

CONFIDENCE INTERVALS FOR QUANTILES WITH STANDARDIZED TIME SERIES

James M. Calvin
Marvin K. Nakayama

Computer Science Department
New Jersey Institute of Technology
Newark, NJ 07102, USA

ABSTRACT

Schruben (1983) developed standardized time series (STS) methods to construct confidence intervals (CIs) for the steady-state mean of a stationary process. STS techniques cancel out the variance constant in the asymptotic distribution of the centered and scaled estimator, thereby eliminating the need to consistently estimate the asymptotic variance to obtain a CI. This is desirable since estimating the asymptotic variance in steady-state simulations presents nontrivial challenges. Difficulties also arise in estimating the asymptotic variance of a quantile estimator. We show that STS methods can be used to build CIs for a quantile for the case of crude Monte Carlo (i.e., no variance reduction) with independent and identically distributed outputs. We present numerical results comparing CIs for quantiles using STS to other procedures.

1 INTRODUCTION

Schruben (1983) introduced the class of *standardized time series* (STS) methods as a way of constructing asymptotically valid confidence intervals (CIs) for the steady-state mean of a stationary stochastic process. STS methods build a CI by *cancelling out* the time-average variance constant (TAVC) appearing in the central limit theorem (CLT) for the time average of the process, in a manner analogous to the Student t statistic. Thus, STS estimators are called *cancellation methods*. This is in contrast to *consistent-estimation methods*, such as the regenerative method (e.g., Chapter IV.4 of Asmussen and Glynn 2007), which consistently estimate the TAVC to construct a CI; see Glynn and Iglehart (1990) for a comparison of these two approaches. But constructing a consistent estimator of the TAVC is nontrivial (e.g., see Chapter IV of Asmussen and Glynn 2007), illustrating a benefit of cancellation methods. STS methods have been proposed and studied in many papers on steady-state simulation, including Goldsman and Schruben (1990), Calvin and Nakayama (2006), and Alexopoulos et al. (2007).

There are other parameters with simulation estimators whose asymptotic variance may be difficult to estimate. In this paper we show that STS methods can be used to construct CIs for one such parameter: a quantile. For a continuous cumulative distribution function (CDF) F and fixed $0 < p < 1$, the p -quantile of F is the smallest constant $Q(p)$ such that $F(Q(p)) = p$. For example, the median is the 0.5-quantile. Quantiles frequently arise in practice to assess risk. In finance, a quantile, which is known as a value-at-risk, is often used as a measure of portfolio risk; e.g., see Jorion (2007). Also, the U.S. Nuclear Regulatory Commission (NRC) requires nuclear-plant licensees to demonstrate that their facilities are complying with federal regulations using a so-called *95/95 criterion*. This entails establishing with 95% confidence that the 0.95-quantile of an output variable of a simulation lies below a mandated threshold (U.S. Nuclear Regulatory Commission 1989). Thus, not only do we need a point estimate for a quantile but also a CI.

The current paper focuses on estimating a quantile using crude Monte Carlo (CMC; i.e., no variance reduction) with independent and identically distributed (i.i.d.) outputs having CDF F . Even in this simple context, there are issues in trying to construct a CI for a quantile. Specifically, one way of developing a CI

for $Q(p)$ first shows that the p -quantile estimator satisfies a CLT (Section 2.3.3 of Serfling 1980). Given a consistent estimator of the asymptotic variance τ_p^2 appearing in the CLT, we can use it when unfolding the CLT to obtain an asymptotically valid (as the sample size $n \rightarrow \infty$) CI for $Q(p)$. The asymptotic variance $\tau_p^2 = p(1-p)/f^2(Q(p))$, where f is the density of F . Methods for consistently estimating $f(Q(p))$ include a finite difference (Bloch and Gastwirth 1968, Bofinger 1975) and kernel estimators (Wand and Jones 1995), but these techniques require the user to specify an appropriate value for a parameter known as a bandwidth, which can be tricky to do in practice. An STS method avoids this complication by instead cancelling out τ_p in the relevant limit theorem, so one does not need to consistently estimate τ_p . We can think of batching and sectioning (Section III.5a of Asmussen and Glynn 2007) as other examples of cancellation methods, and STS generalizes these approaches.

We show how to use STS methods, including the STS weighted area estimator originally developed by Goldsman and Schruben (1990) for constructing a CI for a steady-state mean, to build a CI for a quantile. We present numerical results from running experiments comparing the methods discussed in the previous paragraph with our proposed approach.

Other ways for constructing CIs for a quantile in the CMC i.i.d. setting include a method based on the binomial distribution (e.g., see Section 9.2 of Hogg and Craig 1978) and the bootstrap (Efron 1979). An attractive feature of the binomial approach is that it produces *exact* CIs for *finite* sample sizes n , as opposed to requiring $n \rightarrow \infty$, as in the other (asymptotic) procedures described. However, the technique does not generally appear to apply beyond the CMC i.i.d. framework, and one of our eventual goals is to handle such cases. (An exception is that Hsu and Nelson (1990) extend the binomial argument to a multinomial to work for control variates.) A drawback of the bootstrap estimator of the quantile estimator's variance τ_p^2 is that it converges at a slower rate than finite-difference and kernel estimators (?).

The rest of the paper has the following organization. Section 2 develops the mathematical framework. We describe the STS weighted area estimator in Section 3, and then provide extensions to other STS methods and simultaneous confidence intervals in Section 4. Numerical results are given in Section 5, and we present concluding remarks in Section 6.

2 MATHEMATICAL FRAMEWORK

Consider a random variable X with unknown CDF F and density f . For $0 < p < 1$, we define the p -quantile of F (or equivalently of X) to be $Q(p) = F^{-1}(p) \equiv \inf\{x : F(x) \geq p\}$. Our initial goal is to estimate and construct a confidence interval for $Q(p)$ for a fixed p using simulation. In Section 4 we consider the case of simultaneous confidence intervals for more than one p .

Suppose that \hat{F}_n is an estimator of the CDF F based on a computation budget of n , which often represents the sample size. Let $\hat{Q}_n(p) = \hat{F}_n^{-1}(p)$ be the p -quantile estimator based on the budget of n . We will need to make an assumption on the entire sequence of p -quantile estimators to establish the asymptotic validity of the CI procedures in this paper.

To do this, let $C[0, 1]$ be the space of continuous real-valued functions on the unit interval, and let $D[0, 1]$ be the space of right-continuous real-valued functions with left limits on the unit interval; see Chapters 2 and 3 of Billingsley (1999) for details on these spaces. For fixed $0 < p < 1$, define the centered and scaled p -quantile estimation process $Z_n = [Z_n(t) : 0 \leq t \leq 1] \in D[0, 1]$ with

$$Z_n(t) = \frac{\lfloor nt \rfloor}{\sqrt{n}} (\hat{Q}_{\lfloor nt \rfloor}(p) - Q(p)),$$

where $\lfloor \cdot \rfloor$ is the floor function. Let $W = [W(t) : 0 \leq t \leq 1]$ be a standard Brownian motion; i.e., W is a Gaussian process with $E[W(t)] = 0$ for all $t \in [0, 1]$ and $E[W(t_1)W(t_2)] = t_1 \wedge t_2$, where $a \wedge b = \min(a, b)$; see Section 9 of Billingsley (1999). Let \Rightarrow denote weak convergence; see Chapter 1 of Billingsley (1999). We assume the following functional central limit theorem (FCLT):

Assumption FCLT. $Z_n \Rightarrow \tau_p W$ as $n \rightarrow \infty$ in $D[0, 1]$ endowed with the Skorohod metric, where $0 < \tau_p < \infty$ is some constant.

We focus on the case of crude Monte Carlo with i.i.d. outputs, and we now explain that such an FCLT holds under certain conditions. Let X_1, X_2, \dots, X_n be i.i.d. samples having CDF F . We form the *empirical CDF* \hat{F}_k as an estimator of F from the first $1 \leq k \leq n$ samples with $\hat{F}_k(x) = \frac{1}{k} \sum_{i=1}^k I(X_i \leq x)$, where $I(\cdot)$ denotes the indicator function, which takes value 1 (resp., 0) when its argument is true (resp., false). Note that \hat{F}_k assigns mass $1/k$ to each of the first k samples. Let $[0, 1]^2 = [0, 1] \times [0, 1]$, and define $D[0, 1]^2$ to be the space of right-continuous two-parameter real-valued functions with left limits on $[0, 1]^2$. Define the two-parameter process $\rho_n = [\rho_n(y, t) : 0 \leq y, t \leq 1] \in D[0, 1]^2$ with

$$\rho_n(y, t) = f(Q(y)) \frac{\lfloor nt \rfloor}{\sqrt{n}} (Q(y) - \hat{Q}_{\lfloor nt \rfloor}(y)).$$

Also, define $\hat{\rho}_n = [\hat{\rho}_n(y, t) : 0 \leq y, t \leq 1] \in D[0, 1]^2$ as a truncated version of ρ_n with

$$\hat{\rho}_n(y, t) = \rho_n(y, t) \beta_n(y), \tag{1}$$

where $\beta_n(y) = I(1/(n+1) \leq y \leq n/(n+1))$. Assume that F is twice differentiable on (a, b) , where $a = \sup\{x : F(x) = 0\}$ and $b = \inf\{x : F(x) = 1\}$. Also assume that $f(x) > 0$ for $x \in (a, b)$, and that $\sup_{0 < y < 1} y(1-y)|f'(Q(y))|/f^2(Q(y)) \leq \gamma$ for some $0 < \gamma < \infty$, where f' is the derivative of f . Let $K = [K(y, t) : 0 \leq y, t \leq 1] \in D[0, 1]^2$ be a Kiefer process, which is a real-valued two-parameter Gaussian process with $K(y, 0) = 0, E[K(y, t)] = 0$ for all $(y, t) \in [0, 1]^2$ and $E[K(y_1, t_1)K(y_2, t_2)] = (t_1 \wedge t_2)(y_1 \wedge y_2 - y_1 y_2)$ for all $(y_i, t_i) \in [0, 1]^2, i = 1, 2$. Also, K lies in $C[0, 1]^2$ with probability 1 (w.p. 1), where $C[0, 1]^2$ denotes the space of continuous real-valued functions on $[0, 1]^2$; see Section 1.15 of Csörgő and Révész (1981) for more details on the Kiefer process. Corollary 3.1 of Csörgő and Szyszkowicz (1998) establishes

$$\hat{\rho}_n \Rightarrow K \tag{2}$$

as $n \rightarrow \infty$ in $D[0, 1]^2$. (Under stronger conditions on F , Corollary 3.2 of Csörgő and Szyszkowicz 1998 proves that $\rho_n \Rightarrow K$ as $n \rightarrow \infty$ in $D[0, 1]^2$.) For fixed $0 < y < 1$, define the projection mapping $\pi_y : D[0, 1]^2 \rightarrow D[0, 1]$ such that for $r = [r(y, t) : 0 \leq y, t \leq 1] \in D[0, 1]^2$, we have $\pi_y(r) = [r(y, t) : 0 \leq t \leq 1] \in D[0, 1]$, a one-parameter function of t . The projection mapping π_y is continuous at $r \in C[0, 1]^2$. Since $K \in C[0, 1]^2$ w.p. 1, the continuous-mapping theorem (Theorem 2.7 of Billingsley 1999) and (2) together imply that $\pi_y(\hat{\rho}_n) \Rightarrow \pi_y(K)$ as $n \rightarrow \infty$. Now fix $y = p$, and for n sufficiently large so that $\beta_n(y) = 1$, we have

$$f(Q(p)) \frac{\lfloor nt \rfloor}{\sqrt{n}} (Q(p) - \hat{Q}_{\lfloor nt \rfloor}(p)) = \pi_p(\hat{\rho}_n) \Rightarrow \pi_p(K) \tag{3}$$

as $n \rightarrow \infty$ in $D[0, 1]$. But $\pi_p(K)$ is a one-parameter Gaussian process lying in $C[0, 1]$ w.p. 1 with $E[\pi_p(K)(t)] = 0$ and $E[\pi_p(K)(t_1)\pi_p(K)(t_2)] = (t_1 \wedge t_2)p(1-p)$, so $\pi_p(K)$ is equal in distribution to $\sqrt{p(1-p)}W$, where W is a standard Brownian motion. Hence, dividing (3) by $-f(Q(p))$ shows Assumption FCLT holds with $\tau_p = \sqrt{p(1-p)}/f(Q(p))$ since $-W$ is also a standard Brownian motion.

3 STANDARDIZED TIME SERIES WEIGHTED AREA ESTIMATOR

We now describe how to use the STS weighted area estimator to construct asymptotically valid CIs for $Q(p)$. Define the *standardized time series* $T_n = [T_n(t) : 0 \leq t \leq 1] \in D[0, 1]$ with

$$T_n(t) = \frac{\lfloor nt \rfloor}{\sqrt{n}} (\hat{Q}_n(p) - \hat{Q}_{\lfloor nt \rfloor}(p)),$$

which can be computed from the sequence of CDF estimators $\hat{F}_k, k = 1, 2, \dots, n$. It can be shown that Assumption FCLT implies

$$(\sqrt{n}(\hat{Q}_n(p) - Q(p)), T_n) \Rightarrow (\tau_p W(1), \tau_p B) \tag{4}$$

as $n \rightarrow \infty$, where $B = [B(t) : 0 \leq t \leq 1]$ is a standard Brownian bridge that is independent of $W(1)$, and $W(1)$ has a standard normal distribution; see Schruben (1983) or Glynn and Iglehart (1990).

Now define the *weighted area estimator* $A_n(w)$ as

$$A_n(w) = \left[\frac{1}{n} \sum_{k=1}^n w \left(\frac{k}{n} \right) T_n \left(\frac{k}{n} \right) \right]^2,$$

where the weighting function $w : [0, 1] \rightarrow \mathfrak{R}$ is twice continuously differentiable on $[0, 1]$ and satisfies $E[(\int_0^1 w(t)B(t) dt)^2] = 1$. For estimating the TAVC from the CLT for the time average of a stationary process, some common choices for the weighting function are

$$w_0(t) = \sqrt{12} \quad \text{for all } t \in [0, 1] \tag{5}$$

and

$$w_1(t) = \sqrt{840}(3t^2 - 3t + 1/2); \tag{6}$$

see Schruben (1983), Goldsman, Meketon, and Schruben (1990), and Goldsman and Schruben (1990). Weighting function w_1 leads to a weighted area estimator of the TAVC having a bias expansion in which the highest-order bias term vanishes. However, when estimating the variance constant τ_p^2 in the quantile-estimator CLT, we do not have such a guarantee for w_1 .

Let χ_d^2 be a chi-squared random variable and R_d denote a Student t random variable, both with d degrees of freedom. Following the arguments in Alexopoulos et al. (2007) covering the case of estimating a steady-state mean and the TAVC, we can show the convergence in (4) ensures that as $n \rightarrow \infty$,

$$A_n(w) \Rightarrow \tau_p^2 \chi_1^2 \quad \text{and} \quad \frac{\sqrt{n}(\hat{Q}_n(p) - Q(p))}{(A_n(w))^{1/2}} \Rightarrow \frac{\tau_p W(1)}{(\tau_p^2 \chi_1^2)^{1/2}} = \frac{W(1)}{(\chi_1^2)^{1/2}} \stackrel{\mathcal{D}}{=} R_1,$$

the latter illustrating the cancellation of τ_p in the limit, where $\stackrel{\mathcal{D}}{=}$ denotes equality in distribution. This allows us to construct a (two-sided) asymptotically valid (as $n \rightarrow \infty$) $100(1 - \alpha)\%$ CI for $Q(p)$ as

$$I_n = \left[\hat{Q}_n(p) \pm t_{1,\alpha/2}(A_n(w)/n)^{1/2} \right], \tag{7}$$

where the constant $t_{d,\delta}$ is defined such that $P(R_d \geq t_{d,\delta}) = \delta$.

We can also apply nonoverlapping batching to the weighted area estimator to obtain another CI for $Q(p)$ as follows. Fix a number $b \geq 2$ of batches, and divide the n total outputs into b nonoverlapping batches, each of size $m = n/b$. For each batch j , define the CDF estimator $\hat{F}_{j,k}$ based on the first $1 \leq k \leq m$ samples in batch j as $\hat{F}_{j,k}(x) = (1/k) \sum_{i=(j-1)m+1}^{(j-1)m+k} I(X_i \leq x)$, and let $\hat{Q}_{j,k}(p) = \hat{F}_{j,k}^{-1}(p)$ be the corresponding p -quantile estimator. For batch j , define the standardized time series

$$T_{j,m}(t) = \frac{\lfloor mt \rfloor}{\sqrt{m}} (\hat{Q}_{j,m}(p) - \hat{Q}_{j,\lfloor mt \rfloor}(p))$$

for $t \in [0, 1]$. From each batch j , we also compute a weighted area estimator

$$A_{j,m}(w) = \left[\frac{1}{m} \sum_{k=1}^m w \left(\frac{k}{m} \right) T_{j,m} \left(\frac{k}{m} \right) \right]^2$$

for weighting function w , and we then get the nonoverlapping batched weighted area estimator as

$$\bar{A}_{b,m}(w) = \frac{1}{b} \sum_{j=1}^b A_{j,m}(w).$$

It can then be proven (again by applying arguments in Alexopoulos et al. (2007)) that

$$\bar{A}_{b,m}(w) \Rightarrow \frac{\tau_p^2 \chi_b^2}{b} \quad \text{and} \quad \frac{\sqrt{n}(\hat{Q}_n(p) - Q(p))}{(\bar{A}_{b,m}(w))^{1/2}} \Rightarrow R_b$$

as $n = bm \rightarrow \infty$ with $b \geq 2$ fixed. This allows us to construct an asymptotically valid (as $m \rightarrow \infty$ with $b \geq 2$ fixed) $100(1 - \alpha)\%$ CI for $Q(p)$ as

$$I_{n,b} = \left[\hat{Q}_n(p) \pm t_{b,\alpha/2} (\bar{A}_{b,m}(w)/n)^{1/2} \right]. \tag{8}$$

Another variation uses overlapping batches (Meketon and Schmeiser 1984). We again denote the batch size as $m > 0$, and we define $b = n/m$, which no longer has the interpretation as the number of batches. There are $n - m + 1$ overlapping batches, where for $j = 1, 2, \dots, n - m + 1$, the j th overlapping batch consists of the m samples indexed by $j, j + 1, \dots, j + m - 1$. We define the standardized time series from the j th overlapping batch as

$$\tilde{T}_{j,m}(t) = \frac{\lfloor mt \rfloor}{\sqrt{m}} (\tilde{Q}_{j,m}(p) - \tilde{Q}_{j,\lfloor mt \rfloor}(p))$$

for $t \in [0, 1]$, where $\tilde{Q}_{j,k}(p) = \tilde{F}_{j,k}^{-1}(p)$ and $\tilde{F}_{j,k}(x) = (1/k) \sum_{i=j}^{j+k-1} I(X_i \leq x)$. From the j th overlapping batch, we compute a weighted area estimator

$$\tilde{A}_{j,m}(w) = \left[\frac{1}{m} \sum_{k=1}^m w \left(\frac{k}{m} \right) \tilde{T}_{j,m} \left(\frac{k}{m} \right) \right]^2$$

for weighting function w , and we then get the overlapping batched weighted area estimator as

$$\tilde{A}_{b,m}^O(w) = \frac{1}{n - m + 1} \sum_{j=1}^{n-m+1} \tilde{A}_{j,m}(w).$$

To describe the limiting distribution of $\tilde{A}_{b,m}^O(w)$ as $m \rightarrow \infty$ with $b > 1$ fixed, define $W_b = [W_b(t) : 0 \leq t \leq b]$ as a standard Brownian motion on $[0, b]$, and define $[B_s : 0 \leq s \leq b - 1]$ as a collection of processes such that for each $s \in [0, b - 1]$, the process $B_s = [B_s(t) : 0 \leq t \leq 1]$ has $B_s(t) = [W_b(s+t) - W_b(s)] - t[W_b(s+1) - W_b(s)]$ for $t \in [0, 1]$. Each B_s is a standard Brownian bridge. As in the case when estimating a steady-state mean (Alexopoulos et al. 2007), it can then be proven that for fixed $b > 1$,

$$\tilde{A}_{b,m}^O(w) \Rightarrow \tilde{A}_b^O(w) \equiv \frac{1}{b-1} \int_0^{b-1} \left[\tau_p \int_0^1 w(u) B_s(u) du \right]^2 ds \tag{9}$$

as $m \rightarrow \infty$. The distribution of the limit $\tilde{A}_b^O(w)$ in (9) is complicated to work out, but Alexopoulos et al. (2007) suggest to approximate it as $\tau_p^2 \chi_v^2 / v$ with

$$v = \left\langle \frac{2E^2[\tilde{A}_b^O(w)]}{\text{Var}[\tilde{A}_b^O(w)]} \right\rangle, \tag{10}$$

where $\langle \cdot \rangle$ denotes the rounding function. It turns out that $E[\tilde{A}_b^O(w)] = 1$ for all weighting functions w , but $\text{Var}[\tilde{A}_b^O(w)]$ depends on w . Alexopoulos et al. (2007) show that $\text{Var}[\tilde{A}_b^O(w_0)] = (24b - 31)/[35(b - 1)^2]$ and $\text{Var}[\tilde{A}_b^O(w_1)] = (3514b - 4359)/[4290(b - 1)^2]$. This allows us to construct an approximate (as $m \rightarrow \infty$ with $b > 1$ fixed) $100(1 - \alpha)\%$ CI for $Q(p)$ as

$$I_{n,b}^O = \left[\hat{Q}_n(p) \pm t_{v,\alpha/2} (\tilde{A}_{b,m}^O(w)/n)^{1/2} \right]. \tag{11}$$

4 OTHER STS METHODS AND SIMULTANEOUS CIs

We now extend the results of the previous section to general STS methods, and also develop simultaneous confidence intervals for more than one quantile. Recall $\hat{\rho}_n(y, t)$ in (1). Define the bridge map $\Gamma : D[0, 1]^2 \rightarrow D[0, 1]^2$ by $\Gamma(g)(y, t) = g(y, t) - tg(y, 1)$ for $(y, t) \in [0, 1]^2$. This map is continuous on $D[0, 1]^2$, and

$$\Gamma(\hat{\rho}_n)(y, t) = f(Q(y)) \frac{\lfloor nt \rfloor}{\sqrt{n}} (\hat{Q}_n(y) - \hat{Q}_{\lfloor nt \rfloor}(y)) \beta_n(y) + \frac{\lfloor nt \rfloor - nt}{n} \hat{\rho}_n(y, 1) \equiv g_n(y, t) + r_n(y, t),$$

where $r_n \Rightarrow \mathbf{0}$, the constant zero function, as $n \rightarrow \infty$. The continuity of Γ and (2) ensure $\Gamma(\hat{\rho}_n) \Rightarrow \Gamma(K) \equiv \bar{K}$ as $n \rightarrow \infty$, and together with $r_n \Rightarrow \mathbf{0}$, this implies that $g_n \Rightarrow \bar{K}$, where \bar{K} is a bivariate Brownian bridge; that is, a centered Gaussian field with covariance $E[\bar{K}(s_1, t_1)\bar{K}(s_2, t_2)] = (s_1 \wedge s_2 - s_1 s_2)(t_1 \wedge t_2 - t_1 t_2)$. Furthermore, \bar{K} is independent of $K(\cdot, 1)$.

Let ζ be an STS scaling function (Glynn and Iglehart 1990); that is, $\zeta : D[0, 1] \rightarrow [0, \infty)$ is continuous on $C[0, 1]$, $\zeta(\alpha x) = \alpha \zeta(x)$ for $\alpha \in (0, \infty)$, and $P(\zeta(\bar{K}(y, \cdot)) > 0) = 1$ for any $0 < y < 1$. (The various weighted area estimators in Section 3 are discrete-time versions of scaling functions.) For $n \geq \max\{(1-y)/y, y/(1-y)\}$ (so that $\beta_n(y) = 1$), the homogeneity of ζ implies

$$V_n(y) \equiv \frac{\hat{\rho}_n(y, 1)}{\zeta(g_n(y, \cdot))} = \frac{Q(y) - \hat{Q}_n(y)}{\zeta\left(\frac{\lfloor nt \rfloor}{n} (\hat{Q}_n(y) - \hat{Q}_{\lfloor nt \rfloor}(y)) : 0 \leq t \leq 1\right)} \equiv \frac{Q(y) - \hat{Q}_n(y)}{M_n}.$$

Since K and \bar{K} have continuous sample paths w.p. 1, $V_n(y) \Rightarrow K(y, 1)/\zeta(\bar{K}(y, \cdot))$ as $n \rightarrow \infty$. We then can construct an asymptotic $100(1 - \alpha)\%$ CI for $Q(y)$ as

$$[\hat{Q}_n(y) \pm \gamma M_n], \tag{12}$$

where γ is the $(1 - \alpha/2)$ -critical point of the distribution H of $K(y, 1)/\zeta(\bar{K}(y, \cdot))$; i.e., $H(\gamma) = 1 - \alpha/2$.

Now consider the ‘‘range’’ scaling function

$$\zeta(x) = \max_{0 \leq s \leq 1} x(s) - \min_{0 \leq s \leq 1} x(s). \tag{13}$$

The limiting random variable $K(y, 1)/\zeta(\bar{K}(y, \cdot))$ then has the distribution of the ratio of a standard normal over a standard Brownian excursion height, so $H(x) = 1 - \pi x^2 \sum_{n=1}^{\infty} n K_1(\pi n x)$, where K_1 is the modified Bessel function (Calvin 2004). The 0.95-critical point of H is $\gamma \approx 1.39739$ (i.e., $H(\gamma) = 0.95$), which we can use in (12) to obtain an asymptotic 90% confidence interval for $Q(y)$, where, for $n \geq \max\{(1-y)/y, y/(1-y)\}$,

$$\begin{aligned} M_n &= \max_{0 \leq t \leq 1} \frac{\lfloor nt \rfloor}{n} (\hat{Q}_n(y) - \hat{Q}_{\lfloor nt \rfloor}(y)) - \min_{0 \leq t \leq 1} \frac{\lfloor nt \rfloor}{n} (\hat{Q}_n(y) - \hat{Q}_{\lfloor nt \rfloor}(y)) \\ &= \max_{1 \leq k \leq n} \frac{k}{n} (\hat{Q}_n(y) - \hat{Q}_k(y)) - \min_{1 \leq k \leq n} \frac{k}{n} (\hat{Q}_n(y) - \hat{Q}_k(y)). \end{aligned}$$

We can extend the approach for any STS scaling function ζ to construct simultaneous CIs for more than one value of y . (Chapter 4 of Csörgő (1983) considers other simultaneous CIs.) For $Y \subset (0, 1)$,

$$\sup_{y \in Y} V_n(y) \Rightarrow \sup_{y \in Y} \frac{K(y, 1)}{\zeta(\bar{K}(y, \cdot))} \tag{14}$$

as $n \rightarrow \infty$, and this can be used to construct simultaneous $100(1 - \alpha)\%$ joint CIs for $Q(y)$, $y \in Y$, as

$$\left[\hat{Q}_n(y) \pm v_\alpha \zeta \left(\frac{\lfloor nt \rfloor}{n} (\hat{Q}_n(y) - \hat{Q}_{\lfloor nt \rfloor}(y)) \beta_n(y) : 0 \leq t \leq 1 \right) : y \in Y \right], \tag{15}$$

where v_α is the $(1 - \alpha/2)$ -critical point of the distribution of $\sup_{y \in Y} K(y, 1)/\zeta(\bar{K}(y, \cdot))$. It seems that in most cases it would be necessary to estimate the critical value v_α using simulation.

5 NUMERICAL RESULTS COMPARING STS AND OTHER ESTIMATORS

We ran simulation experiments using a variety of methods to construct confidence intervals for quantiles for a stochastic activity network (SAN). Also known as stochastic PERTs, SANs are widely used in project planning, where a project comprises a collection of activities having precedence relations, and the times to complete the activities follow specified probability distributions; e.g., see Adlakha and Kulkarni (1989). We consider a SAN, previously studied by Hsu and Nelson (1990), consisting of $d = 5$ activities, which correspond to edges labeled $1, 2, \dots, d$ in the network. The length A_i of each edge i is exponential with mean 1, and A_1, A_2, \dots, A_d are independent. There are $q = 3$ paths through the network, with $B_1 = \{1, 2\}$, $B_2 = \{1, 3, 5\}$ and $B_3 = \{4, 5\}$ as the sets of edges on the q paths. Let $X = \max_{1 \leq j \leq q} \sum_{i \in B_j} A_i$ be the (random) length of the longest path, which corresponds to the time to complete the project, and we want to estimate and construct CIs for the p -quantile $Q(p)$ of X . The CDF of X is $F(x) = 1 + (3 - 3x - x^2/2)e^{-x} + (-3 - 3x + x^2/2)e^{-2x} - e^{-3x}$ for $x \geq 0$, and $F(x) = 0$ for $x < 0$.

The goal of our experiments is to study the coverage level and average half-widths of nominal 90% CIs for the p -quantile $Q(p)$ for different values of p . The CIs we developed are asymptotically valid as the sample size $n \rightarrow \infty$, and we study the finite-sample behavior of our intervals for $n = 4^r \times 100$, $r = 0, 1, 2, 3$.

Table 1 contains the results from our experiments using different methods to construct CIs, where the coverage (and average half-widths) are computed from running 10^4 independent replications. The first column gives the sample size n . The next column presents results when using a central finite difference (FD) to estimate $q(p) \equiv 1/f(Q(p)) = \frac{d}{dp}F^{-1}(p)$, which appears in the asymptotic variance $\tau_p^2 = p(1-p)q^2(p)$ of the central limit theorem for $\hat{Q}_n(p)$. Bloch and Gastwirth (1968) and Bofinger (1975) prove that the FD $\hat{q}_n(p) = [\hat{F}_n^{-1}(p+h_n) - \hat{F}_n^{-1}(p-h_n)]/(2h_n) \Rightarrow q(p)$ as $n \rightarrow \infty$ when the bandwidth $h_n \rightarrow 0$ as $n \rightarrow \infty$ with $nh_n \rightarrow \infty$ and f is continuous in a neighborhood of $Q(p)$. The resulting asymptotically valid 90% CI for $Q(p)$ using FD is $[\hat{Q}_n(p) \pm 1.645 \sqrt{p(1-p)\hat{q}_n(p)/\sqrt{n}}]$. In our experiments we chose $h_n = 0.5n^{-1/2}$.

The third column of Table 1 contains results when $f(Q(p))$ is estimated using a plug-in kernel estimator. Specifically, let k be a kernel function, which is typically chosen as a density function symmetric about the origin, and let k_h be a scaled kernel function with $k_h(y) = \frac{1}{h}k(\frac{y}{h})$, where $h > 0$ is the bandwidth. Then the kernel estimator \hat{f}_n of the density f is defined as $\hat{f}_n(y) = \int k_h(y-x) d\hat{F}_n(x) = \frac{1}{n} \sum_{i=1}^n k_{h_n}(y-X_i)$, where h_n is the bandwidth satisfying $h_n \rightarrow 0$ and $nh_n \rightarrow \infty$ as $n \rightarrow \infty$; e.g., see Wand and Jones (1995) for an overview of kernel estimators. Then the plug-in kernel estimator of $f(Q(p))$ is $\hat{f}_n(\hat{Q}_n(p))$. The resulting asymptotically valid 90% CI for $Q(p)$ is $[\hat{Q}_n(p) \pm 1.645 \sqrt{p(1-p)/(\hat{f}_n(\hat{Q}_n(p))\sqrt{n})}]$. In our experiments we chose the Gaussian kernel (i.e., k is the density function of a standard normal) and bandwidth $h_n = 0.5n^{-1/5}$.

The fourth column (labeled ‘‘Batch.’’) of Table 1 gives results when using batching. This method divides the n outputs into $b \geq 2$ batches, each of size $m = n/b$. From the j th batch, a p -quantile estimate $\hat{Q}_{j,m}(p)$ is computed, and we calculate the sample average and sample variance of the b batch quantile estimates $\hat{Q}_{j,m}(p)$, $j = 1, 2, \dots, b$, to construct a CI for $Q(p)$. Specifically, let $\bar{Q}_{b,m}(p) = (1/b) \sum_{j=1}^b \hat{Q}_{j,m}(p)$ be the sample average of the b batch quantile estimates, and let $S^2 = (1/(b-1)) \sum_{j=1}^b (\hat{Q}_{j,m}(p) - \bar{Q}_{b,m}(p))^2$. The resulting asymptotically valid 90% CI for $Q(p)$ is then $[\bar{Q}_{b,m}(p) \pm t_{b-1,0.05}S/\sqrt{b}]$. In our experiments we used $b = 10$ batches.

The fifth column (labeled ‘‘Sect.’’) of Table 1 presents results from applying sectioning; see Section III.5a of Asmussen and Glynn (2007) and Nakayama (2012). Similar to batching, sectioning replaces the batching point estimate $\bar{Q}_{b,m}(p)$ with the overall quantile estimate $\hat{Q}_n(p)$ throughout the CI. Specifically, let $S'^2 = (1/(b-1)) \sum_{j=1}^b (\hat{Q}_{j,m}(p) - \hat{Q}_n(p))^2$, and the resulting asymptotically valid 90% CI for $Q(p)$ is then $[\hat{Q}_n(p) \pm t_{b-1,0.05}S'/\sqrt{b}]$. An advantage of sectioning over batching is that the former’s CI is centered at a less-biased estimator than the latter, and this can lead to better coverage when n is small; see Table 1.

Columns 6 and 7 of Table 1 contain the results for the CIs in (7) and (8) from applying the STS weighted-area (WA) estimator with (constant) weighting function w_0 from (5) and b batches, where the case $b = 1$

corresponds to (7). The eighth column of the table gives results for the STS range estimator (13). The last column of Table 1 presents results where we used the exact value of $q(p)$, which was computed numerically; in this case the resulting asymptotically valid 90% CI for $Q(p)$ is then $\left[\hat{Q}_n(p) \pm 1.645\sqrt{p(1-p)q(p)/\sqrt{n}}\right]$. This is provided as a benchmark, which is typically not implementable since $q(p)$ is unknown.

In general, in terms of coverage, sectioning outperforms batching, especially for small n . FD and Kernel do worse than sectioning when $p \approx 1$ and n is small.

Now we compare sectioning with the STS weighted-area and range methods. Let $WA(w, b)$ denote the STS weighted area estimator with weighting function w and b batches. The $WA(w_0, 1)$ estimator has significantly larger average half-widths than the $WA(w_0, 10)$ estimator. This is because the critical point in the CI for the $WA(w_0, b)$ method is that of a Student t random variable R_b with b degrees of freedom, and the R_b distribution has much heavier tails for $b = 1$ than for $b = 10$. For $p = 0.5$ and 0.8 , sectioning and the various STS methods have about the same coverage for all n . For $p = 0.95$ and 0.99 , $WA(w_0, 1)$ has better coverage than $WA(w_0, 10)$ for small n . This is likely caused by the fact that the asymptotic validity of the STS batching methods relies on the batch size $m = n/b$ to be large for the Brownian approximation to hold; but m is small when n is small, leading to the asymptotics not yet holding. The STS range estimator has about the same average half width as $WA(w_0, 10)$ for large n , but better coverage for small n . For $p = 0.99$, $WA(w_0, 1)$ has better coverage than sectioning for small n .

Table 2 gives results for different weighting functions, w_0 from (5) and w_1 from (6), for nonoverlapping batching (8) and overlapping batching (11) from running 2000 independent replications. In our experiments we chose $b = 10$, so for overlapping batching, this leads to v in (10) equaling 27 for w_0 from (5) and $v = 23$ for w_1 from (6). As noted in Section 3, the weighting function w_1 in (6) results in an asymptotically lower-biased weighted area estimate of the TAVC when estimating a steady-state mean, but there is no such guarantee when estimating a quantile. For $p = 0.5$ and 0.8 , the overlapping estimators have about the same coverage as the other methods and about the same average half-widths as sectioning and nonoverlapping batching with $b = 10$. For $p = 0.95$ the overlapping estimators have about the same average half-widths as sectioning and the nonoverlapping estimators with $b = 10$, and the overlapping estimators have higher coverage than nonoverlapping with $b = 10$ for small n . But for $p = 0.99$ overlapping performs about the same as nonoverlapping with $b = 10$ and worse than $b = 1$. Overall, it appears that when p is extreme, the range estimator does the best job of balancing coverage and average half-widths of the methods considered.

We also ran experiments to construct the simultaneous CIs in (15) for the y -quantiles $Q(y)$ for all $y \in Y$ using the range estimator (13). We chose $Y = \{0.90, 0.95, 0.99\}$, for which the simulated 0.95-quantile of the right side of (14) is 1.6211. Executing 10^4 independent replications led to coverages of 0.492, 0.760, 0.819, and 0.833 for run lengths of 100, 400, 1600, and 6400, respectively. The half-widths for a single quantile (Table 1) were typically about 85% as large as for the simultaneous intervals.

6 CONCLUDING REMARKS

Previous work applying STS methods mainly focuses on constructing CIs for the steady-state mean of a stationary stochastic process. But here we demonstrated how STS estimators can be used to build CIs for a quantile when applying crude Monte Carlo with i.i.d. sampling. This led to not only a CI for a single quantile, but also simultaneous CIs for multiple quantiles. More generally, STS methods can be employed to develop CIs for any parameter θ for which there is a corresponding parameter estimator $\hat{\theta}_n$ satisfying an FCLT. STS may be an attractive approach when it is difficult to consistently estimate the asymptotic variance of $\hat{\theta}_n$. The FCLT required for STS need not necessarily have a Brownian limit nor the canonical scaling \sqrt{n} . Glynn and Whitt (1991) describe many simulation settings for which the estimator satisfies a FCLT with a Brownian limit but with some having non-canonical scaling n^c for $c \neq 1/2$. Of course, the applicability of STS for a non-Brownian limit relies on the analytical tractability of the limiting process.

Our numerical results indicate that the STS weighted area estimator performs about the same as sectioning in terms of coverage of the resulting CIs when p is not too extreme. For $p = 0.99$, the coverage

for the CI of the weighted area estimator with a single batch was closer to the nominal level than for sectioning for small sample sizes, but at the cost of wider average half-widths. This demonstrates that STS methods can provide CIs with improved coverage. Overall, when p is extreme, the range estimator provided the best balance of coverage and average half-widths of the methods considered in our experiments.

We are currently investigating issues related to the work presented. The FCLT in Section 2 for the p -quantile estimator for a single p obtained with CMC i.i.d. sampling is established under conditions that are probably stronger than necessary, and we are working on weaker assumptions ensuring the FCLT. We are also studying applying other STS methods for constructing CIs for quantiles when applying crude Monte Carlo with i.i.d. sampling. In particular, Tables 1 and 2 show that the average half-widths can differ significantly for the various STS scaling functions, and we want to identify scaling functions that result in good coverage with small intervals. Another research direction is using STS methods for building CI for quantiles estimated using variance-reduction techniques (VRTs). This entails first establishing that the resulting p -quantile estimator satisfies a FCLT as the sample size $n \rightarrow \infty$ for fixed $0 < p < 1$; ordinary CLTs have been shown for various VRTs, e.g., in Chu and Nakayama (2012). In addition we are exploring STS methods for developing CIs for steady-state quantiles (?).

ACKNOWLEDGMENTS

This work has been supported in part by the National Science Foundation under Grants No. CMMI-0926949 and CMMI-1200065. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation.

REFERENCES

- Adlakha, V. G., and V. G. Kulkarni. 1989. "A classified bibliography of research on stochastic PERT networks". *INFOR* 27:272–296.
- Alexopoulos, C., N. T. Argon, D. Goldsman, N. M. Steiger, G. Tokol, and J. R. Wilson. 2007. "Efficient Computation of Overlapping Variance Estimators for Simulation". *INFORMS* 19:314–327.
- Alexopoulos, C., N. T. Argon, D. Goldsman, G. Tokol, and J. R. Wilson. 2007. "Overlapping Variance Estimators for Simulation". *Operations Research* 55:1090–1103.
- Asmussen, S., and P. Glynn. 2007. *Stochastic Simulation: Algorithms and Analysis*. 1st ed. New York: Springer.
- Billingsley, P. 1999. *Convergence of Probability Measures*. Second ed. New York: John Wiley & Sons.
- Bloch, D. A., and J. L. Gastwirth. 1968. "On a simple estimate of the reciprocal of the density function". *Annals of Mathematical Statistics* 39:1083–1085.
- Bofinger, E. 1975. "Estimation of a density function using order statistics". *Australian Journal of Statistics* 17:1–7.
- Calvin, J. M. 2004. "Simulation output analysis based on excursions". In *Proceedings of the 2004 Winter Simulation Conference*, edited by R. G. Ingalls, M. D. Rossetti, J. S. Smith, and B. A. Peters, 681–684. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers.
- Calvin, J. M., and M. K. Nakayama. 2006. "Permuted standardized time series for steady-state simulations". *Mathematics of Operations Research* 31:351–368.
- Chu, F., and M. K. Nakayama. 2012. "Confidence Intervals for Quantiles When Applying Variance-Reduction Techniques". *ACM Transactions On Modeling and Computer Simulation* 36:Article 7 (25 pages plus 12–page online–only appendix).
- Csörgő, M. 1983. *Quantile Processes with Statistical Applications*. Philadelphia, PA: SIAM.
- Csörgő, M., and P. Révész. 1981. *Strong Approximations in Probability and Statistics*. New York: Academic Press.

- Csörgő, M., and B. Szyszkowicz. 1998. "Sequential quantile and Bahadur-Kiefer Processes". In *Handbook of Statistics 16, Order Statistics: Theory and Methods*, edited by N. Balakrishnan and C. Rao, 631–688. Amsterdam: Elsevier Science B.V.
- Efron, B. 1979. "Bootstrap methods: Another look at the jackknife". *Annals of Statistics* 7:1–26.
- Glynn, P. W., and D. L. Iglehart. 1990. "Simulation output analysis using standardized time series". *Mathematics of Operations Research* 15:1–16.
- Glynn, P. W., and W. Whitt. 1991. "Estimating the Asymptotic Variance with Batch Means". *Operations Research Letters* 10:431–435.
- Goldsman, D., M. S. Meketon, and L. W. Schruben. 1990. "Properties of standardized time series weighted area variance estimators". *Management Science* 36:602–612.
- Goldsman, D., and L. W. Schruben. 1990. "New confidence interval estimators using standardized time series". *Management Science* 36:393–397.
- Hogg, R. V., and A. T. Craig. 1978. *Introduction to Mathematical Statistics, Fourth Edition*. Macmillan.
- Hsu, J. C., and B. L. Nelson. 1990. "Control variates for quantile estimation". *Management Science* 36:835–851.
- Jorion, P. 2007. *Value at Risk: The New Benchmark for Managing Financial Risk, 3rd Edition*. McGraw-Hill.
- Meketon, M. S., and B. Schmeiser. 1984. "Overlapping batch means: Something for nothing?". In *Proceedings of the 1984 Winter Simulation Conference*, edited by S. Sheppard, U. Pooch, and D. Pegden, 227–230. Institute of Electrical and Electronics Engineers, Piscataway, NJ.
- Nakayama, M. K. 2012. "Using Sectioning to Construct Confidence Intervals for Quantiles When Applying Importance Sampling". In *Proceedings of the 2012 Winter Simulation Conference*, edited by C. Laroque, J. Himmelspach, R. Pasupathy, O. Rose, and A. M. Uhrmacher, to appear. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers.
- Schruben, L. W. 1983. "Confidence interval estimation using standardized time series". *Operations Research* 31:1090–1108.
- Serfling, R. J. 1980. *Approximation Theorems of Mathematical Statistics*. New York: John Wiley & Sons.
- U.S. Nuclear Regulatory Commission 1989. "Best-Estimate Calculations of Emergency Core Cooling Performance". Nuclear Regulatory Commission Regulatory Guide 1.157, Washington, DC.
- Wand, M. P., and M. C. Jones. 1995. *Kernel Smoothing*. London: Chapman & Hall.

AUTHOR BIOGRAPHIES

JAMES M. CALVIN is a professor in the Department of Computer Science at the New Jersey Institute of Technology. He received a Ph.D. in operations research from Stanford University and is an associate editor for *ACM Transactions on Modeling and Computer Simulation* and *INFORMS Journal on Computing*. Besides simulation output analysis, his research interests include global optimization and probabilistic analysis of algorithms. His email address is calvin@njit.edu.

MARVIN K. NAKAYAMA is a professor in the Department of Computer Science at the New Jersey Institute of Technology. He received a Ph.D. in operations research from Stanford University. He won second prize in the 1992 George E. Nicholson Student Paper Competition sponsored by INFORMS and is a recipient of a CAREER Award from the National Science Foundation. He is the Stochastic Models Area Editor for *ACM Transactions on Modeling and Computer Simulation* and the Simulation Area Editor for *INFORMS Journal on Computing*. His email address is marvin@njit.edu.

Table 1: CIs were constructed using different consistent and nonoverlapping STS methods, and the coverage levels (and average half-widths) were estimated from 10^4 independent replications.

n	FD	Kernel	Batch.	Sect.	STS			Exact
					WA($w_0, 1$)	WA($w_0, 10$)	Range	
$p = 0.50$								
100	0.896 (0.356)	0.873 (0.317)	0.782 (0.332)	0.903 (0.349)	0.887 (0.965)	0.883 (0.352)	0.876 (0.360)	0.903 (0.325)
400	0.882 (0.163)	0.889 (0.161)	0.869 (0.174)	0.902 (0.177)	0.891 (0.499)	0.895 (0.178)	0.897 (0.178)	0.901 (0.163)
1600	0.891 (0.081)	0.899 (0.081)	0.893 (0.088)	0.902 (0.088)	0.895 (0.246)	0.895 (0.088)	0.901 (0.088)	0.899 (0.081)
6400	0.894 (0.041)	0.901 (0.041)	0.897 (0.044)	0.900 (0.044)	0.898 (0.124)	0.904 (0.044)	0.898 (0.044)	0.903 (0.041)
$p = 0.80$								
100	0.903 (0.564)	0.836 (0.469)	0.663 (0.479)	0.883 (0.521)	0.888 (1.524)	0.915 (0.634)	0.873 (0.557)	0.898 (0.500)
400	0.880 (0.250)	0.877 (0.245)	0.833 (0.262)	0.901 (0.270)	0.887 (0.764)	0.904 (0.287)	0.895 (0.278)	0.902 (0.250)
1600	0.880 (0.122)	0.893 (0.124)	0.884 (0.135)	0.902 (0.136)	0.895 (0.379)	0.903 (0.138)	0.896 (0.137)	0.900 (0.125)
6400	0.893 (0.062)	0.898 (0.062)	0.897 (0.068)	0.901 (0.068)	0.898 (0.191)	0.902 (0.068)	0.904 (0.068)	0.900 (0.063)
$p = 0.95$								
100	0.947 (1.443)	0.708 (0.697)	0.858 (0.910)	0.862 (0.937)	0.874 (3.039)	0.823 (0.939)	0.852 (1.082)	0.907 (0.951)
400	0.901 (0.506)	0.828 (0.436)	0.670 (0.457)	0.893 (0.497)	0.885 (1.470)	0.931 (0.651)	0.883 (0.538)	0.904 (0.476)
1600	0.895 (0.241)	0.877 (0.231)	0.835 (0.250)	0.900 (0.257)	0.895 (0.728)	0.911 (0.278)	0.895 (0.265)	0.901 (0.238)
6400	0.900 (0.119)	0.894 (0.118)	0.881 (0.127)	0.901 (0.129)	0.894 (0.359)	0.906 (0.132)	0.898 (0.129)	0.905 (0.119)
$p = 0.99$								
100	0.507 (1.109)	0.373 (0.585)	0.037 (0.910)	0.709 (1.389)	0.855 (8.293)	0.536 (0.939)	0.785 (2.343)	0.945 (2.025)
400	0.924 (1.430)	0.648 (0.620)	0.729 (0.874)	0.837 (0.920)	0.878 (3.355)	0.803 (0.918)	0.849 (1.172)	0.915 (1.012)
1600	0.981 (0.824)	0.796 (0.441)	0.902 (0.547)	0.905 (0.562)	0.888 (1.595)	0.885 (0.569)	0.888 (0.584)	0.906 (0.506)
6400	0.936 (0.291)	0.860 (0.244)	0.887 (0.271)	0.900 (0.275)	0.891 (0.771)	0.897 (0.281)	0.892 (0.282)	0.898 (0.253)

Table 2: CIs were constructed using different methods, and the coverage levels (and average half-widths) were estimated from 2000 independent replications, where for the STS weighted area estimators, N stands for nonoverlapping and O stands for overlapping.

n	Sect.	STS Weighted Area (w, b, N or O)					
		$(w_0, 1, N)$	$(w_1, 1, N)$	$(w_0, 10, N)$	$(w_1, 10, N)$	$(w_0, 10, O)$	$(w_1, 10, O)$
$p = 0.50$							
100	0.900 (0.348)	0.884 (0.954)	0.877 (0.972)	0.877 (0.351)	0.886 (0.359)	0.909 (0.343)	0.908 (0.349)
400	0.896 (0.176)	0.885 (0.491)	0.884 (0.489)	0.893 (0.178)	0.888 (0.178)	0.908 (0.172)	0.908 (0.174)
1600	0.898 (0.089)	0.900 (0.253)	0.899 (0.248)	0.892 (0.088)	0.894 (0.089)	0.907 (0.086)	0.909 (0.087)
6400	0.907 (0.045)	0.912 (0.126)	0.899 (0.127)	0.907 (0.044)	0.911 (0.045)	0.920 (0.042)	0.922 (0.043)
$p = 0.80$							
100	0.884 (0.522)	0.876 (1.559)	0.884 (1.575)	0.918 (0.631)	0.868 (0.516)	0.939 (0.615)	0.885 (0.504)
400	0.902 (0.273)	0.881 (0.757)	0.887 (0.783)	0.904 (0.290)	0.894 (0.281)	0.924 (0.278)	0.918 (0.272)
1600	0.896 (0.136)	0.890 (0.379)	0.876 (0.372)	0.907 (0.139)	0.902 (0.139)	0.914 (0.134)	0.920 (0.135)
6400	0.901 (0.069)	0.903 (0.194)	0.902 (0.193)	0.895 (0.068)	0.901 (0.069)	0.903 (0.065)	0.905 (0.066)
$p = 0.95$							
100	0.861 (0.936)	0.868 (2.966)	0.856 (2.835)	0.824 (0.930)	0.793 (0.871)	0.867 (0.937)	0.843 (0.880)
400	0.890 (0.494)	0.885 (1.475)	0.878 (1.452)	0.935 (0.648)	0.893 (0.527)	0.959 (0.630)	0.915 (0.520)
1600	0.905 (0.257)	0.887 (0.719)	0.883 (0.711)	0.913 (0.280)	0.906 (0.268)	0.928 (0.268)	0.920 (0.260)
6400	0.902 0.889 (0.130)	0.897 (0.357)	0.899 (0.364)	0.890 (0.131)	0.913 (0.131)	0.912 (0.127)	(0.128)
$p = 0.99$							
100	0.699 (1.379)	0.865 (8.517)	0.796 (5.365)	0.533 (0.930)	0.500 (0.871)	0.554 (0.937)	0.528 (0.880)
400	0.825 (0.909)	0.876 (3.320)	0.861 (3.015)	0.792 (0.905)	0.741 (0.801)	0.833 (0.906)	0.788 (0.813)
1600	0.908 (0.561)	0.886 (1.559)	0.887 (1.537)	0.884 (0.566)	0.930 (0.690)	0.907 (0.555)	0.956 (0.677)
6400	0.896 (0.272)	0.891 (0.758)	0.889 (0.764)	0.897 (0.282)	0.902 (0.285)	0.909 (0.272)	0.907 (0.277)