MULTIVARIATE STANDARDIZED TIME SERIES FOR STEADY-STATE SIMULATION OUTPUT ANALYSIS

DAVID F. MUÑOZ

Departamento de Ingeniería Industrial, Instituto Tecnológico Autónomo de México, Mexico City 01000, davidm@itam.mx

PETER W. GLYNN

Department of EES & OR, Stanford University, Stanford, California 94305, glynn@leland.stanford.edu

(Received May 1998; revised May 1999; accepted March 2000)

The theory of standardized time series, initially proposed to estimate a single steady-state mean from the output of a simulation, is extended to the case where more than one steady-state mean is to be estimated simultaneously. Under mild assumptions on the stochastic process representing the output of the simulation, namely a functional central limit theorem, we obtain asymptotically valid confidence regions for a (multivariate) steady-state mean based on multivariate standardized time series. We provide examples of multivariate standardized time series, including the multivariate versions of the batch means method and Schruben's standardized sum process. Large-sample properties of confidence regions obtained from multivariate standardized time series are discussed. We show that, as in the univariate case, the asymptotic expected volume of confidence regions produced by standardized time series procedures is larger than that obtained from a consistent estimation procedure. We present and discuss experimental results that illustrate our theory.

Most of the work in steady-state simulation output analysis has been devoted to the estimation of a single steady-state measure of performance. For many real-world simulations, however, it is of interest to estimate more than one steady-state parameter associated with the system performance (e.g., Chen and Seila 1987, Law 1983, or Schruben 1981). In this paper, we consider the steady-state estimation problem of simulation when more than one steady-state mean is to be estimated simultaneously. To be more precise, we consider the output of the simulation as an $\Re^d$-valued stochastic process $Y = \{Y(s) : s \geqslant 0\}$. We assume that the process Y possesses a steady-state mean, that is,

$$r(t) \stackrel{\text{def}}{=} \frac{1}{t}\int_0^t Y(s)\,ds = \frac{1}{t}\left(\int_0^t Y_1(s)\,ds, \ldots, \int_0^t Y_d(s)\,ds\right)^T \Rightarrow (r_1, \ldots, r_d)^T = r, \tag{1}$$

as $t \to \infty$, where $\Rightarrow$ denotes weak convergence, and $r \in \Re^d$ is a given (but unknown) parameter. In what follows, any vector $x \in \Re^d$ will be regarded as a column vector (as in (1)).

Equation (1) states that $r(t)$ is a consistent estimator for r, a property that is required in most point estimation procedures to ensure that the point estimator $r(t)$ approaches the desired parameter r as the run length t increases. In addition, a good estimation procedure should include an assessment of the precision of this estimation. We are interested in methods that produce an asymptotically valid confidence region $R(t) \subset \Re^d$ (depending on the output of the simulation up to the run length t) for the multivariate steady-state mean r. That is, we want to generate a region $R(t) \subset \Re^d$ such that

$$\lim_{t\to\infty} P[(r_1, r_2, \ldots, r_d)^T \in R(t)] = 1 - \alpha. \tag{2}$$

We specialize in cancellation procedures (see Glynn and Iglehart 1990 and Muñoz 1991), that is, we do not attempt to estimate the covariance matrix of the point estimator consistently. Instead we try to scale up the process in such a way that the covariance matrix cancels out in the limiting distribution of the corresponding central limit theorem for the point estimator. In this direction, the theory of standardized time series (Schruben 1982, 1983; Glynn and Iglehart 1990; Goldsman and Schruben 1984; Goldsman et al. 1986; Goldsman et al. 1990) provides an appropriate framework for deriving cancellation procedures that produces asymptotically valid confidence intervals for a single steady-state mean. Based on the work of Glynn and Iglehart (1990) and Chen and Seila (1987), we extend the theory of standardized time series to the multivariate steady-state mean estimation problem, with the objective obtaining asymptotically valid confidence regions $R(t)$ satisfying (2).

We start in §1 by describing the mathematical framework that underlines our study of multivariate standardized time series. We state our functional central limit theorem assumption concerning the underlying stochastic process representing the output of the simulation. Based on well-known results for the case of independent and identically distributed (i.i.d.) observations, we describe how to generate confidence regions for an $\Re^d$-valid steady-state mean r when the covariance matrix of the point estimator can be estimated consistently.

Subject classifications: Simulation, statistical analysis: multivariate steady-state output analysis.
Area of review: SIMULATION.

0030-364X/01/4903-0413 $05.00
1526-5463 electronic ISSN

In §2, we propose a natural extension of the method of standardized time series to estimate a multivariate steady-state mean, and provide examples of multivariate standardized time series procedures. These examples include the multivariate versions of the batch means method and Schruben's standardized sum process as well as a subclass of standardized time series procedures, which we shall refer to as "linear procedures." Large-sample properties of confidence regions obtained from multivariate standardized time series, and its comparison relative to consistent estimation procedures, are given in §3. We show that, as in the univariate case (see Glynn and Iglehart 1990), the asymptotic expected volume of confidence regions produced by standardized time series procedures is larger than that obtained from a consistent estimation procedure. For members of the linear subclass we obtain an explicit expression for the asymptotic volume as a function of the number of batches. This result shows that the asymptotic volume of consistent estimation procedure is approached by a member of the linear subclass as the number of batches becomes large. Finally, in §4 we report experimental results from the application of multivariate standardized time series procedures to estimate more than one steady-state parameter simultaneously.

1. MATHEMATICAL FRAMEWORK

1.1. The Steady-State Mean Confidence Region Problem

Observe that, under the assumption stated in (1), $r(t)$ is a consistent estimator for the steady-state mean r, and it can be used as a point estimator for r. However, the estimation procedure should include an assessment of the precision of this estimator. In the case $d > 1$, the quality of the point estimator $r(t)$ can be assessed if we produce a $100(1-\alpha)\%$ asymptotically valid confidence region $R(t)$ (possibly centered at $r(t)$), such that (2) holds. By analogy with the case $d = 1$, consistent estimation procedures may be used to obtain asymptotically valid confidence regions under a central limit theorem (CLT) assumption. In this setting, the CLT takes the form:

$$t^{1/2}(r(t)-r) \Rightarrow GN_d(0, I), \tag{3}$$

as $t \to \infty$, where G is a nonsingular $d \times d$ matrix, and $N_d(0, I)$ denotes a normal d-variate distribution with mean 0 and covariance matrix I (where I denotes the identity matrix). Note that $\Sigma = GG^T$ is the covariance matrix of the normal random vector appearing in the right-hand side of (3).

If we can estimate Σ consistently, say by $\widehat{\Sigma}(t)$, then $\widehat{\Sigma}(t)$ is positive definite for t large enough (see Dykstra 1970). Hence, $\widehat{\Sigma}(t)^{-1/2}$ is well defined for t large enough (see p. 285 of Searle 1982 for a definition of the square root of a positive definite symmetric matrix). Therefore, it follows from (3) and the converging together principle (see Billingsley 1979) that

$$t^{1/2}\widehat{\Sigma}(t)^{-1/2}(r(t)-r) \Rightarrow N_d(0, I),$$

as $t \to \infty$. Then, from the continuous mapping theorem (see Corollary 1.9 of Ethier and Kurtz 1986), we have

$$t(r(t)-r)^T\widehat{\Sigma}(t)^{-1}(r(t)-r) \Rightarrow \chi^2_{(d)}, \tag{4}$$

as $t \to \infty$, where $\chi^2_{(d)}$ denotes a chi-square distribution with d degrees of freedom.

From (4), a $100(1-\alpha)\%$ asymptotic confidence region for r is given by

$$R(t)=\{x \in \Re^d : t(r(t)-x)^T\widehat{\Sigma}(t)^{-1}(r(t)-x) \leqslant \chi^2_{(d,\alpha)}\}, \tag{5}$$

where $\chi^2_{(d,\alpha)}$ is a constant chosen so that $P[\chi^2_{(d)} \geqslant \chi^2_{(d,\alpha)}] = \alpha$.

If $\widehat{\Sigma}(t)$ is positive definite (as we expect for large t, since we assume that $\widehat{\Sigma}(t)$ is a consistent estimator), the region $R(t)$ is an ellipsoid centered at $r(t)$. Techniques that produce confidence regions in the form of (5) are called *consistent estimation procedures* (see Glynn and Iglehart 1990). A consistent estimation procedure that can be generalized to the multivariate case is the regenerative method (e.g., Iglehart 1978 or Crane and Lemoine 1977). A brief description of how the regenerative method generates a consistent estimator for Σ is provided in Appendix B of Muñoz (1991).

We prefer to study cancellation methods in this paper, because consistent estimation of Σ often requires special conditions on the structure of the process Y (as in the regenerative method). Also, a consistent estimation procedure frequently required the setting of certain parameters that are problem dependent, e.g., the rate at which the number of batches should go to infinity with the run length is a batch means procedure (see Damerdji 1995). Schruben (1983) introduced the method of standardized time series to obtain confidence intervals by using a different approach. This method suggests that we can eliminate Σ from the CLT (3) by scaling the variables appropriately. Standardized time series procedures can be extended to the multivariate case, as we are going to see in §2.

1.2. A Functional Central Limit Theorem Assumption

As discussed before, it is necessary that the output process Y satisfies the law of large numbers (1) in order that the steady-state estimation problem be well defined. However, the development of an asymptotically valid confidence region methodology requires making additional assumptions that permit one to describe the variability of the estimator $r(t)$ about the steady-state mean r. In particular, a standard assumption that (implicitly) underlies much of the existing steady-state simulation methodolology is that $r(t)$ satisfy the CLT (3).

It turns our that the methodology that we propose here requires a slightly stronger type of assumption. Set $\overline{Y}_t = \{\overline{Y}_t(u) : 0 \leqslant u \leqslant 1\}$, and $X_t = \{X_t(u), 0 \leqslant u \leqslant 1\}$, where

$$\overline{Y}_t(u) = \frac{1}{t}\int_0^{ut} Y(s)\,ds,$$
$$X_t(u) = t^{1/2}(\overline{Y}_t(u) - ru).$$

Recall that a d-dimensional standard Brownian motion is a stochastic process $B = \{B(u) : 0 \leqslant u \leqslant 1\}$ with stationary independent increment such that $B(u)$ has a $N_d(0, uI)$ distribution. We shall demand that the following so-called functional central limit theorem (FCLT) for Y be valid:

ASSUMPTION 1. *There exists a nonsingular $d \times d$ matrix G such that*

$$X_t \Rightarrow GB,$$

as $t \to \infty$ (in the topology of weak convergence in $C^d[0, 1]$, the space of $\Re^d$-valued continuous functions defined on $[0, 1]$; see Ethier and Kurtz 1986 for additional discussion).

The reason that the ordinary CLT (3) typically holds for a steady-state simulation depends on the fact that observations taken from Y that are widely separated in time are approximately i.i.d. As a consequence, $r(t)$ behaves very much like an average of i.i.d. random vectors, and one can therefore expect a CLT to hold. The same independence argument leads naturally to the additional structure associated with the FCLT required in Assumption 1 to hold in virtually any real-world discrete-event simulation in which the steady-state simulation problem is well defined. From a mathematical viewpoint, FCLT theorems have been established for Markov processes in discrete and continuous time, stationary processes satisfying so-called "mixing conditions," and associated processes; see Glyn and Iglehart (1990) and Muñoz (1991) for additional details.

We note that Assumption 1 implies the law of large numbers (3). Thus, the steady-state estimation problem is always well defined under Assumption 1. In addition, we remark that while the mathematical discussion of this paper will focus exclusively on a continuous-time output process Y, a discrete-time output process $Z = \{Z_k : k \geqslant 0\}$ can be incorporated into our framework by setting $Y(s) = Z_{\lfloor s \rfloor}$, where $\lfloor s \rfloor$ is the integer part of s.

2. MULTIVARIATE STANDARDIZED TIME SERIES

In this section, we will state the basic definitions that allow us to extend the theory of (univariate) standardized time series, as described in Glynn and Iglehart (1990), to the case where more than one steady-state mean is to be estimated simultaneously. We also show how the most popular univariate standardized time series procedures can be generalized to the case $d > 1$.

2.1. Basic Definitions

To apply the method of standardized time series we need to cancel out the matrix G in (3). The strategy is to consider a quadratic form of the random vector $t^{1/2}(r(t) - r)$, scaled so that the limiting distribution is independent of G. The corresponding matrix associated with this quadratic form will depend on the process $\{Y(s) : 0 \leqslant s \leqslant t\}$. To be more precise, we introduce a function $g\colon C^d[0, 1] \to \Re^{d\times d}$, and we then consider the random variable $t(\overline{Y}_t(1) - r)^T g(X_t)^{-1}(\overline{Y}_t(1) - r)$. If g is suitably continuous, this random variable will converge to $B(1)^T G^T g(GB)^{-1} GB(1)$ (under Assumption 1). To cancel out G in the limiting distribution, we need

$$g(Gx) = Gg(x)G^T \tag{6}$$

for any nonsingular $d \times d$ matrix G, and $x \in C^d[0, 1]$. Since the process X_t depends explicitly on the unknown r, to make $g(X_t)$ independent of r we require

$$g(x - \beta J) = g(x) \tag{7}$$

for $x \in C^d[0, 1]$, and $\beta \in R^d$, where $J(t) = t, 0 \leqslant t \leqslant 1$. Also, to obtain confidence regions that are ellipsoids, we need

$$P[g(B) \text{ is positive definite and symmetric}] = 1. \tag{8}$$

Finally, to apply the continuous mapping theorem, g must satisfy

$$P[B \in D(g)] = 0, \tag{9}$$

where $D(g)$ is the set of discontinuities of g. Let us denote by $\mathcal{M}$ the class of functions $g\colon C^d[0, 1] \to \Re^{d\times d}$ that satisfy (6)–(9). Then we have the following:

THEOREM 1. *Suppose that $g \in \mathcal{M}$ and Y satisfies Assumption 1. Then*

$$(r(t) - r)^T g(\overline{Y}_t)^{-1}(r(t) - r) \Rightarrow B(1)^T g(B)^{-1} B(1),$$

as $t \to \infty$.

PROOF. Let us define $h\colon C^d[0, 1] \to \Re$ by

$$h(x) = x(1)^T g(x)^{-1} x(1).$$

If follows from (6) and (8) that

$$g(GB)^{-1} = (Gg(B)G^T)^{-1} = (G^T)^{-1} g(B)^{-1} G^{-1} a.s.,$$

where the notation *a.s.* ("almost surely") is equivalent to say that this equality holds with probability one. Then

$$h(GB) = (GB(1))^T g(GB)^{-1} GB(1) = B(1)^T g(B)^{-1} B(1).$$

Also, Assumption (9) guarantees that $P[GB \in D(h)] = 0$, so that from the continuous mapping theorem we have

$$h(X_t) \Rightarrow h(GB) = B(1)^T g(B)^{-1} B(1), \tag{10}$$

as $t \to \infty$. However,

$$h(X_t) = t^{1/2}(\overline{Y}_t(1) - r)^T g(X_t)^{-1}(\overline{Y}_t(1) - r)t^{1/2}$$
$$= (\overline{Y}_t(1) - r)^T g(\overline{Y}_t - rJ)^{-1}(\overline{Y}_t(1) - r)$$
$$= (r(t) - r)^T g(\overline{Y}_t)^{-1}(r(t) - r). \quad (11)$$

Hence, the conclusion follows from (10) and (11). □

Based on Theorem 1, we can obtain a $100(1-\alpha)\%$ asymptotically valid confidence region for r. If $z(g, \alpha)$ is a continuity point of the distribution of $B(1)^T g(B)^{-1} B(1)$, and $P[B(1)^T g(B)^{-1} B(1) \leqslant z(g, \alpha)] = 1 - \alpha$, the region

$$R(t) = \{t \in \Re^d : (r(t) - x)^T g(\overline{Y}_t)^{-1}(r(t) - x) \leqslant z(g, \alpha)\} \quad (12)$$

defines a $100(1-\alpha)\%$ asymptotically valid confidence region for r. In all of our examples $g(\overline{Y}_t)$ is positive semidefinite and symmetric, so that the region $R(t)$ is an ellipsoid centered at $\overline{Y}_t(1)$.

To specify completely the confidence region defined in (12), we need to derive the distribution of $B(1)^T g(B)^{-1} B(1)$ for the particular function $g \in \mathcal{M}$. We end this subsection with a characterization of the class of functions $\mathcal{M}$ that is helpful in deriving the distribution of $B(1)^T g(B)^{-1} B(1)$. This characterization is the multivariate version of that obtained in Glynn and Iglehart (1990).

Let $\gamma: C^d[0,1] \to C^d[0,1]$ be defined by

$$(\gamma x)(t) = x(t) - tx(1), \quad (13)$$

and Λ be the class of functions b: $C^d[0,1] \to R^{d \times d}$ that satisfy

$$b(GX) = Gb(x)G^T \quad (14)$$

for any nonsingular $d \times d$ matrix G and any $x \in C^d[0,1]$,

$$P[(b \circ \gamma)(B) \text{ is positive definite and symmetric}] = 1, \quad (15)$$

and

$$P[B \in D(b \circ \gamma)] = 0, \quad (16)$$

where $\circ$ denotes "composition" of function (i.e. $(f_1 \circ f_2)(x) = f_1(f_2(x))$ for any x in the domain of f_2). By denoting $\mathcal{M}^* = \{g : g = b \circ \gamma, b \in \Lambda\}$, we obtain the following proposition.

PROPOSITION 1. $\mathcal{M}^* = \mathcal{M}$.

PROOF. Let us suppose $g \in \mathcal{M}^*$. Then $g = b \circ \gamma$ for some $b \in \Lambda$. Clearly, g satisfies (14), (15), and (16). Also, (13) is verified since $\gamma(x - \beta J) = \gamma(x)$. Therefore, $g \in \mathcal{M}$ and $\mathcal{M}^* \subseteq \mathcal{M}$ holds. Conversely, if $g \in \mathcal{M}$, by taking $\beta = x(1)$ in (13), we see that $g(x) = g(\gamma(x))$, that is $g = g \circ \gamma$. Therefore, $g \in \mathcal{M}^*$ and $\mathcal{M} \subseteq \mathcal{M}^*$ holds. □

Now we can state the following result, which can be useful in deriving the distribution of $B(1)^T g(B)^{-1} B(1)$.

PROPOSITION 2. *If $g \in \mathcal{M}$, then $B(1)$ is independent of $g(B)$.*

PROOF. Since the process $\{B(u) - uB(1) : 0 \leqslant u \leqslant 1\}$ is independent of $B(1)$ (the proof given in p. 84 of Billingsley 1968 extends to the $\Re^d$-valued case since the stochastic processes B_i, $i = 1, 2, \dots, d$, are independent), we have that $\gamma(B)$ is independent of $B(1)$. Hence, it follows from the last proposition that $g(B) = (b \circ \gamma)(B)$ is independent of $B(1)$. □

2.2. Examples of Multivariate Standardize Time Series

In this subsection, we show how the most popular univariate standardized time series procedures can be generalized to the case $d > 1$.

EXAMPLE 1 (BATCH MEANS METHOD). The batch means method is one of the most important techniques for producing asymptotic confidence intervals for a steady-state mean (e.g., Law 1983 or Fishman 1978). To describe the extension of the batch means method to the multivariate case, we proceed as in the univariate case (see Chen and Seila 1987), that is, we subdivide the run length t into m nonoverlapping pieces of equal length. Then, we compute the sample batch means:

$$\overline{X}_i(t) = \frac{m}{t} \int_{\frac{(i-1)t}{m}}^{\frac{it}{m}} Y(s)ds, \quad i = 1, 2, \dots, m,$$

and the sample covariate matrix of the batch means:

$$S_m(t) = \frac{1}{m-1} \sum_{i=1}^{m} [\overline{X}_i(t) - r(t)][\overline{X}_i(t) - r(t)]^T. \quad (17)$$

A confidence region for r based on the batch means method is given by

$$R(t) = \left\{ r \in \Re^d : m(r(t) - r)^T S_m^{-1}(t)(r(t) - r) \leqslant \frac{d(m-1)}{(m-d)} F_{(d, m-d, \alpha)} \right\}, \quad (18)$$

where $F_{(d, m-d, \alpha)}$ is the $(1-\alpha)$-quantile of an F distribution with $(d, m-d)$ degrees of freedom. Note that $m \geqslant (d+1)$ is required since $S_m(t)$ is nonsingualr only if $m \geqslant (d+1)$ (see p. 208 of Searle 1982).

The multivariate version of the batch means method is a standardized time series procedure. To see this, let $m \geqslant (d+1)$, and g_m^b: $C^d[0,1] \to \Re^{d \times d}$ be defined by

$$g_m^b(x) = \frac{m}{m-1} \sum_{i=1}^{m} [\Delta x(i/m) - x(1)/m][\Delta x(i/m) - x(1)/m]^T$$

for $x \in C^d[0,1]$, and $\Delta x(i/m) = x(i/m) - x((i-1)/m)$, $i = 1, 2, \dots, m$.

Then, we can easily verify that $S_m(t) = mg_m^b(\overline{Y}_t)$. To see that $g_m^b \in \mathcal{M}$, note that if $y = Gx$, then $\Delta y(i/m) = G\Delta x(i/m)$ and $y(1) = Gx(1)$, from which we can see that (6) is satisfied. Similarly, we can verify (7). To verify (8), we consider the fact that the sample covariance matrix of a random sample of size m from a $N_d(\mu, \Sigma)$ distribution is positive definite with probability one if Σ is positive definite and $m \geqslant (d+1)$ (see Dykstra 1970), from which (8) follows since $\Delta B(i/m)$ are independent random vectors distributed as $(1/\sqrt{m})N(0, I)$. Finally, (9) is satisfied since our function g_m^b is continuous on $C^d[0, 1]$.

Since $B(1)/m$ and $g(B)/m$ are the sample mean and the sample covariance matrix of $\Delta B(i/m)$, $i = 1, 2, \ldots, m$, respectively, it follows that (see §5.2 of Siotani et al. 1985) $B(1)^T g_m^b(B)^{-1} B(1)$ is distributed as

$$\frac{d(m-1)}{(m-d)} F_{(d, m-d)}.$$

Therefore, from Theorem 1 we have

$$(r(t)-r)^T g_m^b(\overline{Y}_t)^{-1}(r(t)-r) \Rightarrow \frac{d(m-1)}{(m-d)} F_{(d, m-d)},$$

as $t \to \infty$. Since $S_m(t) = m g_m^b(\overline{Y}_t)$, an asymptotically valid $100(1-\alpha)\%$ confidence region for r is given by (18).

EXAMPLE 2 (A LINEAR SUBCLASS). A subclass of $\mathcal{M}$ can be obtained from an "extension" procedure presented in Glynn and Iglehart (1990) for the univariate case. The idea of an extension procedure is to extend the power of the method by applying the procedure separately to each independent increment of the output process, and then "patching" the increments together (this idea was first proposed in Schruben 1983). To generalize the extension procedure presented by Glynn and Iglehart, the process X_t, is subdivided into m pieces, and then we apply a standardized time series procedure to each increment as if it were the output of a single run by itself (recall that each increment of a Brownian motion is itself a Brownian motion). For $0 \leqslant i \leqslant m-1$, we give the $(i+1)$th increment as an element of $C^d[0, 1]$ by defining the map Λ_i^m: $C^d[0, 1] \to C^d[0, 1]$ given by

$$(\Lambda_i^m x)(t) = x((i+t)/m) - x(i/m), \quad 0 \leqslant t \leqslant 1.$$

For a given g satisfying (6), (7), and (9), we define its extension g_m by

$$g_m = \sum_{i=0}^{m-1} g \circ \Lambda_i^m. \tag{19}$$

PROPOSITION 3. *Assume that* g: $C^d[0, 1] \to \mathfrak{R}^{d \times d}$ *satisfies* (6), (7), *and* (9). *If* g_m *as defined in* (19) *satisfies*

$$P[g_m(B) \textit{ is positive definite and symmetric}] = 1,$$

then $g_m \in \mathcal{M}$.

PROOF. Clearly (6) is verified for g_m. Also, we can easily verify (7), since g satisfied (7), and if $y = x - \beta J$, then

$$\Lambda_i^m y = \Lambda_i^m x - (\beta/m)J.$$

Finally, note that $\sqrt{m}\Lambda_i^m B$ is distributed as B, and

$$g(\Lambda_i^m B) = (1/m) g(\sqrt{m}\Lambda_i^m B).$$

Hence, (9) implies that $P[\Lambda_i^m B \in D(g)] = 0$. Therefore, (9) is also satisfied for g_m in place of g. □

Now, we can define a subclass of $\mathcal{M}$ as follows. Let us consider a real valued function v: $C[0, 1] \to \mathfrak{R}$. We say that v is a bounded linear functional if it is linear, that is,

$$v(\alpha x + \beta y) = \alpha v(x) + \beta v(y), \quad \alpha, \beta \in \mathfrak{R}, \ x, y \in C[0, 1],$$

and there exists a constant $M \in \mathfrak{R}$ such that

$$|v(x)| \leqslant M\delta(x, 0), \quad x \in C[0, 1],$$

where δ is the sup-distance (see p. 133 of Royden 1968). As is known, a given function v: $C[0, 1] \to \mathfrak{R}$ is a bounded linear functional if and only if it is linear and continuous (see p. 184 of Royden 1968). Let $C^*[0, 1]$ denote the space of functions v: $C[0, 1] \to \mathfrak{R}$ that are bounded linear functionals, and let $C_d^*[0, 1]$ be the class of functions b: $C^d[0, 1] \to \mathfrak{R}^{d \times d}$ of the form

$$b(x) = h(x)h(x)^T, \quad x \in C^d[0, 1],$$

where $h(x) = (v(x_1), v(x_2), \ldots, v(x_d))^T$ for some $v \in C^*[0, 1]$. Note that, from the Riesz representation theorem (see p. 310 of Royden 1968), $v(x)$ can be written as

$$v(x) = \int_0^1 x(t)\mu(dt)$$

for some finite signed Baire measure μ (this representation relates our subclass to the concept of standardized time series weighted area variance estimators proposed in Goldsman et al. 1986).

The linear subclass $\mathcal{L}$ is defined as the class of functions g: $C^d[0, 1] \to \mathfrak{R}^{d \times d}$ such that

(i) g satisfies

$$P[g(B) \text{ is nonsingular}] = 1,$$

and

(ii) For some $m \geqslant d$, and $b \in C_d^*[0, 1]$, g can be written as

$$g = \sum_{i=0}^{m-1} b \circ \gamma \circ \Lambda_i^m, \tag{20}$$

where γ is defined in (13).

Note that a matrix $b(x) = h(x)h(x)^T$ has rank 1. Hence, in order that $g(x)$ be nonsingular we require $m \geqslant d$ (see p. 208 of Searle 1982). The linear subclass is of particular importance because for members of this class we can explicitly obtain the limiting distribution of Theorem 1. It turns out that the limiting distribution is a multiple of an F distribution, as we see from the following propositions (proofs for these two propositions can be found in Muñoz 1991). □

PROPOSITION 4. $\mathcal{L} \subseteq \mathcal{M}$.

PROPOSITION 5. *Let $g \in \mathcal{L}$ and $m \geqslant d$ such that g satisfies* (20). *Then*

$$(r(t)-r)^T g(\overline{Y}_t)^{-1}(r(t)-r) \Rightarrow \frac{dm}{(m-d+1)\sigma_g^2} F_{(d,\, m-d+1)},$$

as $t \to \infty$, where $\sigma_g^2 > 0$ is a constant determined by g.

We mentioned that σ_g^2 can be obtained from

$$\sigma_g^2 = E[v(W_1)^2], \tag{21}$$

where W_1 is the Brownian bridge defined by $W_1(u) = B_1(u) - uB_1(1)$, $0 \leqslant u \leqslant 1$. The multivariate version of Schruben's standardized sum process belongs to the linear subclass. In this case we have

$$v(x) = \int_0^1 x(t)dt, \quad \text{and} \quad \sigma_g^2 = \frac{1}{12}.$$

EXAMPLE 3 (COMBINED PROCEDURES). An interesting property of members of the linear subclass is that the asymptotic expected volume of their confidence regions is independent of σ_g^2 (as we will see in §3), but it does depend on the degrees of freedom of the corresponding F distribution given in Proposition 5. Smaller asymptotic expected volumes correspond to larger degrees of freedom. This suggests that we can improve the power of standardized time series procedures if we can increase the degrees of freedom of the F distribution that appears in the corresponding limit theorem. An idea in this direction was proposed in Schruben (1983) for the univariate case. The idea is to increase the degrees of freedom of the F distribution by combining a member of the linear subclass with the batch means method. Based on the additivity property of the chi-square distribution (the sum of two independent chi-squares is also chi-square with degrees of freedom obtained by adding the degrees of freedom of the initial ones), the corresponding limiting distribution for the "combined" procedure is an F with more degrees of freedom than the initial ones. By considering the fact that Wishart matrices are also additive (see p. 67 of Siotani et al. 1985), we can develop combined procedures in a similar way for the case $d > 1$ (see Muñoz 1991 for details).

3. LARGE-SAMPLE PROPERTIES OF STANDARDIZED TIME SERIES

In this section, we analyze the asymptotic properties of multivariate standardized time series. The asymptotic performance of an asymptotically valid confidence region can be evaluated by considering the asymptotic (as the run length goes to ∞) expected volume of the confidence region. From a volume standpoint, we view a confidence region as more desirable if it has a smaller volume.

3.1. Asymptotic Properties for the Class $\mathcal{M}$

Let us suppose $g \in \mathcal{M}$. Then its confidence region (defined in (12)) has volume

$$V(g) = \int_{R(t)} ds_1 \cdots ds_d = \det(g(\overline{Y}_t)^{1/2}) z(g, \alpha)^{d/2} q_d, \tag{22}$$

where

$$q_d = \frac{\pi^{d/2}}{\Gamma(\frac{d}{2}+1)}$$

is the volume of the d-dimensional unit sphere $\{x \in \Re^d : x^T x \leqslant 1\}$.

To compare standardized time series procedures to consistent estimation, we consider the volume of the confidence region obtained from a consistent estimation procedure (see (5)). This is given by

$$V(\widehat{\Sigma}(t)) = \det(\widehat{\Sigma}(t)^{1/2}) \left(\frac{\chi^2_{(d,\alpha)}}{t} \right)^{d/2} q_d, \tag{23}$$

where $\widehat{\Sigma}(t)$ is a consistent estimator for the covariance matrix GG^T that appears in Assumption 1. Note that the scaling factor $t^{-d/2}$ of (23) comes from the property $\det(\lambda A) = \lambda^d \det(A)$, where A is a $d \times d$ matrix.

Now, from (6) and (7) we see that, under Assumption 1,

$$tg(\overline{Y}_t) \Rightarrow g(GB) = Gg(B)G^T,$$

as $t \to \infty$. Therefore,

$$\begin{aligned} t^{d/2}\det(g(\overline{Y}_t)^{1/2}) &\Rightarrow [\det(Gg(B)G^T)]^{1/2} \\ &= \det(G)\det(g(B)^{1/2}), \end{aligned} \tag{24}$$

as $t \to \infty$. Hence, we see that $V(g)$ as defined in (22) is of order $t^{-d/2}$. As we can see from (23), convergence rate of $V(g)$ is the same order as that obtained from a consistent estimation procedure.

A more precise comparison can be made if we consider the asymptotic expected volume. If the sequence $\{\det(g(\overline{Y}_t)^{1/2}) : t > 0\}$ is uniformly integrable, we can take expectations of both sides of (22), and from (24) we obtain

$$\lim_{t \to \infty} t^{d/2} E[V(g)] = KE[\det(g(B)^{1/2})] z(g, \alpha)^{d/2}, \tag{25}$$

where $K = \det(G) q_d$.

Similarly, if the sequence $\{\det(\widehat{\Sigma}(t)^{1/2}) : t > 0\}$ is uniformly integrable we have

$$\lim_{t \to \infty} E[\det(\widehat{\Sigma}(t)^{1/2})] = \det(G),$$

so that from (23) we have

$$\lim_{t \to \infty} t^{d/2} E[V(\widehat{\Sigma}(t))] = K\chi^d_{(d,\alpha)}, \tag{26}$$

where $\chi^d_{(d,\alpha)} = (\chi^2_{(d,\alpha)})^{d/2}$. As we see from the following theorem, the right side of (25) cannot be smaller than that of (26) (a proof of this theorem is provided in the appendix).

THEOREM 2. *Let $g \in \mathcal{M}$ and $\widehat{\Sigma}(t)$ be a consistent estimator for GG^T. Assume that $\{\det(\widehat{\Sigma}(t)^{1/2}) : t \geqslant t_0\}$ is uniformly integrable for some $t_0 > 0$. Then, under Assumption* 1,

$$\liminf_{t \to \infty} t^{d/2} E[V(g)] \geqslant \lim_{t \to \infty} t^{d/2} E[V(\widehat{\Sigma}(t))],$$

where $V(g)$ and $V(\widehat{\Sigma}(t))$ are defined in (22) *and* (23), *respectively.*

Theorem 1 shows that the asymptotic expected volume of a standardized time series confidence region is at least as large as the asymptotic expected volume of a confidence region obtained from consistent estimation.

3.2. Asymptotic Properties for the Linear Subclass

For members of the linear subclass defined in §2, we can use Proposition 5 to derive an explicit expression for the asymptotic expected volume of the confidence region. If $g_m = \sum_{i=0}^{m-1} b \circ \gamma \circ \Lambda_i^m \in \mathcal{L}$, a $100(1-\alpha)\%$ asymptotically valid confidence region for r is given by

$$R(t) = \left\{ r \in \Re^d : [r(t) - r]^T g_m(\overline{Y}_t)^{-1} [r(t) - r] \leqslant \frac{dm}{(m-d+1)\sigma_{g_m}^2} F_{(d,\, m-d+1,\, \alpha)} \right\}, \quad (27)$$

where $\sigma_{g_m}^2$ is defined in (21). The volume of the confidence region defined by (27) is

$$V(g_m) = \det(g_m(\overline{Y}_t)^{1/2}) \left(\frac{dm}{(m-d+1)\sigma_{g_m}^2} F_{(d,m-d+1,\alpha)} \right)^{d/2} q_d.$$

As an appropriate measure of this volume we take its dth root (this is expressed in the same units as Y), that is, we consider

$$V_d(g_m) = [\det(g_m(\overline{Y}_t)^{1/2})]^{1/d} \cdot \left(\frac{dm}{(m-d+1)\sigma_{g_m}^2} F_{(d,\, m-d+1,\, \alpha)} \right)^{1/2} q_d^{1/d}.$$

To compare the members of the linear subclass relative to consistent estimation, we consider the dth root of the volume defined in (23):

$$V_d(\widehat{\Sigma}(t)) = t^{-1/2} [\det(\widehat{\Sigma}(t)^{1/2})]^{1/d} \chi(d, \alpha) q_d^{1/d}.$$

As is shown in Muñoz (1991), under the appropriate uniformly integrability assumption, an explicit expression for the asymptotic expected volume's dth of confidence regions obtained from members of the linear subclass is given by

$$\lim_{t\to\infty} t^{1/2} E[V_d(g_m)] = K_1 m^{-1/2} \prod_{i=1}^{d} E[\chi_{(m-i+1)}^{1/d}] \cdot \left(\frac{dm}{(m-d+1)} F_{(d,\, m-d+1,\, \alpha)} \right)^{1/2}, \quad (28)$$

where $K_1 = (\det(G) q_d)^{1/d}$, and

$$E[\chi_{(j)}^{1/d}] = \frac{2^{1/2d} \Gamma(\frac{1}{2d} + \frac{j}{2})}{\Gamma(\frac{j}{2})},$$

$j = 1, 2, \ldots$. Similarly (see Muñoz 1991 for details), under the appropriate uniformly integrability assumption, we have

$$\lim_{t\to\infty} t^{1/2} E[V_d(\widehat{\Sigma}(t))] = K_1 \chi_{(d,\alpha)}. \quad (29)$$

Equations (28) and (29) allows us to compute the asymptotic expected volume's dth root of confidence regions obtained from members of the linear subclass and from consistent estimation procedures (in units of $t^{1/2}\det(G)^{1/d}$). As an example, we show in Table 1 these asymptotic expected volume's dth root for 95% confidence regions, different values of d and m, and consistent estimation procedures ($m = \infty$). As we can see from Table 1, the asymptotic expected volume for members of the linear subclass is strictly bigger than that obtained from consistent estimation, but it can be approached arbitrarily closely by increasing m (an analytical proof of this result is given in Muñoz 1991). This suggests that, as in the case $d = 1$ (see Schmeiser 1982), for a sufficiently large number of batches, additional batches have small effect on the asymptotic expected volume of a member of the linear subclass.

Finally, we mention that, as for members of the linear subclass, the asymptotic expected volume's dth root of confidence regions provided by the batch means method and by combined procedures depend basically on the degrees of freedom of the corresponding F distribution (see Muñoz 1991 for details). To be more precise, if $g_m^b \in \mathcal{M}$ corresponds to the batch means method with m batches, and $g_m^c \in \mathcal{M}$ corresponds to a combined procedure with m batches, the asymptotic expected volume's dth root of their confidence regions are given by

$$\lim_{t\to\infty} t^{1/2} E[V_d(g_m^b)] = (m-1)^{-1/2} K_1 \prod_{i=1}^{d} E[\chi_{(m-i)}^{1/d}] \cdot \left(\frac{d(m-1)}{(m-d)} F_{(d,\, m-d,\, \alpha)} \right)^{1/2},$$

and

$$\lim_{t\to\infty} t^{1/2} E[V_d(g_m^c)] = (2m-1)^{-1/2} K_1 \prod_{i=1}^{d} E\left[\chi_{(2m-i)}^{1/d}\right] \cdot \left(\frac{d(2m-1)}{(2m-d)} F_{(d,\, 2m-d,\, \alpha)} \right)^{1/2}.$$

From which we can see that, the values in Table 1 apply to the batch means method (with $m - d$ in place of $m - d + 1$) and to combined procedures (with $2m - d$ in place of $m - d + 1$).

4. EXPERIMENTAL RESULTS

In this section, we present some experimental results obtained from applying multivariate standardized time series procedures to the simultaneous estimation of more than one steady-state parameter. The system selected to perform our experiments is an $M/M/1$ queue, and we consider the estimation of the steady-state mean $r = (r_1, r_2, \ldots, r_d)^T$, where r_i is the ith moment of the steady-state waiting time (excluding the service time). For a fixed number of replications, three procedures are compared: the

Table 1. Asymptotic expected volume's dth root for 95% confidence regions as a function of d and m.

	d								
$m-d+1$	1	2	3	4	5	6	7	8	9
1	20.248	31.377	41.870	51.613	60.659	69.095	77.006	84.472	91.540
2	7.626	10.270	12.502	14.467	16.241	17.869	19.380	20.796	22.132
3	5.864	7.432	8.664	9.718	10.655	11.508	12.297	13.035	13.729
4	5.220	6.403	7.281	8.013	8.658	9.241	9.778	10.279	10.751
5	4.892	5.881	6.580	7.152	7.649	8.096	8.506	8.889	9.249
6	4.695	5.568	6.160	6.634	7.043	7.408	7.742	8.053	8.346
7	4.564	5.359	5.880	6.290	6.639	6.950	7.233	7.496	7.743
8	4.470	5.211	5.681	6.045	6.351	6.623	6.869	7.098	7.312
9	4.401	5.100	5.532	5.861	6.136	6.378	6.597	6.799	6.989
10	4.344	5.013	5.418	5.720	5.972	6.190	6.389	6.565	6.737
11	4.301	4.944	5.326	5.608	5.831	6.030	6.213	6.383	6.540
12	4.269	4.893	5.249	5.513	5.729	5.914	6.073	6.230	6.374
13	4.240	4.846	5.186	5.435	5.639	5.811	5.958	6.105	6.227
14	4.214	4.805	5.131	5.367	5.559	5.721	5.858	5.995	6.121
15	4.191	4.769	5.090	5.317	5.490	5.643	5.782	5.901	6.018
16	4.172	4.739	5.050	5.267	5.432	5.577	5.709	5.820	5.931
17	4.157	4.715	5.018	5.218	5.385	5.523	5.637	5.754	5.847
18	4.142	4.690	4.985	5.187	5.338	5.470	5.589	5.689	5.790
19	4.131	4.672	4.960	5.156	5.301	5.428	5.532	5.638	5.722
20	4.120	4.654	4.935	5.125	5.266	5.387	5.487	5.589	5.669
30	4.050	4.548	4.784	4.939	5.046	5.138	5.208	5.283	5.338
40	4.015	4.491	4.715	4.853	4.944	5.022	5.078	5.129	5.170

batch means method, Schruben's standardized sum process method, and a methodology derived from univariate confidence intervals (which we refer to as the Bonferroni method).

The Bonferroni method corresponds to the hyperrectangular confidence region obtained from individual confidence intervals based on the univariate batch means method, each with the same confidence level $(1-\alpha/d)$. To be more explicit, if $s_i^2(t)$ denotes the ith diagonal element of the covariance matrix $S_m(t)$ corresponding to the batch means with m batches (as defined in (17)), the confidence region corresponding to the Bonferroni method is the region $I_1 \times I_2 \times \cdots \times I_D$, where

$$I_i = \left[r_i(t) - t_{(m-1,\alpha/2d)} \frac{s_i(t)}{m^{1/2}},\, r_i(t) + t_{(m-1,\alpha/2d)} \frac{s_i(t)}{m^{1/2}} \right],$$
$$i = 1, 2, \ldots, d,$$

where $r_i(t)$ is the ith component of $r(t)$, $t_{(m-1,\beta)}$ denotes the $(1-\beta)$ quantile of a Student t distribution with $(m-1)$ degrees of freedom and $(1-\alpha)$ is the desired confidence level ($\alpha = 0.05$ in our experiments). We point out that the Bonferroni method is a method derived from the well-known Bonferroni inequality, and as is well known, the confidence level of the Bonferroni method is always greater or equal than the desired confidence level $(1-\alpha)$.

Our comparisons are based on both the empirical coverage and the average volume of the confidence regions produced by these methods. Our observations were generated by FORTRAN programs running on a DEC Alpha 386 under UNIX, and using the Learmonth-Lewis random number generator (see LLRANDOMII generator in Lewis and Orav 1989).

Table 2. Performance of 95% confidence regions for the $d = 3$ first waiting time moments from an $M/M/1$ queue based on 1,000 independent replications (t = Run Length = 1,600,000 Observations).

ρ	m	Method	Coverage	Average Volume*	Standard Deviation*
0.5	5	Batch means	0.941	0.0515	0.0149
	5	Schruben	0.945	0.0356	0.0093
	5	Bonferroni	0.962	0.0763	0.0260
	10	Batch means	0.957	0.0244	0.0043
	10	Schruben	0.930	0.0233	0.0042
	10	Bonferroni	0.971	0.0591	0.0130
	20	Batch means	0.937	0.0207	0.0028
	20	Schruben	0.937	0.0204	0.0031
	20	Bonferroni	0.971	0.0535	0.0081
0.8	5	Batch means	0.928	2.5447	0.8507
	5	Schruben	0.919	1.7386	0.5368
	5	Bonferroni	0.967	4.2523	1.7075
	10	Batch means	0.894	1.1987	0.2770
	10	Schruben	0.876	1.1377	0.2751
	10	Bonferroni	0.972	3.2939	0.8479
	20	Batch means	0.883	1.0222	0.2107
	20	Schruben	0.872	1.0016	0.2232
	20	Bonferroni	0.969	3.1057	0.5829

*We report the average volume's dth root and the corresponding standard deviation.

In all of our experiments the output of the simulation is regarded as a discrete-time stochastic process. To be more precise, we consider $Y(s) = Z_{\lfloor s \rfloor}$, where

$$Z_k = (Z_k^1, Z_k^2, \ldots, Z_k^d)^T, \quad k = 1, 2, \ldots, t,$$

Table 3. Performance of 95% confidence regions for the $d = 3$ first waiting time moments from an $M/M/1$ queue based on 2,000 independent replications (t = Run Length = 3,200,000 Observations).

ρ	m	Method	Coverage	Average Volume*	Standard Deviation*
0.5	5	Batch means	0.962	0.0374	0.0110
	5	Schruben	0.942	0.0254	0.0064
	5	Bonferroni	0.975	0.0543	0.0179
	10	Batch means	0.955	0.0175	0.0032
	10	Schruben	0.930	0.0167	0.0029
	10	Bonferroni	0.978	0.0417	0.0095
	20	Batch means	0.947	0.0149	0.0021
	20	Schruben	0.937	0.0147	0.0021
	20	Bonferroni	0.972	0.0379	0.0058
0.8	5	Batch means	0.933	1.8637	0.6217
	5	Schruben	0.921	1.2688	0.3582
	5	Bonferroni	0.957	3.0723	1.1101
	10	Batch means	0.917	0.8817	0.1945
	10	Schruben	0.886	0.8312	0.1831
	10	Bonferroni	0.963	2.3432	0.5982
	20	Batch means	0.906	0.7484	0.1458
	20	Schruben	0.892	0.7286	0.1511
	20	Bonferroni	0.971	2.1381	0.4190

*We report the average volume's dth root and the corresponding standard deviation.

$Z_k^i = W_k^i$, and W_k is the waiting time of the kth customer. For a more detailed explanation on how to implement the standardized time series procedures used in this section, the reader is referred to §1.6.1 of Muñoz (1991).

In Table 2 we summarize the results of 1,000 independent replications, with a run length of $t = 1{,}600{,}000$ observations. In Table 3 we give the results with a larger run length ($t = 3{,}200{,}000$). In all the experiments we fixed the input rate at 1, so that the multivariate steady-state mean is $r = (0.5, 1, 3)^T$ for $\rho = 0.5$, and $r = (3.2, 25.6, 307.2)^T$ for $\rho = 0.8$.

The results of our experiments agree with the observations of §3. As we see from Tables 2 and 3, Schruben's standardized sum process with m batches gives smaller volumes than the batch means method with m batches, but for small sample sizes, the coverage of Schruben's standardized sum process with m batches is smaller than that of the batch means method with m batches.

As we expected, a larger sample size produces a better empirical coverage and a smaller volume. When we increase the traffic intensity, a larger sample size is required to obtain reliable confidence regions. We observe that, when the sample size is not large enough to produce a good empirical coverage, a small number of batches produces better results in terms of empirical coverage. This last observation is explained by the fact that a smaller number of batches leads to a larger batch size.

With respect to the confidence regions obtained from univariate confidence intervals, the Bonferroni method always gives the largest average volume and better coverages than the batch means method or Schruben's standardized sum process. When the run length is good enough to obtain good confidence intervals with the batch means method or Schruben's standardized sum process, the Bonferroni method gives overcoverage and large average volumes compared to the asymptotically valid methods.

APPENDIX

PROOF OF THEOREM 1. To prove Theorem 1 we need the following results.

LEMMA 1. *If X is a random vector distributed according to a $N_d(0, I)$ distribution, and D is a positive definite $d \times d$ matrix, then*

$$P[X^T X \leqslant (\det(D))^{1/d}] \geqslant P[X^T D^{-1} X \leqslant 1].$$

The proof of Lemma 1 is given in Muñoz (1991).

LEMMA 2. *For $g \in \mathcal{M}$, let $f(g) = KE[\det(g(B)^{1/2})] \times z(g, \alpha)^{d/2}$, where K is defined in (25). Then f is scale-invariant, that is, for $\lambda > 0$,*

$$f(\lambda g) = f(g).$$

PROOF. Since

$$P[B(1)^T(\lambda g(B))^{-1}B(1) \leqslant \lambda^{-1} z(g, \alpha)] = P[B(1)^T g(B)^{-1} B(1) \leqslant z(g, \alpha)],$$

we see that $z(\lambda g, \alpha) = \lambda^{-1} z(g, \alpha)$, so that

$$\begin{aligned} f(\lambda g) &= KE[\det(\lambda g(B))^{1/2}] z(\lambda g, \alpha)^{d/2} \\ &= K\lambda^{d/2} E[\det(g(B)^{1/2})] z^{d/2}(g, \alpha)\lambda^{-d/2} \\ &= f(g). \quad \square \end{aligned}$$

LEMMA 3. *For $g \in \mathcal{M}$, let $f(g)$ be as in Lemma 9 and $g \in \mathcal{M}$. Then*

$$f(g) \geqslant K\chi^d_{(d, \alpha)}.$$

PROOF. By Lemma 2, it suffices to consider the case $g' \in \mathcal{M}$ for which

$$z(g', \alpha) = 1. \tag{30}$$

Since $B(1)$ and $g'(B)$ are independent (Proposition 2), it follows from Lemma 7 that:

$$P[B(1)^T B(1) \leqslant \det(g'(B))^{1/d} | g'(B)] \geqslant P[B(1)^T g'(B)^{-1} B(1) \leqslant 1 | g'(B)]. \tag{31}$$

By taking expectations of both sides of (31) and using (30), we obtain:

$$P[B(1)^T B(1) \leqslant \det(g'(B))^{1/d}] \geqslant P[B(1)^T g'(B)^{-1} B(1) \leqslant 1] = 1 - \alpha.$$

Therefore,

$$P[(B(1)^T B(1))^{d/2} \leqslant \det(g'(B))^{1/2}] \geqslant 1-\alpha,$$

that is,

$$E[F(\det(g'(B))^{1/2})] \geqslant 1-\alpha, \tag{32}$$

where

$$F(y) = P[(B(1)^T B(1))^{d/2} \leqslant y] = \int_0^y f(x)dx,$$

and $f(x)$ is the density function of $(B(1)^T B(1))^{d/2}$, given by

$$f(x) = \frac{1}{d2^{d/2-1}\Gamma(d/2)} \exp\left[-\frac{x^{2/d}}{2}\right].$$

Now, since $f(x)$ is strictly decreasing on $[0, \infty)$, $F(y)$ is concave on $[0, \infty)$. Hence, from (32) and Jensen's inequality (see p. 47 of Chung 1974) we obtain

$$F(E[\det(g'(B))^{1/2}]) \geqslant 1-\alpha,$$

so that

$$P[(B(1)^T B(1))^{d/2} \leqslant E[\det(g'(B))^{1/2}]] \geqslant 1-\alpha,$$

from which we obtain $E[\det(g'(B))^{1/2}] \geqslant \chi^d_{(d,\alpha)}$. This gives the desired result. □

From Fatou's lemma (see p. 42 for Chung 1974), (22), and (24) we obtain

COROLLARY 1. *Let $g \in \mathcal{M}$. Then, under Assumption* 1,

$$\liminf_{t\to\infty} t^{d/2} E[V(g)] \geqslant K\chi^d_{(d,\alpha)},$$

where $V(g)$ is defined in (22) *and* $K = \det(G)q_d$.

The conclusion follows from (26) and Corollary 1.

ACKNOWLEDGMENTS

The authors are grateful to Bruce Schmeiser, Barry Nelson, and two anonymous referees for their detailed comments and constructive criticisms that helped to improve the presentation of this paper.

REFERENCES

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

——. 1979. *Probability and Measure*. John Wiley, New York.

Chen D. R., A. F. Seila. 1987. Multivariate inference in stationary simulation using batch means. *Proc. 1987 Winter Simulation Conf.* 302–304.

Chung, K. L. 1974. *A Course in Probability Theory*. Academic Press, New York.

Crane, M. A., A. J. Lemoine. 1977. An introduction to the regenerate method for simulation analysis. *Lecture Notes in Control and Information Sciences*, Vol. 4. Springer Verlag, New York.

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

Dykstra, R. L. 1970. Establishing the positive definiteness of the sample covariance matrix. *Ann. Math. Statist.* **41** 2153–2154.

Ethier, S. N., T. G. Kurtz. 1986. *Markov Processes, Characterization and Convergence*. John Wiley, New York.

Fishman, T. W. 1978. *Principles of Discrete Event Simulation*. John Wiley, New York.

Glynn, P. W., D. L. Iglehart. 1990. Simulation output analysis using standardized time series. *Math. Oper. Res.* **15** 1–16.

Goldsman, D., L. Schruben. 1984. Asymptotic properties of some confidence interval estimators for simulation output. *Management Sci.* **30** 1217–1225.

——, K. Kang, R. G. Sargent. 1986. Large and small sample comparisons of various variance estimators. *Proc. 1986 Winter Simulation Conf.* 278–284.

——, M. Meketon, L. Schruben. 1990. Properties of standardized time series weighted area variance estimators. *Management Sci.* **36** 602–612.

Iglehart, D. L. 1978. The Regenerative Method for Simulation Analysis. K. M. Chandy, R. T. Yeh, eds. *Current Trends in Programming Methodology; Vol III. Software Engineering*. Prentice Hall, New York, 52–71.

Law, A. M. 1983. Statistical analysis of simulation output data. *Oper. Res.* **31** 983–1029.

Lewis, P. A. W., E. J. Orav. 1989. *Simulation Methodology for Statisticians, Operations Analysts, and Engineers, Vol. I.* Wadsworth & Brooks, Cole, CA.

Muñoz, D. F. 1991. *Cancellation Methods in the Analysis of Simulation Output*, Ph.D. dissertation, Department of Operations Research, Stanford University, Stanford, California.

Royden, M. L. 1968. *Real Analysis*. Macmillan Publishing Company, New York.

Schmeiser, B. 1982. Batch size effects in the analysis of simulation output. *Oper. Res.* **30** 556–568.

Schruben, L. 1981. Control of initialization bias in multivariate simulation response. *Comm. Assoc. Comp. Mach.* **24** 246–252.

——. 1982. Detecting initialization bias in simulation output. *Oper. Res.* **31** 569–590.

——. 1983. Confidence interval estimation using standardized time series. *Oper. Res.* **31** 1090–1108.

Searle, S. R. 1982. *Matrix Algebra Useful for Statistics*. John Wiley, New York.

Siotani, M., T. Hayakawa, Y. Fujikoshi. 1985. *Modern Multivariate Statistical Analysis: a Course and Handbook*. American Science Press, Columbus, OH.