Proceedings of the 2022 Winter Simulation Conference
B. Feng, G. Pedrielli, Y. Peng, S. Shashaani, E. Song, C.G. Corlu, L.H. Lee, E.P. Chew, T. Roeder, and P. Lendermann, eds.

OVERLAPPING BATCH CONFIDENCE REGIONS ON THE STEADY-STATE QUANTILE VECTOR

Raghu Pasupathy

Department of Statistics
Purdue University
West Lafayette, IN 47906, USA

Dashi I. Singham

Operations Research Department
Naval Postgraduate School
Monterey, CA 93943, USA

Yingchieh Yeh

Institute of Industrial Management
National Central University
Taoyuan, TAIWAN

ABSTRACT

The ability to use sample data to generate confidence regions on quantiles is of recent interest. In particular, developing confidence regions for multiple quantile values provides deeper information about the distribution of underlying output data that may exhibit serial dependence. This paper presents a cancellation method that employs overlapping batch quantile estimators to generate confidence regions. Our main theorem characterizes the weak limit of the statistic used in constructing such confidence regions, showing in particular that the derived weak limit deviates from the classical multivariate Student's t and the normal distributions depending on the number of batches and the extent of their overlap. We present limited numerical results comparing the effect of fully overlapping versus non-overlapping batches to explore the tradeoff between coverage probability, confidence region volume, and computational effort.

1 INTRODUCTION

Quantiles widely serve as key summary measures of random variables describing the functioning of a system of interest, e.g., the completion time of a construction project, wait time experienced in a vehicular traffic system, or the payouts from an insurance portfolio. Quantiles are almost always *estimated* using output data generated from the system (or a simulation model of the system), making confidence bounds on quantiles of natural interest since they quantify the "uncertainty" associated with the estimated quantile. Due to their obvious utility, we say no more on motivating quantile confidence sets — see Dong and Nakayama (2020) and references therein for further discussion. We emphasize that the "data" $X_1, X_2, \ldots$ come from a time series and can exhibit serial correlation. And, while variance reduction methods can be used as in Chu and Nakayama (2012), Nakayama (2014), Dong and Nakayama (2020) and the numerous other references therein, we do not employ these methods here.

This paper develops what is identified in existing literature as *cancellation methods* to calculate confidence regions on a steady state quantile vector. Motivated by corresponding success in the steady state mean context, the statistic that we consider here (in the service of confidence region construction) is formed from overlapping batches of data. Our main result characterizes the weak limit of such an overlapping batch statistic, in particular demonstrating its dependence on the asymptotic batch size and the extent of batch overlap, and its deviation from the classical Student's t and normal random variables. Limited numerical experience that we present re-affirms the advantages of using large overlapping batches.

978-1-6654-7661-4/22/$31.00 ©2022 IEEE

In Section 2, we present the key problem of determining confidence regions for multiple quantiles. Section 3 reviews related literature. Section 4 presents the main theorem and approach for deriving confidence regions using overlapping batches, and Section 5 concludes with numerical results.

2 NOTATION AND CONFIDENCE REGION DEFINITION

(i) $\mathbb{N}$ refers to the set $\{1,2,\ldots,\}$ of natural numbers. (ii) $\mathbb{I}_A(x)$ is the indicator variable taking the value 1 if $x \in A$ and 0 otherwise. Also, depending on the context, we write $\mathbb{I}(A)$ where $\mathbb{I}(A) = 1$ if the event A is true and 0 otherwise. (iii) I_d refers to the $d \times d$ identity matrix and $\mathscr{M}_d^+$ to the space of symmetric positive-definite matrices. We write A_j and $B_{i,j}$ to refer to the j-th element of the vector A and the (i,j)-th element of the matrix B, respectively. (iv) For a $d \times d$ symmetric positive-definite matrix A, $\sqrt{A}$ refers to a $d \times d$ positive definite matrix that satisfies $\sqrt{A}\sqrt{A} = A$. It is known that a $d \times d$ matrix A is positive definite if and only if there exists a positive definite matrix $\sqrt{A}$ such that $\sqrt{A}\sqrt{A} = A$. (iv) $\|x\|_p, p \geq 1$ refers to the L_p norm $\left(\sum_{j=1}^d |x_j|^p\right)^{1/p}$ of the vector $x \in \mathbb{R}^d$. We use the special notation $\|x\|$ to refer to the L_2 norm. (v) For a $d \times d$ matrix B, $|B|$ refers to its determinant. (vi) $N(0, I_d)$ denotes the standard normal random vector in d dimensions, and χ_ν^2 refers to the chi-square random vector with ν degrees of freedom. (vii) For a random sequence $\{X_n, n \geq 1\}$, we write $X_n \overset{\text{wp1}}{\to} X$ for almost sure convergence, $X_n \overset{\text{p}}{\to} X$ for convergence in probability, and $X_n \overset{\text{d}}{\to} X$ for convergence in distribution (or weak convergence). (viii) $\sigma(X_1, X_2, \ldots, X_n)$ refers to the σ-algebra formed by the random variables $X_1, X_2, \ldots, X_n$. (ix) Let $B(\xi, \delta)$ be a ball centered at ξ with radius δ. (x) For random variables X_i from a stationary stochastic process, let f denote the (marginal) probability density function and F the cumulative distribution function (cdf) of X_i. Additionally, define $F'(x) = f(x)$ and $F''(x)$ to be the first and second derivatives of the cdf. The empirical cdf F_n and the sample quantile estimator Q_n are constructed from $X_j, j = 1, 2, \ldots, n$ as follows:

$$F_n(x) := \frac{1}{n}\sum_{j=1}^{n} \mathbb{I}_{(-\infty,x]}(X_j), \qquad x \in \mathbb{R};$$

$$F_n^{-1}(\eta_i) := \min\{x : F_n(x) \geq \eta_i\}, \quad \eta_i \in [0,1];$$

and the sectioning estimator of the η-quantile is

$$Q_n(\eta) = (F_n^{-1}(\eta_1), F_n^{-1}(\eta_2), \ldots, F_n^{-1}(\eta_d)). \tag{1}$$

In the treatment that follows, we slightly abuse notation and use the same notation (for η) irrespective of whether η is a scalar or a vector. This should cause no confusion since the dimension of η will be clear from the context. Finally, we define ϕ-mixing (Ethier and Kurtz 2009, pp. 59) which represents the type of dependence considered in this paper.

Definition 1 For a strictly stationary stochastic process $\{X_n, n \geq 1\}$, denote

$$\mathscr{A}_1^n := \sigma(X_1, X_2, \ldots, X_n); \quad \mathscr{A}_n^\infty := \sigma(X_n, X_{n+1}, \ldots,),$$

and

$$\phi_n := \sup\left\{|P(A\,|\,B) - P(B)| : A \in \mathscr{A}_1^k, B \in \mathscr{A}_{k+n}^\infty, P(B) > 0\right\}.$$

The process $\{X_n, n \geq 1\}$ is said to be ϕ-mixing if $\phi_n \to 0$ as $n \to \infty$.

Next, we define the confidence region problem. Let $\{X_n, n \geq 1\}$ be a real-valued discrete-time stationary stochastic process, and let

$$\xi := (\xi_1, \xi_2, \ldots, \xi_d); \quad \xi_i := \inf\{x : F(x) \geq \eta_i\}, \quad 0 < \eta_1 < \eta_2 \cdots < \eta_d < 1$$

denote a vector of quantiles associated with F. We seek a method to construct a $(1-\alpha)$-confidence region on ξ, that is, given $\alpha \in (0,1)$, we seek an *ellipsoid* $C_n \subset \mathbb{R}^d$ constructed from the initial segment of data $\{X_j, 1 \leq j \leq n\}$ such that $P(\xi \notin C_n) \to \alpha$ as $n \to \infty$.

Also, observe that we have assumed $\{X_n, n \geq 1\}$ is a real-valued process. More generally, one might assume that $\{X_n, n \geq 1\}$ is an S-valued stationary process, $\theta_k : S \to \mathbb{R}, k = 1,2,\ldots,d$ are functionals such that $F = (F_1, F_2, \ldots, F_d)$ is the distribution function associated with $(\theta_1(X_1), \theta_2(X_1), \ldots, \theta_d(X_1)) \in \mathbb{R}^d$, and the quantiles are

$$\xi_j := \inf\{x \in \mathbb{R} : F_j(x) \geq \eta_j\}.$$

The confidence set construction problem is then that of finding a region

$$C_n := \left\{ y \in \mathbb{R}^d : \|A_n(y - Q_n(\eta))\|_p \leq R_n \right\}, p \geq 1 \tag{2}$$

satisfying, for fixed $\alpha \in (0,1)$,

$$\lim_{n\to\infty} P(\xi \in C_n) = 1 - \alpha,$$

where

$$Q_n(\eta) \in \mathbb{R}^d, R_n \in \mathbb{R}, A_n \in \mathscr{M}_d^+ \text{ and } Q_n(\eta), R_n, A_n \in \sigma(X_1, X_2, \ldots, X_n).$$

Different choices of A_n and p in (2) correspond to different shapes of the confidence region. In this paper, we focus on the case where $p = 2$ and A_n is a symmetric positive-definite matrix, resulting in elliptical confidence regions.

3 LITERATURE

If the sequence $\{X_j, j \geq 1\}$ is ϕ-mixing (see Definition 1), $F(\xi_i) = \eta_i, F'(\xi_i) > 0, 1 \leq i \leq d$, and $\exists \kappa > 0$ and $\delta > 0$ such that $|F''(x)| \leq \kappa$ for $x \in \cup B(\xi_i, \delta)$, then it follows, e.g., using Theorem 2, that for $\eta = (\eta_1, \eta_2, \ldots, \eta_d) \in (0,1)^d$,

$$\sqrt{n}\Sigma^{-1/2}(Q_n(\eta) - \xi) \xrightarrow{\text{d}} N(0,I); \quad \Sigma_{i,j} = \frac{\min(\eta_i, \eta_j) - \eta_i\eta_j}{F'(\xi_i)F'(\xi_j)}. \tag{3}$$

Virtually all existing techniques for constructing confidence regions on quantiles directly or indirectly exploit a central limit theorem (CLT) such as (3). Thus, a useful way of categorizing existing methods for constructing quantile confidence regions is based on how the CLT in (3) is exploited, giving rise to *consistent* and *cancellation* methods.

3.1 Consistent Methods

Consistent methods construct a consistent estimator Σ_n of the variance constant Σ in (3) implying that a simple application of Slutsky's theorem (Serfling 2009) allows us to construct the valid $(1-\alpha)$ elliptical confidence region

$$\left\{ y \in \mathbb{R}^d : n \left\| \sqrt{\Sigma_n}^{-1} (Q_n(\eta) - y) \right\|^2 \leq \chi^2_{d,1-\alpha} \right\},$$

where $\chi^2_{d,1-\alpha} := \min\{x : P(\chi^2_d \leq x) \geq 1-\alpha\}$ is the $(1-\alpha)$ critical value of the the chi-square distribution with d degrees of freedom. While such an approach is attractive due to its simplicity, as Glynn (1996) and Chu and Nakayama (2012) note, constructing a consistent estimator (an estimator Σ_n satisfying $\Sigma_n \xrightarrow{\text{p}} \Sigma$) is a challenge since it entails consistently estimating the density $F'(\xi_i), 1 \leq i \leq d$, and further exacerbated since $F'(\xi)$ appears in the denominator of Σ_n causing its variance to diverge as $\eta_i \to 1$. This challenge has led to various methods aimed specifically at consistent estimation of Σ in the service of constructing confidence regions. For instance, see Chu and Nakayama (2012) for consistent finite-difference estimators of the reciprocal of the density, and the large literature dedicated to the question of estimating the density or its reciprocal. Lei et al. (2020) and Lei et al. (2022) are more recent examples that use a generalized likelihood ratio estimator of the density to construct a consistent method.

3.2 Cancellation Type Methods

A crucial insight is that it is not necessary to consistently estimate Σ in order to construct a valid confidence region on the quantile vector ξ. This fact is exploited by cancellation methods (Glynn and Iglehart 1985; Glynn and Iglehart 1990) which explicitly or implicitly construct a process $\{S_n, n \geq 1\}, S_n \in \mathscr{M}_d^+$ such that

$$(\sqrt{n}(Q_n(\eta) - \xi), S_n^2) \xrightarrow{\text{d}} (\sqrt{\Sigma} N(0, I_d), \Sigma S^2) \text{ as } n \to \infty, \tag{4}$$

where $S^2 \in \mathscr{M}_d^+$ is a $d \times d$ symmetric positive definite random matrix whose distribution can be computed. This implies, among other things, that S does not depend on the unknown quantities Σ and ξ. Under (4), the continuous mapping theorem (Billingsley 1999) allows "cancelling" the unknown Σ:

$$(\sqrt{n}\sqrt{S_n^2}^{-1}(Q_n(\eta) - \xi) \xrightarrow{\text{d}} \left(\cancel{\sqrt{\Sigma}}\sqrt{S^2}\right)^{-1} \cancel{\sqrt{\Sigma}} N(0, I_d) \stackrel{\text{d}}{=} S^{-1} N(0, I_d), \tag{5}$$

giving rise to the asymptotically valid $(1-\alpha)$ elliptical confidence set

$$\left\{ y : n \| \sqrt{S_n^2}^{-1}(Q_n(\eta) - y) \|^2 \leq \tilde{t}_{1-\alpha}^2 \right\}, \tag{6}$$

where $\tilde{t}_{1-\alpha}^2$ is the $(1-\alpha)$ quantile of $\|S^{-1}N(0, I_d)\|^2$ such that $\tilde{t}_{1-\alpha}^2 := \min\left\{t : P(\|S^{-1}N(0, I_d)\|^2 \leq t) \geq 1-\alpha\right\}$. Of course, the choice of S_n^2 is not unique and this constitutes both the challenge and the room for novelty within cancellation methods. Also, in arriving at (6), while nothing has been assumed about the independence between S and $N(0, I_d)$, common choices of S_n will lead to their independence.

A first approach to using batching for quantile estimation for Markov chain output appears in Muñoz (2010). Another application of the cancellation method to the construction of confidence regions for quantiles appears in Calvin and Nakayama (2013), where the authors assume that the following *functional central limit theorem* is in effect:

$$\frac{\lfloor nt \rfloor}{\sqrt{n}} \left(Q_{\lfloor nt \rfloor}(\eta) - \xi)\right) \xrightarrow{\text{d}} \tau_\eta W(t), \ 0 \leq t \leq 1, \tag{FCLT}$$

where

$$\tau_\eta := \frac{\sqrt{\eta(1-\eta)}}{F'(\xi)}; \quad \xi := \inf\{x \in \mathbb{R} : F(x) \geq \eta\}; \quad \eta \in (0,1),$$

and the weak convergence is in $D[0,1]$ endowed with the Skorokhod metric. Calvin and Nakayama (2013) argue that under (FCLT), it can be shown that

$$\left(\sqrt{n}(Q_n(\eta) - \xi), STS_n\right) \xrightarrow{\text{d}} (\tau_\eta W(1), \tau_\eta B), \tag{7}$$

where the *standardized time series* (Schruben 1983) $\{STS_n, n \geq 1\}$ is defined as

$$STS_n := \frac{\lfloor nt \rfloor}{\sqrt{n}} \left(Q_{\lfloor nt \rfloor}(\eta) - Q_n(\eta)\right) \in D[0,1].$$

The importance of (7) is that it allows for construction of a functional of STS_n that can then be used to obtain an analogue of S_n in (4) toward constructing a confidence region using (5) and (6). For example, when $d = 1$, Calvin and Nakayama (2013) argue, based on methods introduced in Alexopoulos et al. (2007) for the steady-state mean context, that the weak limit S appearing in (5) is the chi-square random variable with one degree of freedom (χ_1^2) when S_n is chosen as the *weighted area estimator* (Goldsman et al. 1990) of Σ:

$$S_n := \left(\frac{1}{n}\sum_{j=1}^{n} w\left(\frac{j}{n}\right) STS_n\left(\frac{j}{n}\right)\right)^2,$$

where the weighting function $w : [0,1] \to \mathbb{R}$ is twice differentiable and $\mathbb{E}\left[\int_0^1 w(t)B(t)\,dt\right] = 1$. Calvin and Nakayama (2013) also provide expressions for the weak limit S when the weighted area estimator is constructed using nonoverlapping and overlapping batches. Alexopoulos et al. (2020) extend the work of Calvin and Nakayama (2013) from i.i.d. data to dependent data relying on a geometric moment contraction condition (GMC) (Wu 2005).

Prior to Calvin and Nakayama (2013), Alexopoulos et al. (2012) present an application of a technique similar to cancellation (in the sense of not needing consistent estimation of the variance parameter) toward the construction of a confidence interval on a quantile associated with the steady-state distribution of a real-valued discrete time stochastic process. Specifically, the authors use a fixed number ($b < \infty$) of non-overlapping batches to construct the batch quantile estimators $Q_{j,m}(\eta) := F_{j,m}^{-1}(\eta), F_{j,m}(x) = m^{-1}\sum_{k=(j-1)m+1}^{jm} \mathbb{I}_{(-\infty,x)}(X_k), j = 1,2,\ldots,b; m = n/b$, and then crucially demonstrate under GMC (in lieu of strong mixing) that as $n \to \infty$,

$$\sqrt{m}\,(Q_{1,m}(\eta)-\xi, Q_{2,m}(\eta)-\xi,\ldots,Q_{b,m}(\eta)-\xi) \quad \xrightarrow{\text{d}} \quad N\left(0, \frac{\eta(1-\eta)\sum_{\ell=-\infty}^{\ell=\infty}\rho_\ell}{f^2(\xi)} I_b\right), \tag{8}$$

where $\rho_\ell = \mathrm{Corr}(\mathbb{I}(X_1 \le \xi), \mathbb{I}(X_{\ell+1} \le \xi))$ is the lag-ℓ correlation associated with the process $\{\mathbb{I}(X_j \le \xi), j \ge 1\}$. If $\bar{Q}_n(\eta) = b^{-1}\sum_{j=1}^{b} Q_{j,m}(\eta)$ is the batching estimator and $\bar{S}_n^2$ is the sample variance constructed from $Q_{j,m}, j = 1,2,\ldots,b$, then under (8), the continuous mapping theorem (Billingsley 1999) assures us that $\sqrt{b}(\bar{Q}_n - \xi)/\bar{S}_n$ converges weakly to the Student's t random variable with $b-1$ degrees of freedom, yielding the $(1-\alpha)$ confidence interval $\bar{Q}_n \pm t_{\alpha/2,b-1}\bar{S}_n/\sqrt{b}$. Sequest (Alexopoulos et al. 2019) and Sequem (Alexopoulos et al. 2017) incorporate notable implementation enhancements to the essential idea introduced in Alexopoulos et al. (2012).

The idea presented in Alexopoulos et al. (2012) is of particular relevance to what we present here. In fact, the main theorem that we present can be seen as replacing the batching estimator $\bar{Q}_n$ used in Alexopoulos et al. (2012) with the sectioning estimator Q_n, and then generalizing along the following three directions: (i) allowing any degree of batch overlap ranging from fully overlapping to non-overlapping; (ii) $\xi \in \mathbb{R}^d$ implying that the confidence regions reside in an arbitrarily high but finite dimension; and (iii) number of batches b can be finite or infinite depending on the extent of batch overlap. As our main theorem will make clear, (i), (ii), and (iii) cause deviations from the classical Student's t weak limit, and thereby the nature of the constructed confidence region.

3.3 Bahadur Representations

In this subsection we summarize two strong approximation theorems that are crucially invoked when proving the main results of this paper. Bahadur's remarkable result, now known informally as the Bahadur representation, presents an almost sure bound on the rate at which the error in the sample quantile decays to zero as $n \to \infty$, while Sen extends the results to ϕ-mixing random variables.

Theorem 1 (Bahadur Representation for i.i.d. data, see Bahadur (1966)) Suppose (i) $F(\xi) = \eta$, (ii) F is twice differentiable at ξ, (iii) $F'(\xi) > 0$, and (iv) $\exists \kappa < \infty$ such that $|F''(x)| < \kappa$ for all $x \in B(\xi,\delta)$ and some $\delta > 0$. Then

$$\left|\sqrt{n}\{f(\xi)(Q_n(\eta)-\xi) - (\eta - F_n(\xi))\}\right| = O(n^{-1/4}\log n) \quad \text{a.s.}$$

Theorem 2 (Bahadur Representation Under ϕ-Mixing, see Sen (1972)) Suppose $\{X_j, 1 \le j \le n\}$ is ϕ-mixing with constants $\phi(j) \ge 0$ satisfying $1 \ge \phi(1) \ge \phi(2) \ge \cdots \ge 0$, and $\sum \phi^{1/2}(j) < \infty$. Suppose F is absolutely continuous in some neighborhood of ξ with a continuous density function f such that $0 < f(\xi) < \infty$. Furthermore, suppose $f'(x) = (d/dx)f(x)$ is positive and bounded in some neighborhood of ξ. Then,

$$\left|\sqrt{n}\{f(\xi)\,(Q_n(\eta)-\xi) - (\eta - F_n(\xi))\}\right| = O(n^{-1/8}\log n) \quad \text{a.s.}$$

While Theorem 2 is our essential instrument to handle dependence, we could have equally used a number of other Bahadur representations (Wu 2005) that have appeared recently.

4 MAIN THEOREM

An estimator Σ_n of the $d \times d$ asymptotic covariance matrix Σ appearing in (3) can be constructed using overlapping batches of data as follows. Suppose we partition the data into possibly overlapping batches of size m_n and having initial observations offset by an amount d_n. (See Figure 1 for a clear idea.) The i-th batch consists of data $X_j, j = (i-1)d_n+1, (i-1)d_n+2, \ldots, (i-1)d_n+m_n$ where $i = 1,2,\ldots,b_n$ and the number of batches $b_n = d_n^{-1}(n-m_n)+1$. The empirical distribution from the i^{th} batch is then

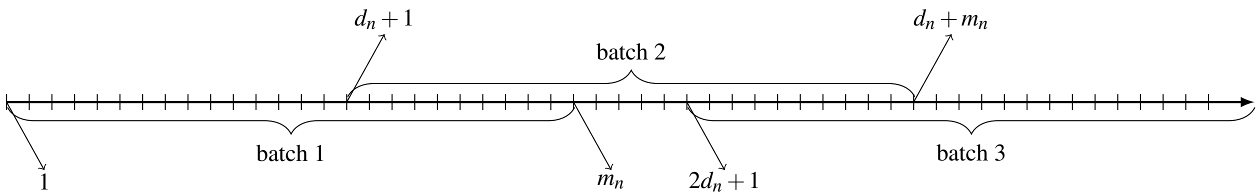

Figure 1: The figure depicts partially overlapping batches. Batch 1 consists of observations $X_j, j = 1,2,\ldots,m_n$; batch 2 consists of observations $X_j, j = d_n+1, d_n+2, \ldots, d_n+m_n$, and so on, with batch i consisting $X_j, j = (i-1)d_n+1, (i-1)d_n+2, \ldots, (i-1)d_n+m_n$. There are thus $b_n := d_n^{-1}(n-m_n)+1$ batches in total, where n is the size of the dataset.

$$F_{i,m_n}(x) = \frac{1}{m_n} \sum_{k=(i-1)d_n+1}^{(i-1)d_n+m_n} \mathbb{I}_{(-\infty,x]}(X_k),$$

yielding the quantile estimators constructed from the various batches:

$$Q_{i,m_n} := (F_{i,m_n}^{-1}(\eta_1), F_{i,m_n}^{-1}(\eta_2), \ldots, F_{i,m_n}^{-1}(\eta_d)), \quad i = 1,2,\ldots,b_n. \tag{9}$$

The sectioning estimator in (1) and the batch quantiles in (9) together suggest the following natural estimator Σ_{m_n} of Σ:

$$\Sigma_{m_n} := \frac{1}{1-(m_n/n)} \frac{m_n}{b_n} \sum_{j=1}^{b_n} (Q_{j,m_n}(\eta) - Q_n(\eta))(Q_{j,m_n}(\eta) - Q_n(\eta))^T. \tag{10}$$

The factor $(1-m_n/n)^{-1}$ ensures that the estimator Σ_{m_n} is asymptotically unbiased. Also define the corresponding "Studentized random vector"

$$T_{m_n} := \sqrt{n}\, \Sigma_{m_n}^{-\frac{1}{2}} \, (Q_n(\eta) - \xi).$$

We are now ready to state the main result that characterizes the weak convergence behavior of the matrix sequence $\{\Sigma_{m_n}, n \geq 1\}$ and the vector sequence $\{T_{m_n}, n \geq 1\}$.

Theorem 3 (OB Limits $\chi^2_{\text{OB}}(\beta, b)$ and $T_{\text{OB}}(\beta, b)$) Suppose that the postulates of Theorem 2 hold and that

$$\beta := \lim_n \frac{m_n}{n} > 0; \quad b := \lim_n b_n \in \{2,3,\ldots,\infty\}.$$

Define the "Brownian bridge" random vector

$$B(u,\beta) := W_d(u+\beta) - W_d(u) - \beta W_d(1), \quad u \in [0,1-\beta], \beta \in (0,1),$$

where $\{W_d(t), 0 \leq t \leq 1\}$ is the d-dimensional standard Wiener process. Then the sequences $\{\Sigma_{m_n}, n \geq 1\}, \{T_{m_n}, n \geq 1\}$ satisfy

$$\Sigma_{m_n} \xrightarrow{\text{d}} \sqrt{\Sigma}\, \chi^2_{\text{OB}}(\beta, b)\, \sqrt{\Sigma}^T; \quad T_{m_n} \xrightarrow{\text{d}} \chi^{-1}_{\text{OB}}(\beta, b)\, W_d(1) =: T_{\text{OB}}(\beta, b), \tag{11}$$

where

$$\chi^2_{\text{OB}}(\beta, b) := \begin{cases} \frac{1}{\kappa_1(\beta,b)} \frac{1}{\beta(1-\beta)} \int_0^{1-\beta} B(u,\beta) B(u,\beta)^T \, \text{d}u & b = \infty \\ \frac{1}{\kappa_1(\beta,b)} \frac{1}{\beta} \frac{1}{b} \sum_{j=1}^{b} B(c_j,\beta) B(c_j,\beta)^T & b \in \mathbb{N} \backslash \{1\}, \end{cases}$$

$c_j = (j-1)(1-\beta)/(b-1)$, and $\chi^{-1}_{\text{OB}}(\beta, b) := \left(\chi^2_{\text{OB}}(\beta, b)\right)^{-\frac{1}{2}}$, and where $\kappa_1(\beta, b) = (1-\beta)$.

Proof Sketch. Recall Equation (10) and substitute $\beta = m_n/n$. Since Σ is a variance matrix there exists $\sqrt{\Sigma}$ such that $\Sigma = \sqrt{\Sigma}\sqrt{\Sigma}^T$ and define

$$\tilde{B}_{j,m_n} := m_n^{-1} \left(W_d\left((j-1)\frac{n-m_n}{b_n - 1} + m_n \right) - W_d\left((j-1)\frac{n-m_n}{b_n-1} \right) \right), \quad j = 1, 2, \ldots, b_n.$$

Then,

$$(1-\beta)\Sigma_{m_n} = \frac{m_n}{b_n} \sum_{j=1}^{b_n} \left[(Q_{j,m_n}(\eta) - Q_n)(Q_{j,m_n}(\eta) - Q_n)^T - (\sqrt{\Sigma}\tilde{B}_{j,m_n} - \sqrt{\Sigma} n^{-1} W_d(n))(\sqrt{\Sigma}\tilde{B}_{j,m_n} - \sqrt{\Sigma} n^{-1} W_d(n))^T \right]$$

$$+ \sqrt{\Sigma} \times \left(\frac{1}{b_n} \sum_{j=1}^{b_n} \left(\sqrt{m_n}\tilde{B}_{j,m_n} - \frac{\sqrt{m_n}}{n} W_d(n) \right) \left(\sqrt{m_n}\tilde{B}_{j,m_n} - \frac{\sqrt{m_n}}{n} W_d(n) \right)^T \right) \times \sqrt{\Sigma}^T =: E_n + L_n.$$

Using Theorem 2, we can show after lots of algebra that $E_n \xrightarrow{\text{wp1}} 0$, and that L_n converges weakly to the appropriate limit as $b_n \to b = \infty$. A similar calculation holds for the $b < \infty$ context. □

4.1 Further Observations

A number of points related to Theorem 3 are salient.

(a) Notice that $\chi^2_{\text{OB}}(\beta, b) \in \mathscr{M}_d^+$ is a random matrix and $T_{\text{OB}}(\beta, b) := \chi^{-1}_{\text{OB}}(\beta, b) W(1) \in \mathbb{R}^d$ is a random vector. They should be seen as the χ^2 and Student's t analogues for the context of constructing confidence regions on objects other than the population mean.

(b) The matrix Σ_{m_n} does not consistently estimate the covariance matrix Σ. This is due to the increased variance stemming from the use of large batches as connoted by $\beta > 0$. It is in this sense that we can "get away with" using large batches and not estimating the covariance matrix consistently. It can be shown that if $\beta = 0$, that is, if $m_n/n \to 0$, then $\chi^2_{\text{OB}}(\beta, b)$ degenerates to the identity matrix and Σ_{m_n} consistently estimates Σ.

(c) Depending on the values of the limiting batch size β and the limiting number of batches b, the random vector $T_{\text{OB}}(\beta, b)$ can deviate quite significantly from the standard normal random vector or the Student's t random vector T_ν. For this reason, it is generally not advisable to substitute $T_{\text{OB}}(\beta, b)$ with standard normal or Student's t critical values in an attempt at approximation. To facilitate using $T_{\text{OB}}(\beta, b)$ as is, code that generates critical values for $\|T_{\text{OB}}(\beta, b)\|_p, p \geq 1$ is available upon request.

(d) The form of the covariance estimator in (10) implies that $O(d^2)$ elements of the covariance matrix need to be estimated. As we show in a forthcoming paper, by "pushing" the underlying correlation structure to the right-hand side, we can construct an statistic that needs to estimate only the diagonal elements of the covariance matrix. In such a case, the weak limit on the right-hand side of (11) will involve a two-parameter Gaussian process called the Kiefer process.

The section estimator Q_n in Theorem 3 can be replaced by what has been called the *batching estimator* (Nakayama 2014):

$$\bar{Q}_n := \frac{1}{b_n}\sum_{j=1}^{b_n} Q_{j,m_n}(\eta).$$

The batching estimator $\bar{Q}_n$ has a higher bias than the sectioning estimator Q_n, and recent analysis by He and Lam (2021) suggests sectioning estimators may have better coverage when there are a large number of batches. The batching estimator and the batch quantiles in (9) together suggest the following alternative to Σ_{m_n} when estimating Σ:

$$\bar{\Sigma}_{m_n} := \frac{1}{\kappa_2(\beta,b)}\frac{m_n}{b_n}\sum_{j=1}^{b_n}(Q_{j,m_n}(\eta)-\bar{Q}_n)(Q_{j,m_n}(\eta)-\bar{Q}_n)^T,$$

where $\kappa_2(\beta,b)$ is the "bias correction" factor. A theorem analogous to Theorem 3 but with $\bar{Q}_n$ replacing $Q_n(\eta)$, and with $\bar{\Sigma}_{m_n}$ replacing Σ_{m_n}, yields parallel results which will be discussed in a forthcoming paper.

4.2 The OB Confidence Ellipsoid

Theorem 3 can be directly used to construct an elliptical $(1-\alpha)$ confidence region on ξ. To see this, note that (11) implies

$$\sqrt{n}\sqrt{\Sigma_{m_n}}^{-1}(Q_n-\xi)\xrightarrow{\text{d}}\sqrt{\chi^2_{\text{OB}}(\beta,b)}^{-1}W_d(1). \tag{12}$$

Using the continuous mapping theorem (Billingsley 1999) on (12) with the mapping function $g(x)=\|x\|^2$, we have

$$n\left\|\sqrt{\Sigma_{m_n}}^{-1}(Q_n-\xi)\right\|^2\xrightarrow{\text{d}}\left\|\sqrt{\chi^2_{\text{OB}}}^{-1}W_d(1)\right\|^2,$$

yielding the $(1-\alpha)$ elliptical confidence set

$$C_n := \left\{y\in\mathbb{R}^d : \sqrt{n}\left\|\sqrt{\Sigma_{m_n}}^{-1}(Q_n-y)\right\|\le t_{\text{OB},2,1-\alpha}\right\}$$

with root half-volumes of

$$\text{vol}(C_n)^{1/d}=\frac{1}{2}\frac{1}{\sqrt{n}}\frac{\sqrt{\pi}}{\Gamma\left(\frac{d}{2}+1\right)^{1/d}}t_{\text{OB},2,1-\alpha}|\Sigma_{m_n}|^{1/2d} \tag{13}$$

where $t_{\text{OB},2,1-\alpha}$ is the $(1-\alpha)$ quantile of $\|T_{\text{OB}}\|=\left\|\sqrt{\chi^2_{\text{OB}}}^{-1}W_d(1)\right\|$. The values of (13) are reported in the tables in Section 5.

5 EXPERIMENTAL RESULTS

This section presents the results of numerical experiments to explore the effects of increasing batch sizes and the effect of overlapping batches. For sample size n, the batch size is $m_n=\beta n$. For nonoverlapping batches, the number of batches is $b_n=n/m_n$ and the limiting number of batches as $n\to\infty$ is $b=\beta^{-1}$. For overlapping batches, we take the maximum level of overlap, so set $d_n=1$ and set $b_n=n(1-\beta)+1$. In this case, the limiting number of batches as $n\to\infty$ is $b=\infty$. We report the coverage of confidence intervals when estimating a single quantile, as well as the size of the confidence interval regions when $d\ge 1$. Note that we use our same estimator for nonoverlapping batches as for overlapping batches (varying the parameter d_n to adjust the level of overlap). Future work will compare the effects of numerous other nonoverlapping batch estimators in the literature that employ alternative cancellation or consistent estimation methods.

5.1 IID Exponential Data

We use i.i.d. values of the exponential distribution with rate 1 to test the performance of the overlapping batch means method. Table 1 shows the results calculating a single quantile using both nonoverlapping and overlapping batch means. We present the coverage and mean half-width estimate from 2000 independent replications for each experiment, using (13) to calculate half-widths and root half-volumes.

Table 1: Independent data, $d = 1$: Coverage values for 95% confidence intervals with average half-widths in parentheses for the quantile estimate for $p = 0.99$ of i.i.d. exponential data using 2000 independent replications.

	NOLB			
n	$\beta = 0.01$	$\beta = 0.05$	$\beta = 0.10$	$\beta = 0.20$
100	0.594 (0.764)	0.840 (1.297)	0.885 (1.586)	0.917 (2.084)
200	0.693 (0.661)	0.865 (1.008)	0.895 (1.197)	0.930 (1.577)
1,000	0.830 (0.415)	0.932 (0.598)	0.935 (0.669)	0.950 (0.797)
	OLB			
n	$\beta = 0.01$	$\beta = 0.05$	$\beta = 0.10$	$\beta = 0.20$
100	0.587 (0.758)	0.816 (1.244)	0.891 (1.501)	0.884 (1.770)
200	0.677 (0.658)	0.849 (0.968)	0.892 (1.113)	0.895 (1.342)
1,000	0.823 (0.414)	0.917 (0.572)	0.936 (0.640)	0.929 (0.697)

Table 1 shows that increasing the value of β yields greater coverage and greater halfwidths. Meanwhile, the OLB approach has slightly worse coverage, but narrower half-widths. Using the same data length n, overlapping batches with a larger β may deliver adequate coverage with smaller half-widths compared to using NOLB methods. We observe that for small $\beta = 0.01$, the performance of nonoverlapping and overlapping are similar, and the discrepancy increases with larger batches as the effect of the overlap becomes greater.

Next, we explore simultaneous confidence intervals for multiple quantiles of i.i.d. data where a confidence region is generated. Table 2 displays these results for the exponential distribution (with rate 1) for simultaneous confidence intervals for $p_i = 0.01, 0.30, 0.50, 0.70, 0.99$, with dimension $d = 5$. We use the estimator Σ_{m_n} as the estimate of the covariance matrix Σ and report the mean observed root half-volume of the multidimensional confidence ellipsoid. Table 2 reveals that using overlapping batches appears to reduce the volume of the confidence regions. Values marked N/A imply the batch size was not large enough to generate a d-dimensional estimate. Larger confidence interval volumes for larger β may be related to an observation in Corollary 4.16 of Glynn and Iglehart (1990), in which consistent estimators achieve lower expected half-widths for the steady state mean than cancellation estimators such as standardized time series.. We note some overcoverage for large batch sizes (large β when n is large and $d > 1$) when the sample size n is not large enough, for example, in the NOLB results coverage is high when $n = 1,000$ but decreases towards nominal as n increases. The OLB results take much longer to compute so we cannot see the convergence to nominal given computing limits.

5.2 Autoregressive Data

We simulate values of the AR(1) autoregressive process with lag 1 and coefficient $\alpha = 0.5$. In this case, the marginal distribution of the data series is $\mathcal{N}(0, \sigma^2/(1-\alpha^2))$ where σ^2 is the variance of the noise terms and is set to 1 in our experiments. Table 3 displays the results for confidence intervals calculated for a single quantile of 0.90. As before, we observe narrower confidence interval half-widths for OLB resulting from the greater number of batches. Coverage improves as the batch size increases and as n increases.

Table 2: Independent data, $d = 5$: Coverage values for 95% confidence intervals with mean root half-volumes in parentheses for the joint quantile estimate for $p_i = 0.01, 0.30, 0.50, 0.70, 0.99$ of i.i.d. exponential data using 2000 independent replications.

	NOLB		
n	$\beta = 0.01$	$\beta = 0.05$	$\beta = 0.10$
100	N/A	0.799 (0.287)	0.915 (0.383)
200	N/A	0.876 (0.189)	0.951 (0.272)
1,000	0.795 (0.073)	0.956 (0.082)	0.967 (0.112)
10,000	0.954 (0.022)	0.956 (0.025)	0.961 (0.035)
100,000	0.950 (0.007)	0.951 (0.008)	0.957 (0.011)
1,000,000	0.949 (0.002)	0.947 (0.002)	0.952 (0.003)
	OLB		
n	$\beta = 0.01$	$\beta = 0.05$	$\beta = 0.10$
100	N/A	0.814 (0.275)	0.897 (0.286)
200	N/A	0.861 (0.178)	0.924 (0.199)
1,000	0.828 (0.072)	0.960 (0.077)	0.968 (0.082)
10,000	0.952 (0.021)	0.956 (0.023)	0.966 (0.025)

Table 3: Dependent data, $d = 1$: Coverage values for 95% confidence intervals with mean half-widths in parentheses for the joint quantile estimate for $p = 0.90$ with AR(1) data using 2000 independent replications.

	NOLB			
n	$\beta = 0.01$	$\beta = 0.05$	$\beta = 0.10$	$\beta = 0.20$
100	0.830 (0.370)	0.893 (0.458)	0.918 (0.519)	0.937 (0.663)
200	0.875 (0.289)	0.899 (0.342)	0.930 (0.395)	0.938 (0.488)
1,000	0.905 (0.145)	0.952 (0.173)	0.943 (0.189)	0.959 (0.225)
	OLB			
n	$\beta = 0.01$	$\beta = 0.05$	$\beta = 0.10$	$\beta = 0.20$
100	0.820 (0.369)	0.898 (0.443)	0.920 (0.493)	0.939 (0.573)
200	0.877 (0.291)	0.898 (0.331)	0.929 (0.375)	0.935 (0.428)
1,000	0.919 (0.146)	0.941 (0.168)	0.953 (0.180)	0.955 (0.194)

Table 4 shows results for $d = 3$ where the goal is to estimate simultaneously the 0.90, 0.95, and 0.99 quantiles for AR(1) dependent data. We observe higher sample sizes are needed for higher dimensional settings to achieve adequate coverage. While the overlapping approach with large batch sizes appear to achieve better coverage with smaller sample sizes, we note the computation time is much higher than for nonoverlapping interval estimation.

ACKNOWLEDGEMENTS

The authors are grateful to the anonymous referees of this paper, who suggested many comments that greatly improved the content and presentation of this work. The first author thanks the Department of Computer Science and Engineering, IIT Madras, Chennai, INDIA and the Operations Research Department, Naval Postgraduate School, Monterey, CA for the hospitality during his sabbatical visit in 2021–2022.

Table 4: Dependent data, $d = 3$: Coverage values for 95% confidence intervals with mean half-volumes in parentheses for the joint quantile estimate for $p = 0.90, 0.95, 0.99$ with AR(1) data using 2000 independent replications.

	NOLB		
n	$\beta = 0.01$	$\beta = 0.05$	$\beta = 0.10$
500	N/A	0.834 (0.242)	0.948 (0.330)
1,000	0.501 (0.114)	0.927 (0.195)	0.946 (0.238)
2,000	0.664 (0.095)	0.928 (0.141)	0.956 (0.175)
4,000	0.832 (0.081)	0.943 (0.104)	0.958 (0.126)
	OLB		
n	$\beta = 0.01$	$\beta = 0.05$	$\beta = 0.10$
500	N/A	0.812 (0.232)	0.940 (0.288)
1,000	0.470 (0.114)	0.925 (0.187)	0.949 (0.207)
2,000	0.672 (0.095)	0.932 (0.135)	0.947 (0.152)
4,000	0.851 (0.080)	0.946 (0.100)	0.965 (0.110)

REFERENCES

Alexopoulos, C., N. T. Argon, D. Goldsman, G. Tokol, and J. R. Wilson. 2007. "Overlapping Variance Estimators for Simulation". *Operations Research* 55(6):1090–1103.

Alexopoulos, C., J. H. Boone, D. Goldsman, A. Lolos, K. D. Dingeç, and J. R. Wilson. 2020. "Steady-State Quantile Estimation Using Standardized Time Series". In *Proceedings of the 2020 Winter Simulation Conference*, edited by K.-H. Bae, B. Feng, S. Kim, S. Lazarova-Molnar, Z. Zheng, T. Roeder, and R. Theising, 289–300. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.

Alexopoulos, C., D. Goldsman, A. C. Mokashi, K.-W. Tien, and J. R. Wilson. 2019. "Sequest: A Sequential Procedure for Estimating Quantiles in Steady-State Simulations". *Operations Research* 67(4):1162–1183.

Alexopoulos, C., D. Goldsman, A. C. Mokashi, and J. R. Wilson. 2017. "Automated Estimation of Extreme Steady-State Quantiles via the Maximum Transformation". *ACM Transactions on Modeling and Computer Simulation (TOMACS)* 27(4):1–29.

Alexopoulos, C., D. Goldsman, and J. R. Wilson. 2012. "A New Perspective on Batched Quantile Estimation". In *Proceedings of the 2012 Winter Simulation Conference*, edited by C. Laroque, J. Himmelspach, R. Pasupathy, O. Rose, and A. M. Uhrmacher, 1–11. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.

Bahadur, R. R. 1966. "A Note on Quantiles in Large Samples". *The Annals of Mathematical Statistics* 37(3):577–580.

Billingsley, P. 1999. *Convergence of Probability Measures*. 2nd ed. ed. Wiley series in Probability and Statistics. New York: Wiley.

Calvin, J. M., and M. K. Nakayama. 2013. "Confidence Intervals for Quantiles with Standardized Time Series". In *Proceedings of the 2013 Winter Simulations Conference*, edited by R. Pasupathy, S.-H. Kim, A. Tolk, R. Hill, and M. E. Kuhl., 601–612. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.

Chu, F., and M. K. Nakayama. 2012. "Confidence Intervals for Quantiles when Applying Variance-Reduction Techniques". *ACM Transactions on Modeling and Computer Simulation (TOMACS)* 22(2):1–25.

Dong, H., and M. K. Nakayama. 2020. "A Tutorial on Quantile Estimation via Monte Carlo". In *Monte Carlo and Quasi-Monte Carlo Methods*, Springer Proceedings in Mathematics & Statistics, 3–30. Cham: Springer International Publishing.

Ethier, S. N., and T. G. Kurtz. 2009. *Markov Processes: Characterization and Convergence*, Volume 282. John Wiley & Sons.

Glynn, P., and D. Iglehart. 1985. "Large-Sample Theory for Standardized Time Series: An Overview". In *Proceedings of the 1985 Winter Simulation Conference*, edited by D. Gantz, G. Blais, and S. Solomon, 129–134: ACM.

Glynn, P. W. 1996. "Importance Sampling for Monte Carlo Estimation of Quantiles". In *Mathematical Methods in Stochastic Simulation and Experimental Design: Proceedings of the 2nd St. Petersburg Workshop on Simulation*, 180–185. St. Petersberg, Russia: Publishing House of St. Petersberg University.

Glynn, P. W., and D. L. Iglehart. 1990. "Simulation Output Analysis Using Standardized Time Series". *Mathematics of Operations Research* 15(1):1–16.

Goldsman, D., M. Meketon, and L. Schruben. 1990. "Properties of Standardized Time Series Weighted Area Variance Estimators". *Management Science* 36(5):602–612.

He, S., and H. Lam. 2021. "Higher-Order Coverage Error Analysis for Batching and Sectioning". In *Proceedings of the 2021 Winter Simulation Conference*, edited by S. Kim, B. Feng, K. Smith, S. Masoud, Z. Zheng, C. Szabo, and M. Loper, 1–12. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.

Lei, L., C. Alexopoulos, Y. Peng, and J. R. Wilson. 2020. "Confidence Intervals and Regions for Quantiles using Conditional Monte Carlo and Generalized Likelihood Ratios". In *Proceedings of the 2020 Winter Simulation Conference*, edited by K.-H. Bae, B. Feng, S. Kim, S. Lazarova-Molnar, Z. Zheng, T. Roeder, and R. Theising, 2071–2082. Piscataway, New Jersey: IEEE: Institute of Electrical and Electronics Engineers, Inc.

Lei, L., C. Alexopoulos, Y. Peng, and J. R. Wilson. 2022. "Estimating Confidence Intervals and Regions for Quantiles by Monte Carlo Simulation". Preprint available at SSRN.

Muñoz, D. F. 2010. "On the Validity of the Batch Quantile Method for Markov Chains". *Operations Research Letters* 38(3):223–226.

Nakayama, M. K. 2014. "Confidence Intervals for Quantiles using Sectioning when Applying Variance-Reduction Techniques". *ACM Transactions on Modeling and Computer Simulation (TOMACS)* 24(4):1–21.

Schruben, L. W. 1983. "Confidence Interval Estimation Using Standardized Time Series". *Operations Research* 31(6):1090–1108.

Sen, P. K. 1972. "On the Bahadur Representation of Sample Quantiles for Sequences of φ-Mixing Random Variables". *Journal of Multivariate Analysis* 2(1):77–95.

Serfling, R. J. 2009. *Approximation Theorems of Mathematical Statistics*, Volume 162. John Wiley & Sons.

Wu, W. B. 2005. "On the Bahadur Representation of Sample Quantiles for Dependent Sequences". *The Annals of Statistics* 33(4):1934–1963.

AUTHOR BIOGRAPHIES

RAGHU PASUPATHY is Professor of Statistics at Purdue University. His current research interests lie broadly in general simulation methodology, statistical computation (especially optimization) and statistical inference. He has been actively involved with the Winter Simulation Conference for the past 15 years. Raghu Pasupathy's email address is pasupath@purdue.edu, and his web page https://web.ics.purdue.edu/~pasupath contains links to papers, software codes, and other material.

DASHI I. SINGHAM is a Research Associate Professor in the Operations Research Department at the Naval Postgraduate School. She obtained her Ph.D. in Industrial Engineering & Operations Research from the University of California at Berkeley and her B.S.E. in Operations Research and Financial Engineering from Princeton University. Her email address is dsingham@nps.edu, and her website is https://faculty.nps.edu/dsingham/.

YINGCHIEH YEH is an Associate Professor in Institute of Industrial Management at National Central University, Taiwan. He received his Ph.D. from the School of Industrial Engineering at Purdue University. His primary research interests include simulation output analysis, applied probability and statistics, and applied operations research. His email address is yeh@mgt.ncu.edu.tw.