, CONCLUSIONS, AND RECOMMENDATIONS

A summary of the results of the research is given in Section 9.1, followed by conclusions in Section 9.2, and recommendations for future research in Section 9.3.

9.1 Summary of Results

Results of this research are summarized here. They are listed in the same order as presented in the body of this dissertation.

Chapter 2. Background.

1. Derivation of an approximation formula to compute the weighted sum of autocorrelations of the M/M/1-QWT process (Equation 2.35, Section 2.7).

Chapter 3. DPSS: The d-State Bernoulli-Demand (s, S)-Inventory Markov Chain.

1. Development of a mathematical model of a general process based on the (s, S)-Inventory Markov Chain with demand given by any discrete distribution (Section 3.2).
2. Specialization to Bernoulli demands: the DPSS process (Section 3.3).
   (a) Development of the mathematical model (Section 3.3).
   (b) Derivation of the steady-state distribution, mean, variance, lag-h autocovariance, and lag-h autocorrelation (Section 3.3.3).

(c) Derivation of the sum of autocorrelations (Theorem 3.8, Section 3.3.3).

(d) Derivation of the weighted sum of autocorrelations (Theorem 3.9, Section 3.3.3).

(e) Derivation of formulas to compute the four parameters of the DPSS process as functions of desired statistical properties (Section 3.4).

(f) Development of FORTRAN code to generate a DPSS pseudo-random variate (Appendix B).

Chapter 4. Covariances and Correlations of Overlapping Batch Means and Bartlett Estimators.

1. Derivation of the asymptotic covariance between two OBM, two Bartlett, or one OBM and one Bartlett estimator of the variance of the sample mean (Theorem 4.1, Section 4.4).

2. Derivation of the asymptotic variance of OBM and Bartlett estimators (Corollary 4.1, Section 4.4). A similar result for OBM estimators was used earlier as a conjecture.

3. Derivation of the asymptotic correlation between two OBM, two Bartlett, or one OBM and one Bartlett estimator of the variance of the sample mean (Corollary 4.2, Section 4.4).

4. Development of Monte Carlo experimentation to study the applicability of the asymptotic formulas for finite samples (Section 4.5). The experimental results indicate that the quality of the approximation provided by the asymptotic formulas is good if batch sizes are roughly in the range $(\gamma_0, n/2)$.

Chapter 5. Covariance between Bartlett Estimators of the Spectral Density.

1. Derivation of the asymptotic covariance between Bartlett estimators of the spectral density function (Theorem 5.1, Section 5.3).

2. Derivation of the asymptotic variance of Bartlett estimators of the spectral density function (Corollary 5.1, Section 5.3). Priestley [1992] derives a similar result using a different technique.

3. Derivation of the asymptotic correlation between Bartlett estimators of the spectral density function (Corollary 5.2, Section 5.3).

Chapter 6. Equivalence of Overlapping Batch Means and Bartlett Estimators.

1. Analytical and numerical comparison of OBM and simplified Bartlett estimators (Section 6.2).

2. Derivation of the expected value of OBM estimators (Theorem 6.1, Section 6.3).

3. Derivation of the bias of OBM estimators (Corollary 6.1, Section 6.3). A similar result was used earlier as a conjecture.

4. Derivation of the expected value of Bartlett estimators (Theorem 6.2, Section 6.3).

5. Derivation of the bias of Bartlett estimators (Corollary 6.2, Section 6.3). Politis and Romano [1993] derive a similar result in the context of spectral density estimation.

6. Derivation of the correlation between OBM and Bartlett estimators with the same batch/window size (Theorem 6.3, Section 6.4).

7. Characterization of the asymptotic equivalence between OBM and Bartlett estimators (Theorem 6.4, Section 6.4).

Chapter 7. Estimating Variance of the Sample Mean: Optimal Batch Size Estimation and 1-2-1 Overlapping Batch Means.

1. Definition of MSE-consistency in the context of the estimation of the variance of the sample mean. Derivation of sufficient conditions providing MSE-consistency for selecting OBM batch size (Theorem 7.1 and Corollary 7.1, Section 7.3.1).

2. Derivation of an MSE-consistent estimator of the sum of autocorrelations (Theorem 7.2, Section 7.3.2).

3. Derivation of a formula relating the weighted sum of autocorrelations with the difference of the expectations of two OBM estimators with two different batch sizes (Theorem 7.3, Section 7.3.3).

4. Derivation of an estimator of the weighted sum of autocorrelations and characterization of the conditions for this estimator to be MSE-consistent (Theorem 7.4, Section 7.3.3).

5. Derivation of a guideline for selecting OBM batch sizes to estimate the weighted sum of autocorrelations (Corollary 7.2, Section 7.3.3).

6. Development of the 1–2–1 OBM algorithm to estimate the variance of the sample mean (Section 7.4.3).

7. Derivation of the convergence in probability of the 1–2–1 OBM batch size to the approximate optimal batch size, as defined by Equation 7.6 (Theorem 7.5, Section 7.5.1).

8. Development of Monte Carlo experimentation to evaluate the finite-sample performance of the 1–2–1 OBM algorithm (Sections 7.5.2 and 7.5.3).

9. Development of a FORTRAN implementation of the 1–2–1 OBM algorithm (Appendix C).

Chapter 8. Regression of Overlapping Batch Means

1. Development of the linear regression model of OBM estimators (Section 8.4).

2. Derivation of asymptotically unbiased estimators $\widehat{V}^{\mathrm{RO}}$, $\widehat{\gamma}_0^{\mathrm{RO}}$, and $\widehat{\gamma}_1^{\mathrm{RO}}$ of V, γ_0, and γ_1, respectively (Theorem 8.1 and Corollaries 8.1 and 8.2, Section 8.4).

3. Discussion of implementation aspects of regression-based algorithms (Section 8.4).

9.2 Conclusions

Conclusions from this research are presented here.

1. This research is directed toward developing methodology to estimate the variance of the sample mean with four desired properties: (i) automatic, (ii) robust, and (iii) computationally and (iv) statistically efficient.

   (a) The 1–2–1 OBM algorithm described in Chapter 7 satisfies these four desirable properties. The user only has to provide n observations from any stationary data process and the 1–2–1 OBM algorithm estimates the variance of the sample mean in $O(n)$ units of time with good MSE performance.

   (b) Let us emphasize that (i) in respect to computational efficiency, no reasonable algorithm can estimate the variance of the sample mean in less than

$O(n)$ units of time; (ii) in respect to statistical efficiency, we compare the MSE performance of the 1–2–1 OBM algorithm with the corresponding idealistic performance that could be obtained if the data sum of autocorrelations and the data weighted sum of autocorrelations were known; and (iii) the algorithm can be implemented easily and immediately by any user (we provide a FORTRAN implementation in Appendix C).

(c) The overlapping-batch-means estimator is the basis for most of the methodology developed in this research. Analogous development can be considered for other estimators parameterized by batch size, e.g., non-overlapping batch means and standardized time series.

(d) The structure of the 1–2–1 OBM algorithm is not the only possible. Other algorithms with similar computational and statistical performance can be developed based on the theoretical guidelines that we derive.

2. This research goes beyond the estimation of the variance of the sample mean and involves more general theoretical and practical guidelines, as discussed in Chapters 4 through 8. For example, (i) the results of Chapter 5 are applicable in the estimation of the spectral density function; (ii) the results of Chapters 4 and 6 are applicable in the estimation of properties like the sum of autocorrelations and weighted sum of autocorrelations, as shown in Chapters 7 and 8.

3. In addition, we also develop in Chapter 3 the DPSS family of stochastic processes for evaluating simulation methods. Because (i) DPSS has a correlation structure more complex than, for example, AR(1) processes; (ii) DPSS is simpler than, for example, ARMA(p,q) processes; (iii) the fundamental statistical properties (for evaluation of output analysis methods) of the DPSS process are

related to its four parameters through simple closed-form equations; (iv) the DPSS marginal distribution is discrete uniform; and (v) DPSS variates are, therefore, easily and time-efficiently generated, the DPSS family of stochastic processes proved useful in the evaluation of output analysis methods.

9.3 Recommendations for Future Research

This work has laid some foundations on which future research can be carried out. Some interesting research directions include the following.

1. Developing algorithms to estimate the variance of the sample mean using the regression-based methodology that we develop in Chapter 8. This methodology is asymptotically unbiased and potentially may lead to better MSE performance than the idealistic performance that could be obtained by one OBM estimator (or other estimator parameterized batch size) using the optimal batch size.

2. Generalizing the regression based methodology to non-zero frequencies, by using Bartlett spectral estimators rather than OBM estimators.

3. Developing methodology to estimate the number of observations needed to provide a "good" estimate of the variance of the sample mean. Defining "good" is central to develop this idea, but it is not easy because there are different measures of goodness involving, for example, moments (bias, variance, and MSE) and confidence-interval procedures (coverage probability, expected width, and variance of the width). One possible direction is (i) to define the goodness of the estimates as a function of their MSE (this implies the need to relate the MSE

to the other measures of goodness); and (ii) to compute the number of observations needed to obtain a desired MSE performance from the MSE asymptotic formula.

4. Developing methodology analogous to the methodology developed in this dissertation, but based on other estimators parameterized by batch size. For example, from a computational point of view, NBM estimators may have advantages over OBM estimators. In particular, NBM is sometimes computationally more efficient than OBM when standard errors are computed for increasing values of n, such as occurs when using sequential stopping rules.

LIST OF REFERENCES

LIST OF REFERENCES

[1] ADAM, N. 1983. Achieving a confidence interval for parameters estimated by simulation. *Management Science*, 29, 856–866.

[2] ANDERSON, T. W. 1971. *The Statistical Analysis of Time Series.* John Wiley & Sons, New York.

[3] BHAT, U. N. AND LAL, R. 1990. A sequential inspection plan for Markov dependent production processes. *IIE Transactions*, 22, 56–64.

[4] BICKEL, P. J., AND DOKSUM, K. 1977. *Mathematical Statistics: Basic Ideas and Selected Topics.* Holden-Day, San Francisco, California.

[5] BRATLEY, P., FOX, B. L., AND SCHRAGE, L. E. 1987. *A Guide to Simulation.* Springer-Verlag, New York.

[6] BROCKWELL, P. J., AND DAVIS, R. A. 1991. *Time Series: Theory and Methods.* Springer-Verlag, New York.

[7] CASELLA, G., AND BERGER, R. L. 1990. *Statistical Inference.* Wadsworth, Belmont, California.

[8] CHUN, Y., AND KELTON, W. D. 1988. Time series corrections to batch means variance estimators. *ORSA/TIMS Bulletin*, Number 26, 59.

[9] CRANE, M. A., AND IGLEHART, D. L. 1975. Simulating stable stochastic systems, III: regenerative processes and discrete-event simulations. *Operations Research*, 23, 33–45.

[10] CRANE, M. A., AND LEMOINE, A. J. 1977. *An Introduction to the Regenerative Method for Simulation Analysis.* Lecture Notes in Control and Information Sciences, vol. 4. Springer-Verlag, New York.

[11] DALEY, D. J. 1968. The serial correlation coefficients of waiting times in a stationary single server queue. *Journal of Australian Mathematical Society*, 8, 683–699.

[12] DAMERDJI, H. 1991. Mean-square consistency of the variance estimator in steady-state simulation output analysis. Research Memorandum 91-17, School of Industrial Engineering, Purdue University, West Lafayette, Indiana.

[13] DEVROYE, L. 1986. *Non-Uniform Random Variate Generation.* Springer-Verlag, New York.

[14] DEWALD, L. S., LEWIS, P. A. W., AND MCKENZIE, E. 1989. A bivariate first-order autoregressive time series model in exponential variables (BEAR(1)). *Management Science*, 35, 1236–1246.

[15] FISHMAN, G. S. 1978. Grouping observations in digital simulations. *Management Science*, 24, 510–521.

[16] FOLEY, R. D., AND GOLDSMAN, D. 1988. Confidence intervals using orthonormally weighted standardized time series. In *Proceedings of the 1988 Winter Simulation Conference.* IEEE, New York, 422–424.

[17] FOX, B. L., GOLDSMAN, D., AND SWAIN, J. J. 1991. Spaced batch means. *Operations Research Letters*, 10, 255–263.

[18] GLYNN, P. W., AND IGLEHART, D. L. 1986. Estimation of steady-state central moments by the regenerative method of simulation. *Operations Research Letters*, 5, 271–276.

[19] GLYNN, P. W., AND IGLEHART, D. L. 1990. Simulation output analysis using standardized time series. *Mathematics of Operations Research*, 15, 1–16.

[20] GOLDSMAN, D., KANG, K., AND SARGENT, R. G. 1986. Large and small sample comparisons of various variance estimators. In *Proceedings of the 1986 Winter Simulation Conference.* IEEE, New York, 278–284.

[21] GOLDSMAN, D., KANG, K., AND SEILA, A. F. 1993. Cramér-von Mises variance estimators for simulations. Technical Report, School of Industrial and Systems Engineering, Georgia Institute of Technology, Atlanta, Georgia.

[22] GOLDSMAN, D., AND 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, Georgia.

[23] GOLDSMAN, D., AND SCHRUBEN, L. 1984. Asymptotic properties of some confidence interval estimators for simulation output. *Management Science*, 30, 1217–1225.

[24] GRASSMANN, W. K., TAKSAR, M. I., AND HEYMAN, D. P. 1985. Regenerative analysis and steady state distributions for Markov chains. *Operations Research*, 33, 1107–1116.

[25] HANNAN, E. J. 1957. The variance of the mean of a stationary process. *Journal of the Royal Statistical Society*, B 19, 282–285.

[26] HEIDELBERGER, P., AND WELCH, P. D. 1981. A spectral method for confidence interval generation and run length control in simulation. *Communications of the ACM*, 24, 233–245.

[27] HOEL, P. G., PORT, S. C., AND STONE, C. J. 1972. *Introduction to Stochastic Processes*, Waveland Press, Prospect Heights, Illinois.

[28] JOHNSON, R. A., AND WICHERN, D. W. 1988. *Applied Multivariate Statistical Analysis*. Prentice Hall, Englewood Cliffs, New Jersey.

[29] KANG, K., AND GOLDSMAN, D. 1990. The correlation between mean and variance estimators in computer simulation. *IIE Transactions*, 22, 15–23.

[30] KANG, K., AND SCHMEISER, B. W. 1987. Properties of batch means from stationary arma time series. *Operations Research Letters*, 6, 19–24.

[31] KANG, K., AND SCHMEISER, B. W. 1990. Graphical methods for evaluating and comparing confidence-interval procedures. *Operations Research*, 38, 546–553.

[32] KENDALL, M. G., AND STUART, A. 1969. *The Advanced Theory of Statistics*, volume 1. Hafner Publishing Company, New York.

[33] LAVENBERG, S. S., AND SLUTZ, D. R. 1975. Introduction to regenerative simulation. *IBM Journal of Research Development*, 19, 463–475.

[34] LAW, A. M. 1983. Statistical analysis of simulation output data. *Operations Research*, 31, 983–1029.

[35] LAW, A. M., AND KELTON, W. D. 1991. *Simulation Modeling and Analysis*. McGraw-Hill, New York.

[36] LAWRANCE, A. J. AND LEWIS, P. A. W. 1981. New autoregressive time series model in exponential variables (NEAR(1)). *Adv. Appl. Prob.*, 13, 826–845.

[37] LAWRANCE, A. J. AND LEWIS, P. A. W. 1982. A mixed exponential time series model. *Management Science*, 28, 1045–1053.

[38] LEHMANN, E. L. 1991. *Theory of Point Estimation*. Wadsworth, Belmont, California.

[39] LEWIS, P. A. W., AND ORAV, E. J. 1989. *Simulation Methodology for Statisticians, Operations Analysts, and Engineers*, vol. 1. Wadsworth, Belmont, California.

[40] LEWIS, P. A. W., MCKENZIE, E., AND HUGUS, D. K. 1989. Gamma processes. *Communications in Statistics — Stochastic Models*, 5, 1–30.

[41] LOEVE, M. 1977. *Probability Theory*, vols. 1 and 2. Springer-Verlag, New York, New York.

[42] MEKETON, M. S. 1980. *The Variance Time Curve: Theory, Estimation and Application.* Doctoral Dissertation, School of Operations Research and Industrial Engineering, Cornell University, Ithaca, New York.

[43] MEKETON, M. S., AND SCHMEISER, B. W. 1984. Overlapping batch means: Something for nothing? In *Proceedings of the 1984 Winter Simulation Conference.* IEEE, New York, 227–230.

[44] MELAMED, B. 1991. TES: A class of methods for generating autocorrelated uniform variates. *ORSA Journal on Computing*, 3, 317–329.

[45] MORAN, P. A. P. 1975. The estimation of standard errors in Monte Carlo simulation experiments. *Biometrika*, 62, 1–4.

[46] MUÑOZ, D. 1991. *Cancellation Methods in the Analysis of Simulation Output.* Doctoral Dissertation, Department of Operations Research, Stanford University, Stanford, California.

[47] NEAVE, H. R. 1971. The exact error in spectrum estimates. *The Annals of Mathematical Statistics*, 42, 961–975.

[48] PAPOULIS, A. 1984. *Probability, Random Variables and Stochastic Processes.* McGraw-Hill, Singapore.

[49] PEDROSA, A. C., AND SCHMEISER, B. W. 1993a. DPSS: The d-state Bernoulli-demand (s, S)-inventory Markov chain. Technical Report SMS93-14, School of Industrial Engineering, Purdue University, West Lafayette, Indiana.

[50] PEDROSA, A. C., AND SCHMEISER, B. W. 1993b. Asymptotic and finite-sample correlations between OBM estimators. In *Proceedings of the 1993 Winter Simulation Conference.* IEEE, New York, 481–488.

[51] PEDROSA, A. C., AND SCHMEISER, B. W. 1993c. Covariances and correlations of overlapping batch means and Bartlett estimators. Technical Report SMS93-15, School of Industrial Engineering, Purdue University, West Lafayette, Indiana.

[52] PEDROSA, A. C., AND SCHMEISER, B. W. 1993d. Equivalence of overlapping batch means and Bartlett spectral estimators. Technical Report SMS93-16, School of Industrial Engineering, Purdue University, West Lafayette, Indiana.

[53] PEDROSA, A. C., AND SCHMEISER, B. W. 1994a. Covariance between Bartlett estimators of the spectral density. Technical Report SMS94-1, School of Industrial Engineering, Purdue University, West Lafayette, Indiana.

[54] PEDROSA, A. C., AND SCHMEISER, B. W. 1994b. Estimating variance of the sample mean: Optimal batch size estimation and 1-2-1 overlapping batch means. Technical Report SMS94-3, School of Industrial Engineering, Purdue University, West Lafayette, Indiana.

[55] POLITIS, D. N., AND ROMANO, J. P. 1993. On the sample variance of linear statistics derived from mixing sequences. *Stochastic Processes and their Applications*, 45, 155–167.

[56] POLITIS, D. N., AND ROMANO, J. P. 1994. Bias-corrected nonparametric spectral estimation. Technical Report 94-5, Department of Statistics, Purdue University, West Lafayette, Indiana.

[57] PRIESTLEY, M. B. 1992. *Spectral Analysis and Time Series.* Academic Press, San Diego.

[58] ROSENBLATT, M. 1985. *Stationary Sequences and Random Fields.* Birkhäuser, Boston, Massachusetts.

[59] ROSS, S. M. 1983. *Stochastic Processes*, John Wiley & Sons, New York.

[60] ROSS, S. M. 1989. *Introduction to Probability Models.* Academic Press, San Diego, California.

[61] SARGENT, R. G., KANG, K., AND GOLDSMAN, D. 1992. An investigation of finite-sample behavior of confidence interval estimators. *Operations Research*, 40, 898–913.

[62] SCHMEISER, B. W. 1982. Batch size effects in the analysis of simulation output. *Operations Research*, 30, 556–568.

[63] SCHMEISER, B. W. 1990. Simulation Experiments. In *Stochastic Models.* D.P. Heyman and M. J. Sobel Eds., 295–330. North Holland, New York.

[64] SCHMEISER, B. W., AVRAMIDIS, T. N., AND HASHEM, S. 1990. Overlapping batch statistics. In *Proceedings of the 1990 Winter Simulation Conference.* IEEE, New York, 395–398.

[65] SCHMEISER, B. W., AND LAL, R. 1982. Bivariate gamma random vectors. *Operations Research*, 30, 355–374.

[66] SCHMEISER, B. W., AND SONG, W.-M. T. 1989. Inverse-transformation algorithms for some common stochastic processes. In *Proceedings of the 1989 Winter Simulation Conference.* IEEE, New York, 490-496.

[67] SCHRIBER, T. J., AND ANDREWS, R. W. 1979. Interactive analysis of simulation output by the method of batch means. In *Proceedings of the 1979 Winter Simulation Conference*. ACM, New York, 513–525.

[68] SCHRUBEN, L. W. 1983. Confidence interval estimation using standardized time series. *Operations Research*, 31, 1090–1108.

[69] SEARLE, S. R. 1971. *Linear Models*. John Wiley & Sons, New York.

[70] SEILA, A. F. 1982a. A batching approach to quantile estimation in regenerative simulation. *Management Science*, 28, 573–581.

[71] SEILA, A. F. 1982b. Estimation of percentiles in discrete event simulation. *Simulation*, 39, 193–200.

[72] SONG, W.-M. T. 1988. *Estimators of the Variance of the Sample Mean: Quadratic Forms, Optimal Batch Sizes, and Linear Combinations*. Doctoral Dissertation, School of Industrial Engineering, Purdue University, West Lafayette, Indiana.

[73] SONG, W.-M. T. 1993. On the estimation of optimal batch size in the analysis of simulation output. Technical Report 93-1, Department of Industrial Engineering, National Tsing-Hua University, Republic of China.

[74] SONG, W.-M. T., AND SCHMEISER, B. W. 1988a. Minimal-mse linear combinations of variance estimators of the sample mean. In *Proceedings of the 1988 Winter Simulation Conference*. IEEE, New York, 414–421.

[75] SONG, W.-M. T., AND SCHMEISER, B. W. 1988b. On the dispersion matrix of variance estimators of the sample mean in the analysis of simulation output. *Operations Research Letters*, 7, 259–266.

[76] SONG, W.-M. T., AND SCHMEISER, B. W. 1993. Variance of the sample mean: Properties and graphs of quadratic-form estimators. *Operations Research*, 41, 501–517.

[77] SONG, W.-M. T., AND SCHMEISER, B. W. 1994. Optimal mean-squared-error batch sizes. Technical Report SMS89-3, School of Industrial Engineering, Purdue University, West Lafayette, Indiana. *Management Science*, forthcoming.

[78] WELCH, P. D. 1987. On the relationship between batch means, overlapping batch means and spectral estimation. In *Proceedings of the 1987 Winter Simulation Conference*. IEEE, New York, 320–323.

APPENDICES

Appendix A: Derivation of Quadratic-Form Coefficients of Equation 2.27

We derive here the quadratic-form coefficients of Equation 2.27.

By Equation 2.26,

$$\begin{aligned}\hat{f}(\omega) &= \frac{n}{2\pi}\sum_{r=1}^{n}\sum_{s=1}^{n} q_{rs}\left[X_r X_s - \bar{X}\,(X_r + X_s) + \bar{X}^2\right] e^{-j(r-s)\omega} \\ &= \frac{n}{2\pi}(S_1 - S_2 + S_3),\end{aligned}$$

where

$$\begin{aligned}S_1 &= \sum_{r=1}^{n}\sum_{s=1}^{n} q_{rs}\, X_r X_s\, e^{-j(r-s)\omega}, \\ S_2 &= 2\sum_{r=1}^{n}\sum_{s=1}^{n} q_{rs}\, \bar{X}\, X_r\, e^{-j(r-s)\omega},\end{aligned}$$

and

$$S_3 = \sum_{r=1}^{n}\sum_{s=1}^{n} q_{rs}\, \bar{X}^2\, e^{-j(r-s)\omega}.$$

By definition, $\bar{X} = n^{-1}\sum_{t=1}^{n} X_t$, yielding

$$\begin{aligned}S_2 &= \sum_{r=1}^{n}\sum_{s=1}^{n}\sum_{t=1}^{n} \frac{2}{n}\, q_{rs}\, X_r X_t\, e^{-j(r-s)\omega} \\ &= \sum_{r=1}^{n}\sum_{t=1}^{n} X_r X_t\, e^{-j(r-t)\omega}\left(\frac{2}{n}\sum_{s=1}^{n} q_{rs}\, e^{-j(t-s)\omega}\right).\end{aligned}$$

By Equations 2.26 and 2.17 and because $|e^{-j(t-s)\omega}| = 1$

$$\begin{aligned}S_2 &= \sum_{r=1}^{n}\sum_{t=1}^{n} X_r X_t\, e^{-j(r-t)\omega}\, O\left(\frac{1}{n}\sum_{s=r-m}^{r+m}\frac{1}{n^2}\right) \\ &= \sum_{r=1}^{n}\sum_{t=1}^{n} O\left(\frac{m}{n^3}\right) X_r X_t\, e^{-j(r-t)\omega}.\end{aligned}$$

Letting $s = t$

$$S_2 = \sum_{r=1}^{n}\sum_{s=1}^{n} O\left(\frac{m}{n^3}\right) X_r X_s\, e^{-j(r-s)\omega}.$$

The term S_3 is

$$\begin{aligned} S_3 &= \sum_{r=1}^{n}\sum_{s=1}^{n}\sum_{t=1}^{n}\sum_{u=1}^{n} \frac{1}{n^2}\, q_{rs}\, X_t X_u\, e^{-j(r-s)\omega} \\ &= \sum_{t=1}^{n}\sum_{u=1}^{n} X_t X_u\, e^{-j(t-u)\omega} \left(\frac{1}{n^2} \sum_{r=1}^{n}\sum_{s=1}^{n} q_{rs}\, e^{-j(r-s-t+u)\omega} \right). \end{aligned}$$

By Equation 2.26 and 2.17 and because $|e^{-j(r-s-t+u)\omega}| = 1$

$$\begin{aligned} S_3 &= \sum_{t=1}^{n}\sum_{u=1}^{n} X_t X_u\, e^{-j(t-u)\omega} O\left(\frac{1}{n^2} \sum_{r=1}^{n} \sum_{s=r-m}^{r+m} \frac{1}{n^2} \right) \\ &= \sum_{t=1}^{n}\sum_{u=1}^{n} O\left(\frac{m}{n^3}\right)\, X_t X_u\, e^{-j(t-u)\omega}. \end{aligned}$$

Letting $r = t$ and $s = u$

$$S_3 = \sum_{r=1}^{n}\sum_{s=1}^{n} O\left(\frac{m}{n^3}\right)\, X_r X_s\, e^{-j(r-s)\omega}.$$

Therefore

$$\hat{f}(\omega) = \frac{n}{2\pi} \sum_{r=1}^{n}\sum_{s=1}^{n} \left[q_{rs} + O\left(\frac{m}{n^3}\right) \right] X_r X_s\, e^{-j(r-s)\omega}.$$

By Equation 2.27

$$p_{rs} = q_{rs} + O\left(\frac{m}{n^3}\right).$$

Appendix B: FORTRAN Code for the DPSS Process

B.1 Program to Generate a DPSS Random Variate

```
c     ..........................................................
      subroutine rdpss( id, p, smalls, bigs, iseed, x )
c     ..........................................................
c
c     Antonio Pedrosa and Bruce Schmeiser
c       May 1993
c       Purdue University
c     Purpose: Generate one observation from dpss: the id-state
c       Bernoulli-demand (s,S)-inventory Markov chain.
c
c       The one-step transition probability matrix (for id=5) is
c
c                    state0 state1 state2 state3 state4
c
c                   | p      1-p    0      0      0     |  state0
c                   | 0      p      1-p    0      0     |  state1
c             P =   | 0      0      p      1-p    0     |  state2
c                   | 0      0      0      p      1-p   |  state3
c                   | 1-p    0      0      0      p     |  state4
c
c       The limiting probabilities are
c          P{X=state0} = P{X=state1} = ... = 1/id
c
c     Method:  Inverse transformation.
c     Input:
c         id     : (integer) number of states.              (id > 1)
c         p      : (float)   probability of not changing state.
c         bigs   : (float)   value of state(0).    (S).
c         smalls : (float)   value of state(id-1). (s).     (s <= S)
c         iseed  : (integer) random-number seed.
c         x      : (float)   value of the current state.
c     Output:
c         iseed  : (integer) random-number seed.
c         x      : (float)   value of the generated state.
c     Working variables:
c         delta  : (float)   distance between adjacent state values.
c         istate : (integer) state number of the current state.
c                               istate=0     corresponds to bigs.
c                               istate=id-1  corresponds to smalls.
c     Other routine used:
c         rand   : (float)   uniform (0,1) random-number generator.
c
      implicit real    (a-h, o-z)
      implicit integer (i-n)
c
      u = rand( iseed )
c
```

```
      if ( (x. lt. smalls) .or. (x .gt. bigs) ) then
c         ...generate from the steady-state distribution
          delta  = (bigs - smalls) / (id - 1)
          istate = u * id
      else
c         ...generate from the conditional distribution
          if ( u .le. p ) return
          delta  = (bigs - smalls) / (id - 1)
          istate = ((bigs - x) / delta) + 1.5
      endif
c
      x = bigs - (delta * istate)
      if ( istate .ge. id ) x = bigs
      return
      end
```

B.2 Program to Calculate γ_1

```
c ..................................................................
      subroutine sdpss( id, gam0, gam1, niter, error )
c ..................................................................
c
c Written by:     Antonio Pedrosa and Bruce Schmeiser
c                 School of Industrial Engineering, Purdue University
c                 May 1993
c Language  :     Fortran
c Computer  :     Sun/Sparc workstation
c Purpose   :     Numerically compute the weighted sum of correlations
c                    gamma1 = 2*SUM( h * corr(h) ), h=1,...,infinity
c                 for the d-state Markov chain with Bernoulli transitions.
c Input Parameters:
c     id    :     (integer) number of states.
c     gam0  :     (float)   SUM( corr(h) ), h=1,...,infinity,
c                    which for the d state Markov chain is equal to
c                         = p / (1 - p),
c                    where p is the transition probability from
c                    state i to state i, itself.
c                    For the dpss(5,p) we have     corr(1) = p
c                    For the dpss(2,p) we have     corr(h) = (2p-1)**h.
c Output Parameters:
c     gam1  :     (float) 2*SUM( h * corr(h) ), h=1,...,infinity.
c     niter :     (float) number of iterations used to calculate gamma1.
c     error :     (float) relative error estimate
c                     (based on the error for gamma0)
c
      implicit real    (a-h, o-z)
      implicit integer (i-n)
c
      parameter     (ldim=10)
      data ld       / ldim    /
```

```
      data maxiter / 1000000 /
      data errmax  / 1.d-8   /
c
      double precision t(ldim,ldim), th(ldim,ldim)
      double precision a(ldim,ldim), b(ldim,ldim), c(ldim,ldim), x(ldim)
      double precision errmax, p, corr, cov, cov0
      double precision wgam0, wgam1, error0, error1, bigs
c
c     ...Calculate transition probability from state i to state i.
c
      p = gam0 / (1.d0 + gam0)
c
c     ...Set minimum number of iterations.
c
      miniter = 10. * ( gam0 + 1. / gam0 )
      if (miniter .le. (id * 1000))   miniter = id * 1000
c
c     ...Set the state values centered at zero with delta = 1.
c
      bigs = (id - 1.d0) / 2.d0
      do 1000 i=1,id
         x(i) = bigs - i + 1
 1000 continue
c
c     ...Set t (transition probability matrix).
c
      do 1100 i=1,id
         do 1200 j=1,id
            t(i,j) = 0.d0
 1200    continue
         t(i,i) = p
         k = i+1
         if (i .eq. id) k = 1
         t(i,k) = 1.d0 - p
 1100 continue
c
c     ...Set th = identity matrix.
c
      do 2000 i=1,id
         do 2100 j=1,id
            th(i,j) = 0.d0
 2100    continue
         th(i,i) = 1.d0
 2000 continue
c
c     ...Set lag-0 autocovariance
c
      cov0 = (id**2 - 1.d0) / 12.d0
c
c     ...Initialize for looping.
c
      ih     = 0
```

```
      wgam0  = 1.d0
      wgam1  = 0.d0
      error1 = 1.d10
c
c     ======================
c.....     Start iteration.
c     ======================
c
c     ...Update the iteration number.
c
 3000 ih = ih + 1
c
c     ...Calculate power ih of t (th = t ** ih).  Need only the first
c           row, since all rows are identical, but shifted one position.
c
      call matmul( ld, 1, id, id, th, t, c )
c
      do 3200 j=1,id
         th(1,j) = c(1,j)
 3200 continue
      do 3300 i=2,id
         th(i,1) = th(i-1,id)
         do 3400 j=2,id
            th(i,j) = th(i-1,j-1)
 3400    continue
 3300 continue
c
c    ...Calculate  c = (x) * (t**ih)
c
      do 3500 j=1,id
        a(1,j) = x(j)
 3500 continue
      call matmul( ld, 1, id, id, a, th, c )
c
c     ...Calculate  c = (x) * (t**ih) * (x)
c
      do 3600 j=1,id
         a(1,j) = c(1,j)
         b(j,1) = x(j)
 3600 continue
      call matmul( ld, 1, id, 1, a, b, c )
c
c     ...Compute the lag-ih auto-correlation.
c
      cov  = c(1,1) / id
      corr = cov / cov0
c
c     ...Update SUM[corr(ih)]   (this quantity is the calculated gamma0).
c
      wgam0 = wgam0 + (corr+corr)
c
c     ...Update SUM[(ih)*corr(ih)] (this quantity is the calculated gamma1).
```

```
c
      wgam1 = wgam1 + ((corr+corr) * ih)
c
c     ...Compute maximal absolute error based on gamma0 and on gamma1.
c
      error0 = ih * abs( gam0 - wgam0 )
      error1 = ih * abs( corr )
      if ( error1 .lt. error0 )  error1 = error0
c
c     ...Check whether we can stop iterating.
c
      if ((ih .le. maxiter)    .and.
     &    ((error1 .ge. errmax) .or. (ih .le. miniter))) go to 3000
c
c     ========================
c.....      End iteration.
c     ========================
c
      niter = ih
c
c     ...Estimate relative error on gamma 1
c
      error1 = error1 / abs( wgam1 )
      error0 = abs( (gam0 - wgam0) / gam0 )
      if (error1 .ge. error0) then
         error = error1
      else
         error = error0
      endif
c
c     ...Return gamma1
c
      gam1 = wgam1
c
      return
      end
c
c     ............................................................
      subroutine matmul( ld, l, m, n, a, b, c )
c     ............................................................
c     Antonio Pedrosa and Bruce Schmeiser
c       Purdue University
c       May 1993
c     purpose:
c       multiply matrices    a (l*m)  b (m*n) = c (l*n)
c
      implicit real    (a-h, o-z)
      implicit integer (i-n)
c
      double precision a(ld,ld), b(ld,ld), c(ld,ld), sum
c
      do 1000 i=1,l
```

```
        do 1100 j=1,n
           sum = 0.d0
           do 1200 k=1,m
              sum = sum + (a(i,k) * b(k,j))
 1200      continue
           c(i,j) = sum
 1100   continue
 1000 continue
c
      return
      end
```

Appendix C: FORTRAN Code for the 1-2-1 OBM Algorithm

```
c     ***************************************************************************
      subroutine m121obm(n,data,dmean,dvar,xobm)
c     ***************************************************************************
c     Written by:     Antonio Pedrosa and Bruce Schmeiser
c                     School of Industrial Engineering, Purdue University
c                     April 1994
c
c     Purpose:        Estimate variance of the sample mean
c
c                     ====================
c     Reference:      ALGORITHM 1--2--1 OBM
c                     ====================
c                     "Estimating Variance of the Sample Mean: Optimal Batch
c                        Size Estimation and 1-2-1 Overlapping Batch Means"
c                     Technical Report SMS94-3
c                     School of Industrial Engineering, Purdue University
c
c     Method:         Estimate optimal batch size using overlapping batch means
c                     (a) estimate gamma0 using obm(m0)
c                            m0 = n**0.5
c                            gamma0 = n * obm(m0) / dvar
c                     (b) estimate center of gravity
c                         set m1
c                            m1 = max{ n**(1/6), 0.75*max{gamma0, 1/gamma0} }
c                         call gravity to get gamma0 and gamma1
c                     (c) estimate variance of sample mean
c                         estimate optimal batch size
c                            mhat = [1.5*n*((gamma1/gamma0)**2)]**(1/3) + 1
c                         estimate vhat using obm(mhat)
c      INPUT
c         n      (sample size)
c         data   (vector of n data points)
c         dmean  (sample mean of the data)
c         dvar   (sample variance of the data)
c
c      OUTPUT
c         xobm   (estimate of variance of the sample mean)
c
       implicit real    (a-h, o-z)
       implicit integer (i-n)
       dimension    data(n)
c
       fn = n
       m0 = sqrt(fn)
       call obm(n,data,dmean,m0,xobm0)
       xgamma0 = fn * xobm0 / dvar
c
       x  = 0.75 * max(xgamma0, 1./xgamma0)
       y = fn**(1./6.)
```

```
c
      m1 = max(x,y)
      m1 = min(m1, n/4)
c
      call gravity(n,data,dmean,dvar,m1,xgamma0,xgamma1)
c
      mhat = (1.5*fn*(xgamma1/xgamma0)**2)**(1.0/3.0) + 1.
      mhat = min(mhat, n/4)
c
      call obm(n,data,dmean,mhat,xobm)
c
      return
      end
c
c     ************************************************************************
      subroutine gravity(n,data,dmean,dvar,m1,xgamma0,xgamma1)
c     ************************************************************************
c     Written by:    Antonio Pedrosa and Bruce Schmeiser
c                    School of Industrial Engineering, Purdue University
c                    April 1994
c
c     Purpose:       Estimate center of gravity
c                         (more exactly, estimate gamma0 and gamma1)
c
c                    =====================
c     Reference:     ALGORITHM 1--2--1 OBM
c                    =====================
c                    "Estimating Variance of the Sample Mean: Optimal Batch
c                        Size Estimation and 1-2-1 Overlapping Batch Means"
c                    Technical Report SMS94-3
c                    School of Industrial Engineering, Purdue University
c
c     Method:        Estimate gamma0 from 1 OBM and estimate gamma1 from 2 OBM's
c
c     INPUT
c        n     (sample size)
c        data  (vector of n data points)
c        dmean (sample mean of the data)
c        dvar  (sample variance of the data)
c        m1    (smallest batch size to be used)
c
c     OUTPUT
c        xgamma0 (estimate of gamma0, the sum of autocorrelations)
c        xgamma1 (estimate of gamma1, the weighted sum of autocorrelations)
c
      implicit real    (a-h, o-z)
      implicit integer (i-n)
      dimension    data(n)
c
      m2 = m1 + 1
      call obm(n,data,dmean,m1,xobm1)
      xgamma0 = float(n) * xobm1 / dvar
```

```
      call obm(n,data,dmean,m2,xobm2)
      xgamma1 = (xobm2 - xobm1) * float(m1) * float(m2) * float(n) / dvar
c
      return
      end
c
c     ***************************************************************************
      subroutine obm(n,data,dmean,m,xobm)
c     ***************************************************************************
c     Written by:    Antonio Pedrosa and Bruce Schmeiser
c                    School of Industrial Engineering, Purdue University
c                    April 1994
c
c     Purpose:       Obtain OBM estimate of the variance of the sample mean
c
c     Reference:     "Estimating Variance of the Sample Mean: Optimal Batch
c                        Size Estimation and 1-2-1 Overlapping Batch Means"
c                    Technical Report SMS94-3
c                    School of Industrial Engineering, Purdue University
c
c     Method:        OBM
c
c     INPUT
c        n     (sample size)
c        data  (vector of n data points)
c        dmean (sample mean of the data)
c        m     (batch size)
c
c     OUTPUT
c        xobm  (estimate of variance of sample mean)
c
      implicit real    (a-h, o-z)
      implicit integer (i-n)
      dimension    data(n)
c
      sum = 0.
      do 1000 j=1,m
         sum = sum + data(j)
 1000 continue
c
      fm = m
      sumsq = 0.
      do 2000 i=1,n-m+1
         sumsq = sumsq + ( sum / fm - dmean )**2
         sum = sum - data(i) + data(m+i)
 2000 continue
c
      xobm = (fm * sumsq / float(n-m+1) ) / float(n-m)
c
      return
      end
```

VITA

VITA

António Manuel de Carvalho Pedrosa was born on July 1, 1958, in Porto, Portugal. António received his undergraduate degree in Electrical Engineering in 1982 at Universidade do Porto, Portugal. In July 1990 he received his Master of Science in Engineering Degree at Purdue University. Between March 1981 and August 1989 he worked in the industry as a systems engineer, and taught at Universidade do Porto and at Escola Superior de Biotecnologia, Universidade Católica Portuguesa.

In January 1992 António came to the School of Industrial Engineering at Purdue University, majoring in Simulation and Applied Probability. In May 1994 he completed his Ph.D. dissertation entitled "Automatic Batching in Simulation Output Analysis" under the supervision of Professor Bruce Schmeiser.

António has accepted a position as Assistant Professor, at Escola Superior de Biotecnologia, Universidade Católica Portuguesa, Porto, Portugal.