A Central Limit Theorem For Empirical Quantiles in the Markov Chain Setting

Peter W. Glynn and Shane G. Henderson

Abstract We provide a new proof of a central limit theorem for empirical quantiles in the positive-recurrent Markov process setting under conditions that are essentially tight. We also establish the validity of the method of nonoverlapping batch means with a fixed number of batches for interval estimation of the quantile. The conditions of these results are likely to be difficult to verify in practice, and so we also provide more easily verified sufficient conditions.

1 Introduction

Given a real-valued random variable Y with cumulative distribution function (CDF) F, the pth quantile q (for $0 < p < 1$) is $q = F^{-1}(p) = \inf\{x : F(x) \geq p\}$. The problem of quantile estimation is, given p, to determine $q = F^{-1}(p)$.

We focus on the case where Y is a random variable associated with the steady-state regime of a Markov chain. To be more precise, let $X = (X_t : t \geq 0)$ be a positive (Harris) recurrent Markov chain on a general state space S in discrete or continuous time, and denote the stationary distribution of X by π. Let $f : S \to \mathbb{R}$ be a real-valued function defined on the state space S of X. We consider the problem of computing the pth quantile q of the random variable $Y = f(X_0)$, where X_0 has distribution π. Under mild additional conditions, the pth quantile Q_t of the empirical CDF

$$F(\cdot, t) = \frac{1}{t} \int_0^t \mathbb{1}(f(X_s) \leq \cdot)\, ds. \tag{1}$$

Peter W. Glynn
Stanford University, Stanford CA, e-mail: glynn@stanford.edu

Shane G. Henderson
Cornell University, Ithaca NY, e-mail: sgh9@cornell.edu

converges to q almost surely. Our main result is a central limit theorem (CLT) for Q_t. We further show that this CLT can be leveraged to establish the validity of the non-overlapping batch means procedure for reporting asymptotically valid confidence intervals for q.

A CLT for empirical quantiles can be established by appealing to regenerative theory. This is the approach taken in [1, 2, 3], for example, and indeed we use this approach in this paper. We exploit the "1-dependent regenerative property" of Harris processes to obtain our main results. In addition to smoothness conditions on the target CDF at the quantile q, our main assumption is that the second moment of the cycle lengths is finite. As we will show, one cannot expect the CLT to hold in general if this condition is relaxed.

Since the conditions of our main result are hard to verify in practice, we also provide more easily-verifiable conditions under which the required properties hold. These conditions are Lyapunov drift criteria, together with a condition that ensures that the target distribution is appropriately smooth at the quantile q.

Why are these particular results of interest to the simulation community? It is known that any discrete-event simulation that is "well-posed", in a certain precise sense, can be modelled as a positive Harris recurrent Markov chain [4]. If the state space of the simulation is continuous, as is often the case, then the analysis in this paper is relevant. To buttress this point we provide an example in Section 6.

Another application area where this problem is of great interest is in Markov chain Monte Carlo (MCMC); see, e.g., [5], [6, Chapter 5], [7, Chapter XIII] and especially [3]. In this setting, one is typically interested in exploring a given distribution π that is known only up to a normalizing constant. A Markov chain may be produced whose steady-state distribution is the given distribution π, and one then attempts to infer properties of the distribution π from Markov chain simulations. Unlike most work in MCMC, we neither assume nor require reversibility.

Quantile estimation has received a great deal of attention in the simulation community. In the case where the observations are i.i.d., [8] developed a number of important results including bias expansions that expand on the general theory for the i.i.d. case available in, e.g., [9, Section 2.3]. [10] derived large-deviations results for quantile estimators and explored the use of stratification techniques in estimating quantiles. Other papers that explore the use of variance reduction techniques in quantile estimation for i.i.d. observations include [11, 12, 13, 14, 15]. Additional work that develops sufficient conditions for quantile estimators that employ variance reduction techniques to satisfy a CLT includes [16, 17, 18].

In the case of estimating steady-state quantiles as in the present paper, work includes [19], where sufficient conditions for the validity of the method of nonoverlapping batch quantiles are presented, along with a practical algorithm for providing a confidence interval for a quantile. Additional practical algorithms may be found in [20, 21, 22]. Asymptotic results for the method of overlapping batch means are stated in [23]. In closely related work, [24] gives sufficient conditions for the quantile estimator to satisfy a central limit theorem in the Markov-chain setting. The sufficient conditions ensure that the chain is geometrically ergodic through the use of Lyapunov drift criteria, along with additional conditions on the time-dependent distribution of the chain. In contrast, our conditions are much weaker and we do not require conditions on the time-dependent distribution. The sufficient conditions of [24] permit a comparison of the bias and mean-squared error of 3 estimators of steady-state quantiles in [25].

Perhaps the closest work to ours is [3], though that paper has a more practical focus on estimation methods while we strive for minimal conditions for the CLT. [3] establishes the quantile CLT and describes how to estimate the variance constant that appears in the CLT using both batch means and regenerative methods. The central assumption there is *polynomial ergodicity* of an order strictly greater than 1, which implies that the length of regenerative cycles in Harris chains have a finite $(2 + \epsilon)$ moment for some $\epsilon > 0$ (see the proof of Theorem 5 in [3]), while we only require a finite 2nd moment. Moreover, polynomially ergodic chains are necessarily aperiodic; we do not require aperiodicity. Finally, [3] assumes independent regenerative cycles, yet some Harris chains arising in practice cannot have independent cycles [26].

In early work, [27, 28] established CLTs and laws of the iterated logarithm for empirical quantiles obtained from ϕ-mixing stochastic processes. Given that ergodic Markov chains are strong mixing [29, 30] one could apply these results to the Markov setting. However, we believe that the hypotheses of these results are difficult to verify in practice. The assumptions of [31] are more readily verified and were employed in the quantile estimation context by [19], but may require stronger conditions than does our analysis. For example, in the single-server example in Section 6 where we require a finite second moment condition, [32] instead requires a finite moment-generating function to verify a key assumption in [31]. Still, the machinery of [31] may be more directly applicable to some stochastic processes than ours, so the two approaches are complementary.

The remainder of this paper is organized as follows. Section 2 proves a CLT for empirical quantiles under very general hypotheses. The key hypothesis there is a uniform CLT for the empirical distribution function in a neighbourhood of the true quantile. Section 3 proves a uniform CLT for 1-dependent sequences. Section 4 specializes the results of the previous sections to obtain the desired quantile CLT for Harris processes in discrete or continuous time. Section 5 establishes the validity of non-overlapping batch means, partly through the development of a Bahadur-Ghosh representation of the quantile estimator, which may be of independent interest. Finally, Section 6 gives some sufficient conditions for the quantile CLT to hold, and presents a small example.

2 A Quantile Central Limit Theorem

Given a real-valued stochastic process $(W(t) : t \geq 0)$, let

$$F(\cdot, t) = t^{-1} \int_0^t \mathbb{1}(W(s) \leq \cdot)\, ds$$

be its empirical CDF. For a real-valued process $(W_k : k = 0, 1, \ldots)$ in discrete time, define $W(t) = W_{\lfloor t \rfloor}$ and $F(\cdot, t)$ as above.

For any fixed $x \in \mathbb{R}$, we say that $F(x, \cdot)$ satisfies a CLT if there exist constants $\sigma^2(x) > 0, F(x)$ such that for any $y \in \mathbb{R}$

$$\mathbb{P}\left(\frac{t^{1/2}[F(x,t) - F(x)]}{\sigma(x)} \le y\right) - \Phi(y) \to 0$$

as $t \to \infty$, where Φ denotes the distribution function of a standard normal random variable. If this CLT holds, then the pointwise convergence is uniform in y, i.e.,

$$\sup_y \left|\mathbb{P}\left(\frac{t^{1/2}[F(x,t) - F(x)]}{\sigma(x)} \le y\right) - \Phi(y)\right| \to 0$$

as $t \to \infty$; see, e.g., [9, p. 18].

We say that $F(\cdot, \cdot)$ satisfies a CLT uniformly in the set N if

$$\sup_{x \in N} \sup_y \left|\mathbb{P}\left(\frac{t^{1/2}[F(x,t) - F(x)]}{\sigma(x)} \le y\right) - \Phi(y)\right| \to 0 \tag{2}$$

as $t \to \infty$.

Theorem 1 *Fix $q \in \mathbb{R}$ and suppose that $F(\cdot, \cdot)$ satisfies a CLT uniformly in an open neighborhood N of q. Suppose further that $F(\cdot)$ is differentiable at q, $F'(q) > 0$, $\sigma^2(q) > 0$ and $\sigma^2(\cdot)$ is continuous at q. Let $p = F(q)$ and let $Q_t = F^{-1}(p, t) = \inf\{x : F(x, t) \ge p\}$ be the pth quantile of $F(\cdot, t)$. Let*

$$G(y, t) = \mathbb{P}\left[\frac{\sqrt{t}(Q_t - q)}{\sigma(q)/F'(q)} \le y\right].$$

Then $G(\cdot, t) \Rightarrow \Phi$ as $t \to \infty$.

Proof We employ a similar proof to the one for empirical quantiles in the i.i.d. case given in [9, p. 78]. Define $q_t = q + t^{-1/2}\sigma(q)y/F'(q)$. Then

$$G(y, t) = \mathbb{P}[Q_t \le q_t] = \mathbb{P}[p \le F(q_t, t)],$$

since for any cumulative distribution function H and arbitrary real x and $u \in (0, 1)$, $H^{-1}(u) \le x$ if and only if $u \le H(x)$ [9, Lemma 1.1.4(iii)]. Now,

$$\begin{aligned} G(y, t) &= \mathbb{P}\left[t^{1/2}\frac{F(q_t, t) - F(q_t)}{\sigma(q_t)} \ge t^{1/2}\frac{p - F(q_t)}{\sigma(q_t)}\right] \\ &= \mathbb{P}(U(q_t, t) \ge -y_t), \end{aligned}$$

where

$$U(z, t) = \frac{t^{1/2}(F(z, t) - F(z))}{\sigma(z)} \quad \text{and} \quad y_t = \frac{t^{1/2}[F(q_t) - p]}{\sigma(q_t)},$$

and so

$$\begin{aligned}\Phi(y) - G(y,t) &= \mathbb{P}[U(q_t,t) < -y_t] - (1-\Phi(y)) \\ &= [\mathbb{P}[U(q_t,t) < -y_t] - \Phi(-y_t)] + [\Phi(y) - \Phi(y_t)]. \qquad (3)\end{aligned}$$

We now show that the two bracketed terms in (3) converge to 0 as $t \to \infty$.

For the first term in (3), for t sufficiently large that $q_t \in N$,

$$|\mathbb{P}(U(q_t,t) < -y_t) - \Phi(-y_t)| \leq \sup_{x\in N} \sup_{-\infty<w<\infty} |\mathbb{P}(U(x,t) < w) - \Phi(w)|.$$

The uniform CLT assumption ensures that this term converges to 0 as $t \to \infty$.

To show that the second term in (3) converges to 0, it suffices to show that $y_t \to y$ as $t \to \infty$. Since F is differentiable at q,

$$F(q_t) - p = F(q_t) - F(q) = F'(q)(q_t - q) + o(q_t - q),$$

where a quantity r_t is said to be $o(h_t)$ if $r_t/h_t \to 0$ as $t \to \infty$. Thus,

$$y_t = \frac{F'(q)t^{1/2}(q_t - q)}{\sigma(q_t)} = \frac{\sigma(q)y + o(1)}{\sigma(q_t)}$$

as $t \to \infty$. Since $\sigma(q_t) \to \sigma(q) > 0$ as $t \to \infty$, $y_t \to y$ as $t \to \infty$.

3 A Uniform CLT for 1-Dependent Sequences

The key ingredient in Theorem 1 is the uniform CLT. In this section we establish a uniform CLT for 1-dependent processes. We then apply this result to Harris processes in Section 4.

Let $Z(\theta) = (Z_n(\theta) : n \geq 1)$ be a stationary sequence of real-valued, 1-dependent, mean 0 random variables for each $\theta \in \Theta$. Let

$$S_n(\theta) = \sum_{i=1}^{n} Z_i(\theta).$$

It is well known that if $EZ_1^2(\theta) < \infty$ and

$$\eta^2(\theta) = EZ_1^2(\theta) + 2EZ_1(\theta)Z_2(\theta) > 0$$

then as $n \to \infty$, we have the CLT

$$\sup_{y\in\mathbb{R}} \left| \mathbb{P}\left(\frac{S_n(\theta)}{\eta(\theta)\sqrt{n}} \leq y \right) - \Phi(y) \right| \to 0.$$

We seek a uniform (in θ) version of this result, which requires a linking assumption.

A1 The family of random variables $(Z_1^2(\theta) : \theta \in \Theta)$ is uniformly integrable.

Theorem 2 *Let $Z(\theta)$ and $\eta^2(\theta)$ be defined as above for all $\theta \in \Theta$. Suppose that Assumption A1 holds and $\eta^2(\theta)$ is bounded away from 0 for $\theta \in \Theta$. Then, as $n \to \infty$,*

$$\sup_{\theta\in\Theta}\sup_{y\in\mathbb{R}}\left|\mathbb{P}\left(\frac{S_n(\theta)}{\eta(\theta)\sqrt{n}} \le y\right) - \Phi(y)\right| \to 0.$$

We need some preliminary results before proving Theorem 2. First, we show that in proving a uniform CLT, we can ignore terms that are uniformly small in θ. The proof is an extension of that of the converging together lemma [9, p. 19] and omitted.

Lemma 1 *Let $U_n(\theta), V_n(\theta), W_n(\theta)$ and $X_n(\theta)$ be real-valued random variables for all $n \ge 1$ and all $\theta \in \Theta$. Suppose that for all $n \ge 1$ and all $\theta \in \Theta$, $X_n(\theta) = U_n(\theta)V_n(\theta) + W_n(\theta)$. Suppose that for all $\epsilon > 0$,*

$$\lim_{n\to\infty}\sup_{\theta\in\Theta}\mathbb{P}(|U_n(\theta) - 1| > \epsilon) = 0 \quad \textit{and} \quad \lim_{n\to\infty}\sup_{\theta\in\Theta}\mathbb{P}(|W_n(\theta)| > \epsilon) = 0, \textit{ and}$$

$$\lim_{n\to\infty}\sup_{\theta\in\Theta}\sup_{y}|\mathbb{P}(V_n(\theta) \le y) - G(y)| = 0$$

for some distribution function G. Then $(X_n(\cdot) : n \ge 1)$ satisfies

$$\lim_{n\to\infty}\sup_{\theta\in\Theta}\sup_{y}|\mathbb{P}(X_n(\theta) \le y) - G(y)| = 0.$$

Lemma 2 is a special case of Theorem 18.1 and Corollary 18.3 of [33].

Lemma 2 *Suppose that $(U_n : n \ge 1)$ is an i.i.d. sequence of r.v.'s with mean 0 and variance 1, and let $\mathcal{N}$ denote a standard normal random variable. Let $g(x, a) = x^2 I(|x| > a)$, and G_n denote the distribution function of $n^{-1/2}\sum_{i=1}^n U_i$. Then*

1. $\left|\mathbb{E}g\left(\frac{1}{\sqrt{n}}\sum_{i=1}^n U_i, a\right) - \mathbb{E}g(\mathcal{N}, a)\right| \le c\delta_n$, *and*
2. $\sup_y |G_n(y) - \Phi(y)| \le c\delta_n$,

where the constant c does not depend on a, n, or the distribution of U_1, and

$$\delta_n = \inf_{\epsilon\in[0,1]}\left(\epsilon + \mathbb{E}[U_1^2; U_1^2 > n\epsilon^2]\right).$$

Let $\tilde{Z}(\theta) = (\tilde{Z}_n(\theta) : n \ge 1)$ be an i.i.d. sequence of real-valued random variables with $\tilde{Z}_1(\theta)$ having the same distribution as $Z_1(\theta)$ for all $\theta \in \Theta$. Define the variance $\gamma^2(\theta) = \mathbb{E}Z_1^2(\theta)$, and for $n \ge 1$ let $\tilde{S}_n(\theta) = \sum_{i=1}^n \tilde{Z}_i(\theta)$.

Lemma 3 (Uniform integrability assuming independence)
Under the conditions of Theorem 2, as $n \to \infty$,

$$\sup_{\theta\in\Theta}\mathbb{E}\left[\frac{\tilde{S}_n^2(\theta)}{n\gamma^2(\theta)}; \tilde{S}_n^2(\theta) > n\gamma^2(\theta)a_n\right] \to 0$$

for any sequence of positive constants $\{a_n\}$ with the property that $a_n \to \infty$ as $n \to \infty$.

Proof For $a > 0$, define $g(x, a) = x^2 I(|x| > a)$, and let $\mathcal{N}$ denote a standard normal random variable. Since $\mathbb{E}\mathcal{N}^2 < \infty$, $\mathbb{E}g(\mathcal{N}, a_n) \to 0$ as $n \to \infty$. Part 1 of Lemma 2 implies that

$$\left|\mathbb{E}g\left(\frac{\tilde{S}_n(\theta)}{\sqrt{n}\gamma(\theta)}, a_n\right) - \mathbb{E}g(\mathcal{N}, a_n)\right| \le c\delta_n(\theta),$$

where c is a constant that does not depend on a_n, n or θ and

$$\delta_n(\theta) = \inf_{\epsilon\in[0,1]}\left(\epsilon + \mathbb{E}\left[\frac{\tilde{Z}_1^2(\theta)}{\gamma^2(\theta)}; \tilde{Z}_1^2(\theta) > n\epsilon^2\gamma^2(\theta)\right]\right). \tag{4}$$

We assumed that $\eta^2(\theta)$ is bounded away from 0, and therefore so is $\gamma^2(\theta)$, since

$$\eta^2(\theta) = \gamma^2(\theta) + 2\mathbb{E}Z_1(\theta)Z_2(\theta) \le \gamma^2(\theta) + \mathbb{E}Z_1^2(\theta) + \mathbb{E}Z_2^2(\theta) = 3\gamma^2(\theta).$$

Let $\gamma^2 > 0$ be a lower bound on $\gamma^2(\theta)$ over $\theta \in \Theta$. It follows that the second term in the infimum in (4) is bounded above by

$$\gamma^{-2}\mathbb{E}[\tilde{Z}_1^2(\theta); \tilde{Z}_1^2(\theta) > n\epsilon^2\gamma^2] = \gamma^{-2}\mathbb{E}[Z_1^2(\theta); Z_1^2(\theta) > n\epsilon^2\gamma^2].$$

If we now choose $\epsilon = \epsilon(n)$ in such a way that $n\epsilon^2(n) \to \infty$ and $\epsilon(n) \to 0$ as $n \to \infty$, then A1 ensures that $\sup_{\theta\in\Theta}\delta_n(\theta) \to 0$ as $n \to \infty$, proving the result. □

Lemma 4 (Uniform integrability assuming 1-dependence)

Under the conditions of Theorem 2, as $n \to \infty$,

$$\sup_{\theta\in\Theta}\mathbb{E}\left[\frac{S_n^2(\theta)}{n\eta^2(\theta)}; S_n^2(\theta) > n\eta^2(\theta)a_n\right] \to 0$$

for any sequence of positive constants $\{a_n\}$ with the property that $a_n \to \infty$ as $n \to \infty$.

Proof We can write

$$\begin{aligned} S_n(\theta) &= \sum_{i=1, i \text{ odd}}^{n} Z_i(\theta) + \sum_{i=1, i \text{ even}}^{n} Z_i(\theta) \\ &= \tilde{S}_n(\theta, 1) + \tilde{S}_n(\theta, 2) \text{ (say)}. \end{aligned}$$

Let $M_n(\theta) = \max\{|\tilde{S}_n(\theta, 1)|, |\tilde{S}_n(\theta, 2)|\}$ so that $S_n^2(\theta) \le [2M_n(\theta)]^2$. Now, $|S_n(\theta)| > u$ implies that $|\tilde{S}_n(\theta, i)| > u/2$ for at least one of $i = 1, 2$, which is, in turn, equivalent to $M_n(\theta) > u/2$. Thus,

$$\begin{aligned} \mathbb{E}\left[\frac{S_n^2(\theta)}{n\eta^2(\theta)}; S_n^2(\theta) > n\eta^2(\theta)a_n\right] &\le \mathbb{E}\left[\frac{4M_n^2(\theta)}{n\eta^2(\theta)}; M_n^2(\theta) > n\eta^2(\theta)a_n/4\right] \\ &\le \sum_{i=1}^{2}\mathbb{E}\left[\frac{4\tilde{S}_n^2(\theta, i)}{n\eta^2(\theta)}; \tilde{S}_n^2(\theta, i) > n\eta^2(\theta)a_n/4\right]. \end{aligned} \tag{5}$$

We now apply Lemma 3 to each of the summands in (5) to complete the proof. We use the fact that each of the summands consists of essentially $n/2$ terms, and also that $\gamma^2(\theta)/\eta^2(\theta)$ is bounded away from 0 and bounded above. □

Proof (of Theorem 2) We use the "big block, little block" argument (e.g., [34, Theorem 7.3.1]) to reduce the problem for 1-dependent summands to one for independent summands. The big blocks are sums of consecutive $Z_i(\theta)$s, which are separated by small blocks of size 1 that ensure, together with 1-dependence, that the big blocks are independent. When the big blocks grow at an appropriate rate with n, the result follows. Let $m_n = \lfloor n^\alpha \rfloor$ be the size of the blocks, where $\alpha \in (0, 1)$. Let $k_n = \lfloor n/m_n \rfloor$ be the number of big blocks. For $1 \le j \le k_n$, define the jth big block to be

$$\Gamma_j(\theta, n) = \sum_{i=(j-1)m_n+1}^{jm_n-1} Z_i(\theta).$$

Then for $n \ge 1$,

$$\begin{aligned} S_n(\theta) &= \sum_{j=1}^{k_n} \Gamma_j(\theta, n) + \sum_{j=1}^{k_n} Z_{jm_n}(\theta) + \sum_{i=k_n m_n+1}^{n} Z_i(\theta) \\ &= S'_n(\theta) + S''_n(\theta) + S'''_n(\theta) \text{ say.} \end{aligned}$$

The hypothesis of 1-dependence ensures that for n sufficiently large, the $\Gamma_j(\theta, n)$s are i.i.d. (in j). Furthermore, so are the $Z_{jm_n}(\theta)$s provided that $m_n > 1$, which is again assured for n large enough. For any $\epsilon > 0$ and n sufficiently large that $m_n > 1$,

$$\mathbb{P}\left(\frac{|S''_n(\theta)|}{\eta(\theta)\sqrt{n}} > \epsilon\right) \le \frac{\mathbb{E}S''_n(\theta)^2}{n\epsilon^2\eta^2(\theta)} \le \frac{k_n\gamma^2(\theta)}{n\epsilon^2\underline{\eta}^2}, \tag{6}$$

where $\underline{\eta}^2 > 0$ is a lower bound on $\eta^2(\theta)$ over $\theta \in \Theta$. Assumption A1 implies that $\gamma^2(\theta)$ is bounded above, so that (6) converges to 0 uniformly in $\theta \in \Theta$ as $n \to \infty$. Similarly, we can show that $S'''_n(\theta)$ does not figure in the asymptotics (uniformly in $\theta \in \Theta$). So by Lemma 1 it suffices to show a uniform CLT for $n^{-1/2}S'_n(\theta)/\eta(\theta)$. Let

$$v_n^2(\theta) = \operatorname{Var}\Gamma_j(\theta, n) = (m_n - 1)\eta^2(\theta) - 2EZ_1(\theta)Z_2(\theta)$$

be the variance of the big blocks. Applying Part 2 of Lemma 2 to a normalized version of $S'_n(\theta)$, we get

$$\sup_{y\in\mathbb{R}} \left| \mathbb{P}\left(\frac{S'_n(\theta)}{v_n(\theta)\sqrt{k_n}} \le y\right) - \Phi(y) \right| \le c\delta_n(\theta),$$

where c is a constant that does not depend on n or θ, and

$$\delta_n(\theta) = \inf_{\epsilon\in[0,1]} \left(\epsilon + \mathbb{E}\left[\frac{\Gamma_1^2(\theta, n)}{v_n^2(\theta)}; \Gamma_1^2(\theta, n) > k_n\epsilon^2 v_n^2(\theta)\right]\right). \tag{7}$$

Now choose $\epsilon = \epsilon(n)$ in such a way that $k_n \epsilon^2(n) \to \infty$ and $\epsilon(n) \to 0$ as $n \to \infty$. We then apply Lemma 4 to the second term in the infimum in (7), using the facts that $\Gamma_1(\theta, n) = S_{m_n}(\theta)$ and $v_n^2(\theta)/(m_n \eta^2(\theta)) \to 1$ as $n \to \infty$ uniformly in θ. We can then conclude that (7) converges to 0 uniformly in θ as $n \to \infty$.

To complete the proof, observe that

$$\frac{S_n'(\theta)}{v_n(\theta)\sqrt{k_n}} - \frac{S_n'(\theta)}{\eta(\theta)\sqrt{n}} = \frac{\beta_n(\theta) S_n'(\theta)}{v_n(\theta)\sqrt{k_n}},$$

where $\beta_n(\theta) \to 0$ as $n \to \infty$ uniformly in θ. Thus,

$$\mathbb{P}\left(\left|\frac{\beta_n(\theta) S_n'(\theta)}{v_n(\theta)\sqrt{k_n}}\right| > \epsilon\right) \leq \frac{\mathbb{E} S_n'(\theta)^2}{k_n v_n^2(\theta)} \frac{\beta_n^2(\theta)}{\epsilon^2} = \frac{\beta_n^2(\theta)}{\epsilon^2} \to 0$$

as $n \to \infty$, uniformly in θ. The result now follows from Lemma 1. □

4 A Quantile Central Limit Theorem for Harris Processes

We now specialize the preceding results to positive-recurrent Harris processes X on state space S in both discrete and continuous time. These processes possess 1-dependent structure that we exploit. Suppose that S is a complete, separable metric space equipped with Borel sigma algebra $\mathcal{S}$. We assume without further comment that if X is a continuous-time process, then it is non-explosive and strong Markov, and that its sample paths are right-continuous with left limits. (See [35, pp. 198–206, 407–410] for background.) Let $\mathbb{P}_x$ and $\mathbb{E}_x$ be the probability and expectation over path space when $X_0 = x$. We first define a Harris chain in discrete time.

Definition 1 We say that $X = (X_n : n = 0, 1, 2, \ldots)$ is a Harris chain on $(S, \mathcal{S})$ if there exists a set $C \in \mathcal{S}$, a $\gamma > 0$, a probability measure φ and an $m \geq 1$ such that

A2 $\mathbb{P}_x(X_m \in A) \geq \gamma\varphi(A)$ for all $x \in C$ and all $A \in \mathcal{S}$, and
A3 $\mathbb{P}_x(\sum_{n=0}^{\infty} I(X_n \in C) = \infty) = 1$ for all $x \in S$.

Harris processes in continuous time can be defined as follows.

Definition 2 We say that $X = (X_t : t \in [0, \infty))$ is a Harris process on $(S, \mathcal{S})$ if there exists a probability measure ν on $(S, \mathcal{S})$ such that whenever $\nu(A) > 0$, $\mathbb{P}_x(\int_{t=0}^{\infty} \mathbb{1}(X_t \in A) = \infty) = 1$ for all $x \in S$.

A Harris process X in discrete or continuous time automatically possesses a unique (up to a multiplicative constant) stationary measure π. If $\pi(S) < \infty$, then we can normalize π to a probability and we then say that X is positive Harris recurrent.

Harris processes are regenerative. For Harris chains (in discrete time), regeneration times can be defined through the famous split-chain construction; see [36] for a complete treatment. For Harris processes (in continuous time), regeneration times can be defined using the fact that Harris processes in continuous time observed at the event times of an independent homogeneous Poisson process are Harris chains (let us call the resulting chain the *sampled chain*), and then using the split-chain construction as discussed in [37]; see also [35, p. 199]. (Asmussen uses a non-standard definition of Harris recurrence in continuous time, but the basic ideas are present.) Here we sketch the key ideas behind this construction of regeneration times, as we will need the construction later.

Let $(\Lambda(i) : i \geq 0)$ be the event times in a homogeneous Poisson process that is independent of X, where $\Lambda(0) = 0$, and let $N(t) = \max\{i \geq 0 : \Lambda(i) \leq t\}$ for $t \geq 0$ be the associated counting process. Define $\tilde{X}_i = X_{\Lambda(i)}$ for $i \geq 0$. Then $\tilde{X} = (\tilde{X}_i : i \geq 0)$ is an embedded discrete-time Harris chain.

Proposition 1 *Let $(\tilde{X}_n : n = 0, 1, 2, \ldots)$ be the sampled chain as constructed above from a unit-rate Poisson process. Then we may assume that A2 holds with $m = 1$.*

Proof Sample the Harris process $X = (X_t : t \in [0, \infty))$ at the event times of a Poisson process with rate 2 that is independent of X to obtain a sampled chain $\hat{X} = (\hat{X}_n : n = 0, 1, 2, \ldots)$. The sampled chain $\hat{X}$ then satisfies A2 for some $m \geq 1$, C and $\gamma > 0$. Thus, for all $x \in C$, $P_x(\hat{X}_m \in \cdot) \geq \gamma\varphi(\cdot)$, i.e.,

$$\int_0^\infty \frac{2^m t^{m-1} e^{-2t}}{(m-1)!} P_x(X_t \in \cdot)\, dt \geq \gamma\varphi(\cdot).$$

We can find some $c > 0$ so that

$$ce^{-t} \geq \frac{2^m t^{m-1} e^{-2t}}{(m-1)!}$$

for all $t \geq 0$, and it follows that for all $x \in C$,

$$\int_0^\infty e^{-t} P_x(X_t \in \cdot)\, dt \geq \frac{\gamma}{c}\varphi(\cdot),$$

i.e., $P_x(\tilde{X}_1 \in \cdot) \geq (\gamma/c)\varphi(\cdot)$ for all $x \in C$, as required. □

Turning to the construction of regeneration times, if the chain is to be initiated with distribution φ then define $T(0) = 0$ (the "zeroth" regeneration time), set the number of complete regeneration cycles $\ell = 0$, a counter of "attempted splits" $n = 0$, the "wall clock time" $t = 0$ and generate $\tilde{X}_0$ from φ. Otherwise, set $T(-1) = 0$, set the number of complete regenerative cycles $\ell = -1$, the counter $n = 0$ and $t = 0$, and generate $\tilde{X}_0$ from the desired distribution of the process X at time 0. Next, generate $(\tilde{X}_1, \tilde{X}_2, \ldots, \tilde{X}_N)$, where $N = \inf\{j \geq 0 : \tilde{X}_j \in C\}$ is the (discrete) first hitting time of the set C. Also generate the (continuous) time process X up to time $\Lambda(N)$ from its appropriate conditional distribution. Next, independent of all else, set $n = n + 1$ and generate a Bernoulli random variable I_n, with $\mathbb{P}(I_n = 1) = \gamma$. If $I_n = 1$, then a regeneration occurs on the next (discrete-time) step, so set $\ell = \ell + 1$, set the ℓth regeneration time $T(\ell)$ equal to $\Lambda(N + 1)$, the time of the next event in the Poisson process beyond time $\Lambda(N)$, and generate $\tilde{X}_{N+1}$ according to φ. If $I_n = 0$, then generate $\tilde{X}_{N+1}$ according to $(\tilde{P}(x, \cdot) - \gamma\varphi(\cdot))/(1 - \gamma)$, where $\tilde{P}$ is the transition kernel for the sampled chain and $x = \tilde{X}_N$. Then generate the (continuous-time) intervening values $(X_s : \Lambda(N) < s < \Lambda(N+1))$ from the appropriate conditional distribution given the endpoint values. Set the "current time" $t = \Lambda(N+1)$ and repeat this process, thereby inductively constructing the continuous time process and its regeneration times $(T(k) : k \geq 0)$.

In the remainder of this section we exploit the fact that Harris processes are regenerative. In order to simultaneously treat Harris processes in both discrete and continuous time, in the remainder of this section we view a Harris chain $(X_n : n \geq 0)$ as a continuous-time process $(X_t : t \geq 0)$ where $X_t = X_{\lfloor t \rfloor}$. Such a process is no longer a Markov process, but it is regenerative.

For $i \geq 0$, let $\tau_i = T(i) - T(i-1)$ be the length of the ith regenerative cycle, and define the ith cycle to be $W_i = (X_{T(i-1)+s} : 0 \leq s < \tau_i, \tau_i)$. As discussed in [35], the cycles $(W_0, W_1, W_2, \ldots)$ are 1-dependent and the cycles $(W_1, W_2, \ldots)$ are identically distributed. This structure allows us to define the stationary measure π as follows.

For a function $g : S \to [0, \infty)$ define, for $i \geq 0$, $Y_i(g) = \int_{T(i-1)}^{T(i)} g(X_s)\, ds$. Define $Y_i(g)$ for signed g by splitting g into its positive and negative components. Now, for $A \in \mathcal{S}$, define $\pi(A) = \mathbb{E}[Y_1(\mathbb{1}(\cdot \in A))]$. Then $\pi(S) = \mathbb{E}\tau_1$, so that π has finite total mass and the process is positive Harris recurrent if and only if $\mathbb{E}\tau_1 < \infty$. We now restrict our attention to the positive Harris recurrent case, and normalize π to a probability measure by redefining $\pi(A) = \mathbb{E}[Y_1(\mathbb{1}(\cdot \in A))]/\mathbb{E}\tau_1$. Also, for $g : S \to \mathbb{R}$, define $\pi(g) = \int_S g(x)\pi(dx)$.

Now, let $f : S \to \mathbb{R}$ and for real x and $t > 0$, let $F(x, t) = t^{-1} \int_0^t \mathbb{1}(f(X_s) \leq x)\, ds$ be the empirical distribution function at time t. The strong law for positive Harris recurrent processes (see, e.g., [35, p. 203]), asserts that $F(x, t) \to F(x)$ as $t \to \infty$ almost surely, where

$$F(x) = \pi(\mathbb{1}(f(\cdot) \leq x)) = \frac{\mathbb{E} \int_{T(0)}^{T(1)} \mathbb{1}(f(X_s) \leq x)\, ds}{\mathbb{E}\tau_1}.$$

Also, let Q_t be the pth quantile associated with $F(\cdot, t)$ and q be the pth quantile of F. Our goal in this section is a CLT for Q_t.

For $t \geq 0$, let $\ell(t) = \max\{k : T(k) \leq t\}$ be the number of identically-distributed cycles completed by time t and let $\lambda = 1/\mathbb{E}\tau_1$. Also, for $i \geq 1$, define the cycle quantity

$$Z_i(x) = \int_{T(i-1)}^{T(i)} [\mathbb{1}(f(X_s) \leq x) - F(x)]\, ds.$$

Lemma 5 *Suppose that* $\mathbb{E}\tau_1^2 < \infty$. *Then*

$$\sqrt{t}(F(x, t) - F(x)) = \frac{1}{\sqrt{t}} \sum_{i=1}^{\lfloor \lambda t \rfloor} Z_i(x) + R(x, t)$$

where $\lim_{t \to \infty} \sup_x \mathbb{P}(|R(x, t)| > \epsilon) = 0$ *for any* $\epsilon > 0$.

Proof Observe that

$$
\begin{aligned}
t(F(x,t)-F(x)) &= \int_0^t [\mathbb{1}(f(X_s)\le x)-F(x)]\,ds \\
&= O(\tau_0)+\int_{T(0)}^{T(\ell(t))}[\mathbb{1}(f(X_s)\le x)-F(x)]\,ds+O(\tau_{\ell(t)+1}) \quad (8)
\end{aligned}
$$

where $O(x)$ denotes a value z such that $|z|\le cx$ for some constant $c>0$. We can then write (8) as

$$\sum_{i=1}^{\ell(t)} Z_i(x)+O(\tau_0+\tau_{\ell(t)+1}). \quad (9)$$

Now, $(Z_i(x) : i\ge 1)$ is a 1-dependent, identically distributed sequence of random variables, and

$$
\begin{aligned}
\operatorname{Var} Z_1(x) &= \mathbb{E}Z_1^2(x) \\
&= \mathbb{E}\left(\int_{T(0)}^{T(1)}[\mathbb{1}(f(X_s)\le x)-F(x)]\,ds\right)^2 \\
&\le \mathbb{E}\left(\int_{T(0)}^{T(1)}|\mathbb{1}(f(X_s)\le x)-F(x)|\,ds\right)^2 \\
&\le \mathbb{E}\tau_1^2. \quad (10)
\end{aligned}
$$

From (9) we see that

$$R(x,t)=t^{-1/2}O(\tau_0+\tau_{\ell(t)+1})+t^{-1/2}\sum_{i=1}^{\ell(t)}Z_i(x)-t^{-1/2}\sum_{i=1}^{\lfloor \lambda t\rfloor}Z_i(x). \quad (11)$$

The first term on the right-hand side of (11) does not depend on x, and furthermore, converges almost surely to 0 as $n\to\infty$; see, e.g., [36, p. 420]. So it suffices to study the second and third terms on the right-hand side of (11). We use a modification of a standard argument (see, e.g., [36, p. 420] or [34, p. 216] for the standard case) that accounts for the 1-dependence of the sequence $(Z_i(x) : i\ge 1)$ and our goal of uniformity in x. Let ϵ and δ be arbitrary positive quantities. Then

$$
\begin{aligned}
&\mathbb{P}\left(\left|\sum_{i=1}^{\ell(t)}Z_i(x)-\sum_{i=1}^{\lfloor \lambda t\rfloor}Z_i(x)\right|>\epsilon t^{1/2}\right) \\
&\le \mathbb{P}\left(\left|\sum_{i=1}^{\ell(t)}Z_i(x)-\sum_{i=1}^{\lfloor \lambda t\rfloor}Z_i(x)\right|>\epsilon t^{1/2};\,|\ell(t)-\lambda t|>\delta t\right) \\
&+\mathbb{P}\left(\left|\sum_{i=1}^{\ell(t)}Z_i(x)-\sum_{i=1}^{\lfloor \lambda t\rfloor}Z_i(x)\right|>\epsilon t^{1/2};\,|\ell(t)-\lambda t|\le\delta t\right)
\end{aligned}
$$

$$\le \mathbb{P}(|\ell(t) - \lambda t| > \delta t) + \mathbb{P}\left(\max_{k=1}^{\lceil \delta t \rceil}\left|\sum_{i=1}^{k} Z_i(x)\right| > \epsilon t^{1/2}\right)$$

$$\le \mathbb{P}(|\ell(t) - \lambda t| > \delta t) + \mathbb{P}\left(\max_{k=1}^{\lceil \delta t \rceil}\left|\sum_{i=1, i \text{ odd}}^{k} Z_i(x)\right| > \epsilon t^{1/2}/2\right)$$

$$+ \mathbb{P}\left(\max_{k=1}^{\lceil \delta t \rceil}\left|\sum_{i=1, i \text{ even}}^{k} Z_i(x)\right| > \epsilon t^{1/2}/2\right) \tag{12}$$

$$\le \mathbb{P}(|\ell(t) - \lambda t| > \delta t) + \frac{(\delta t + 1) \operatorname{Var} Z_i(x)}{\epsilon^2 t/4} \tag{13}$$

where (13) follows from Kolmogorov's maximum inequality; see, e.g., [34, Theorem 5.3.1]. We add only over odd cycles or even cycles, so the terms in the sums in (12) are independent, and there are a total of at most $\delta t + 1$ terms. Now choose $\delta = \epsilon^3$, so that the bound (13) becomes

$$\mathbb{P}(|\ell(t)-\lambda t| > \delta t)+4(\epsilon+\epsilon^{-2}t^{-1}) \operatorname{Var} Z_i(x) \le \mathbb{P}(|\ell(t)-\lambda t| > \delta t)+4(\epsilon+\epsilon^{-2}t^{-1})\mathbb{E}\tau_1^2.$$

This bound does not depend on x, and a standard renewal-theoretic result ensures that $\ell(t)/t \to \lambda$ as $t \to \infty$ almost surely and hence in probability. Since $\epsilon > 0$ was arbitrary, this completes the proof.

The representation given in Lemma 5 is sufficient to obtain a CLT for $F(x, \cdot)$ for any fixed x. In particular, using a CLT for 1-dependent sequences and assuming that $\mathbb{E}\tau_1^2 < \infty$ we see that $\sqrt{t}(F(x,t) - F(x)) \Rightarrow \sigma(x)N(0,1)$ as $t \to \infty$, where

$$\sigma^2(x) = \frac{\mathbb{E}Z_1^2(x) + 2\mathbb{E}Z_1(x)Z_2(x)}{\mathbb{E}\tau_1}.$$

To apply Theorem 1 we need $\sigma^2(\cdot)$ to be continuous in a neighborhood of q. To this end we have the following result.

Lemma 6 *Suppose that $\mathbb{E}\tau_1^2 < \infty$ and that F is continuous at q. Then $\sigma^2(\cdot)$ as defined above is continuous at q.*

Proof Since $F(\cdot) = \mathbb{P}_\pi(f(X_0) \le \cdot)$ is continuous at q, it follows that $\mathbb{P}_\pi(f(X_0) = q) = 0$. But then, for all $i \ge 1$,

$$0 = \mathbb{P}_\pi(f(X_0) = q) = \frac{\mathbb{E}\int_{T(i-1)}^{T(i)} \mathbb{1}(f(X_s) = q)\, ds}{\mathbb{E}\tau_1},$$

so that $\mathbb{E}\int_{T(i-1)}^{T(i)} \mathbb{1}(f(X_s) = q)\, ds = 0$. It immediately follows that $Z_i(x) \to Z_i(q)$ as $x \to q$ almost surely for any $i \ge 1$. Hence

$$Z_1^2(x) + 2Z_1(x)Z_2(x) \to Z_1^2(q) + 2Z_1(q)Z_2(q)$$

as $x \to q$ almost surely. Furthermore, $|Z_1^2(x) + 2Z_1(x)Z_2(x)| \le \tau_1^2 + 2\tau_1\tau_2$ for any x and $\mathbb{E}(\tau_1^2 + 2\tau_1\tau_2) \le 3\mathbb{E}\tau_1^2 < \infty$. The dominated convergence theorem then gives

$$\sigma^2(x) = \frac{\mathbb{E}(Z_1^2(x) + 2Z_1(x)Z_2(x))}{\mathbb{E}\tau_1} \to \frac{\mathbb{E}(Z_1^2(q) + 2Z_1(q)Z_2(q))}{\mathbb{E}\tau_1} = \sigma^2(q)$$

as $x \to q$ as desired. □

We are now in a position to state and prove the main result of the paper.

Theorem 3 *Suppose that $\mathbb{E}\tau_1^2 < \infty$, F is differentiable at q with $F'(q) > 0$ and $\sigma^2(q) > 0$. Then, as $t \to \infty$,*

$$\frac{\sqrt{t}(Q_t - q)}{\sigma(q)/F'(q)} \Rightarrow N(0,1).$$

Proof Lemma 6 together with the assumption that $\sigma^2(q) > 0$ ensures that $\sigma^2(\cdot)$ is bounded away from 0 in a neighborhood N of q. Furthermore, the random variables $(Z_1(x) : x \in (-\infty,\infty))$ are uniformly integrable, as can be seen from (10). These observations, together with the representation given in Lemma 5 and Theorem 2 ensure that the uniform CLT holds, i.e.,

$$\sup_{x\in N}\sup_{y}\left|\mathbb{P}\left(\frac{t^{1/2}[F(x,t) - F(x)]}{\sigma(x)} \le y\right) - \Phi(y)\right| \to 0$$

as $t \to \infty$. The result now follows from Theorem 1. □

Remark 1 The following example indicates that we cannot relax the assumption that $\mathbb{E}\tau_1^2 < \infty$. Let $\tau_1, \tau_2, \dots$ be i.i.d. nonnegative random variables where $0 < \mathbb{E}\tau_1 < \infty$ and $\mathbb{E}\tau_1^2 = \infty$. For $n \ge 1$ let $T(n) = \tau_1 + \cdots + \tau_n$ and let $T(0) = 0$. For $t \ge 0$ let $\ell(t) = \sup\{n : T(n) \le t\}$ be the number of completed cycles by time t. Let $X_t = t - T(\ell(t))$, so that $X = (X_t : t \ge 0)$ is the age process associated with the renewal process $(\ell(t) : t \ge 0)$. Take f to be the identity function, and note that

$$F(x) = \frac{\mathbb{E}\int_0^{\tau_1} I(X_s \le x)}{\mathbb{E}\tau_1}\, ds = \frac{\mathbb{E}[x \wedge \tau_1]}{\mathbb{E}\tau_1},$$

where $a \wedge b = \min(a,b)$. To simplify things, we assume that $\tau_1 > 1$ a.s., so that for $x \in [0,1]$, $F(x) = x/\mathbb{E}\tau_1$. Choose $p \in (0, 1/\mathbb{E}\tau_1)$ so that $q = F^{-1}(p) = p\mathbb{E}\tau_1 < 1$. Then for $y \in \mathbb{R}$ and t sufficiently large that $q + yt^{-1/2} < 1$,

$$\begin{aligned}
\mathbb{P}(t^{1/2}(Q_t - q) \le y) &= \mathbb{P}(Q_t \le q + yt^{-1/2}) \\
&= \mathbb{P}(p \le F(q + yt^{-1/2}, t)) \qquad (14) \\
&= \mathbb{P}(pt \le [q + yt^{-1/2}]\ell(t) + R_t), \qquad (15)
\end{aligned}$$

where $R_t = (q + yt^{-1/2}) \wedge X_t$. Equality (14) follows from [9, Lemma 1.1.4], and (15) since $tF(x,t) = x\ell(t) + x \wedge X_t$ for $x < 1$. Now, since $q = pE\tau_1$,

$$
\begin{aligned}
\mathbb{P}(t^{1/2}(Q_t - q) \le y) &= \mathbb{P}\left(\frac{pt^{3/2}}{\ell(t)} - p\mathbb{E}\tau_1\sqrt{t} - \frac{R_t\sqrt{t}}{\ell(t)} \le y\right) \\
&= \mathbb{P}\left(\frac{pt}{\ell(t)}\frac{t - \ell(t)\mathbb{E}\tau_1}{\sqrt{t}} - \frac{R_t\sqrt{t}}{\ell(t)} \le y\right) \\
&= \mathbb{P}\left(\frac{pt}{\ell(t)}\frac{1}{\sqrt{t}}\sum_{i=1}^{\ell(t)}(\tau_i - \mathbb{E}\tau_1) + \frac{pt}{\ell(t)}\frac{X_t}{\sqrt{t}} - \frac{R_t\sqrt{t}}{\ell(t)} \le y\right). \quad (16)
\end{aligned}
$$

Now, $t/\ell(t) \to \mathbb{E}\tau_1$ as $t \to \infty$ a.s. Furthermore, the set of random variables $(X_t : t \ge 0)$ is tight (see [38, Proposition 1] and [39, Proposition 9]), and so

$$
\frac{pt}{\ell(t)}\frac{X_t}{\sqrt{t}} - \frac{R_t\sqrt{t}}{\ell(t)} \Rightarrow 0
$$

as $t \to \infty$. Thus, from the converging together lemma (e.g., [34, Theorem 4.4.6 and corollary]) and (16), $t^{1/2}(Q_t - q)$ converges in distribution to a normal random variable if and only if

$$
t^{-1/2}\sum_{i=1}^{\ell(t)}(\tau_i - \mathbb{E}\tau_1) \quad (17)
$$

converges in distribution to a normal random variable. From [38], (17) converges in distribution to a normal random variable if and only if $E\tau_1^2 < \infty$, so the desired CLT does not hold. The other conditions of Theorem 3 are easily seen to hold for this example. Thus, the condition $\mathbb{E}\tau_1^2 < \infty$ is, in a certain sense, sharp.

5 The Validity of Non-Overlapping Batch-Means Estimation

Theorem 3 establishes conditions under which the quantile estimator Q_t is asymptotically normally distributed. One would like to leverage this result to provide confidence intervals for q. Constructing such confidence intervals by directly estimating the variance constant $\sigma(q)/F'(q)$ is difficult, because both terms in this expression are challenging to estimate. Indeed, regenerative estimators of $\sigma(q)$ require the ability to identify the cycle boundaries $(T(i) : i \ge 0)$, and this is, at best, extremely difficult in general discrete-event simulations [26]. Furthermore, the density term $F'(q)$ requires some form of density estimator, and such estimators typically converge at a rate that is slower than the canonical $t^{-1/2}$ rate [40].

An alternative is the method of non-overlapping batch quantiles; see, e.g., [24, 41]. In this method, the sample path $(X_s : 0 \le s \le t)$ is divided into b batches, with the ith batch given by $(X_s : (i-1)t/b < s \le it/b)$, $i = 1, 2, \ldots, b$. Let $F_i(\cdot, t)$ denote the empirical CDF based on the ith batch, so that

$$
F_i(x, t) = \frac{b}{t}\int_{(i-1)t/b}^{it/b} \mathbb{1}(f(X_s) \le x)\, ds,
$$

for all $x \in \mathbb{R}$ and all $i = 1, 2, \ldots, b$. Let $Q_i(t) = F_i^{-1}(p, t)$ be the estimator of the p quantile based on the ith batch. Theorem 3 basically establishes that, for each i, $Q_i(t)$ is approximately normal. If, in addition, $Q_i(t)$ is asymptotically independent of $Q_j(t)$ for $i \neq j$, then standard confidence interval theory ensures that an approximate $100(1-\alpha)\%$ confidence interval is given by

$$\bar{Q}(t) \pm t_{\alpha,b-1} \frac{s_b}{\sqrt{b}}, \tag{18}$$

where $\bar{Q}(t) = b^{-1} \sum_{i=1}^{b} Q_i(t)$ is the average of the batch quantiles, s_b^2 is the sample variance of $Q_1(t), Q_2(t), \ldots, Q_b(t)$, and $t_{\alpha,b-1}$ is the $1-\alpha/2$ quantile of a t distribution with $b-1$ degrees of freedom.

This procedure is rigorously justified through a joint CLT for $(Q_i(t) : i = 1, 2, \ldots, b)$, which we provide in Theorem 4 below.

Quantile estimators are known to exhibit bias, with the bias being on the order of the inverse of the runlength [8]. Accordingly, the estimator $\bar{Q}(t)$ has a bias that can be expected to be approximately b times as large as that of the estimator Q_t of the quantile based on the entire length-t sample path. The coverage of the confidence interval (18) can be expected to be improved if the average of the batch quantiles $\bar{Q}(t)$ is replaced by Q_t. The asymptotic validity of confidence intervals constructed in this way is assured through the joint CLT, Theorem 4, and a result that establishes that Q_t and $\bar{Q}(t)$ are "close" in the sense that $t^{1/2}(Q_t - \bar{Q}(t)) \Rightarrow 0$ as $t \to \infty$. This latter result is a direct consequence of Proposition 2 below, which gives a so-called Bahadur-Ghosh representation of quantile estimators in the Markov chain setting.

Our first result in this section provides a representation for the batch empirical CDFs along the lines of Lemma 5. The proof follows almost exactly the same lines as that of Lemma 5, using a vector version of Lemma 1, and so is omitted.

Lemma 7 *If $E\tau^2 < \infty$ then for $x \in \mathbb{R}^b$,*

$$\sqrt{\frac{t}{b}} \begin{pmatrix} F_1(x_1, t) - F(x_1) \\ F_2(x_2, t) - F(x_2) \\ \vdots \\ F_b(x_b, t) - F(x_b) \end{pmatrix} = \frac{1}{\sqrt{\frac{t}{b}}} \begin{pmatrix} \sum_{j=1}^{l} Z_j(x_1) \\ \sum_{j=l+1}^{2l} Z_j(x_2) \\ \vdots \\ \sum_{j=(b-1)l+1}^{bl} Z_j(x_2) \end{pmatrix} + R(x, t),$$

where $l = \lfloor \lambda t/b \rfloor$ and the vector-valued error term $R(x,t)$ satisfies, for any $\epsilon > 0$,

$$\lim_{t\to\infty} \sup_x \mathbb{P}(\|R(x,t)\| > \epsilon) = 0.$$

The next result is a vector version of the uniform CLT, Theorem 2. The proof is very similar to that of Theorem 2 and so we only provide a sketch of the proof.

Lemma 8 *Let $(q_1, q_2, \ldots, q_b) \in \mathbb{R}^b$ and let N_i be an open neighbourhood of q_i for each $i = 1, 2, \ldots, b$. Let $\tilde{N} = N_1 \times N_2 \times \cdots \times N_b$. If $\mathbb{E}\tau^2 < \infty$ and $\eta(x) = \mathbb{E}Z_1^2(x) + 2\mathbb{E}[Z_1(x)Z_2(x)]$ is bounded away from 0 for $x \in \cup_{i=1}^{b} N_i$, then*

$$\sup_{x\in\tilde{N}}\sup_{y\in\mathbb{R}^b}\left|P\left(\frac{\sum_{j=(i-1)l+1}^{il} Z_j(x_i)}{\eta(x_i)\sqrt{l}}\le y_i, i=1,2,\dots,b\right)-\prod_{i=1}^{b}\Phi(y_i)\right|\to 0$$

as $t\to\infty$, where $l=l(t)=\lfloor \lambda t/b\rfloor$.

Proof (Sketch) Within each batch, apply the "big block little block" argument to obtain asymptotic (marginal) normality, as in the proof of Theorem 2 for each batch. To obtain the desired asymptotic independence, drop the last cycle in each batch, i.e., write the ith batch sum as

$$\sqrt{\frac{l-1}{l}}\frac{\sum_{j=(i-1)l+1}^{il-1} Z_j(x_i)}{\eta(x_i)\sqrt{l-1}}+\frac{Z_{il}(x_i)}{\eta(x_i)\sqrt{l}}$$

and now apply the matrix version of Lemma 1. □

Theorem 4 *Suppose that $\mathbb{E}\tau^2<\infty$. Suppose further that $F(\cdot)$ is differentiable at q, $F'(q)>0$, $\sigma^2(q)>0$ and $\sigma^2(\cdot)$ is continuous at q. For $y\in\mathbb{R}^b$ let*

$$G(y,t)=\mathbb{P}\left(\frac{\sqrt{t/b}(Q_i(t)-q)}{\sigma(q)/F'(q)}\le y_i, i=1,\dots,b\right).$$

Then $G(y,t)\to\prod_{i=1}^{b}\Phi(y_i)$ as $t\to\infty$.

Proof The proof is very similar to that of Theorem 1. Define

$$q_{t,i}=q+\frac{\sigma(q)y_i}{F'(q)\sqrt{t/b}},$$

for $i=1,2,\dots,b$. Then

$$\begin{aligned}G(y,t)&=\mathbb{P}(Q_i(t)\le q_{t,i}, i=1,2,\dots,b)\\&=\mathbb{P}(p\le F_i(q_{t,i},t), i=1,2,\dots,b)\\&=\mathbb{P}(U_i(q_{t,i},t)\ge -y_{t,i}, i=1,2,\dots,b),\end{aligned}$$

where

$$U_i(z,t)=\sqrt{\frac{t}{b}}\frac{F_i(z,t)-F(z)}{\sigma(z)}\quad\text{and}\quad y_{t,i}=\sqrt{\frac{t}{b}}\frac{F(q_{t,i})-p}{\sigma(q_{t,i})}.$$

Defining $\bar{\Phi}(a)=1-\Phi(a)$, we have that $\Phi(a)=\bar{\Phi}(-a)$, and so

$$\begin{aligned}G(y,t)-\prod_{i=1}^{b}\Phi(y_i)&=\mathbb{P}(U_i(q_{t,i},t)\ge -y_{t,i}, i=1,2,\dots,b)-\prod_{i=1}^{b}\bar{\Phi}(-y_{t,i})\\&\quad+\prod_{i=1}^{b}\Phi(y_{t,i})-\prod_{i=1}^{b}\Phi(y_i).\end{aligned}$$

The first line of the right-hand side converges to 0 by the uniform law of large numbers. The second line converges to 0 because $y_{t,i} \to y_i$ as $t \to \infty$. □

This is the desired multivariate CLT. Thus batch means using the average of the batch quantiles is asymptotically valid.

Recall that if $t^{1/2}(Q_t - \bar{Q}(t)) \Rightarrow 0$, then we can replace the average of the batch quantiles, $\bar{Q}(t)$, in the joint CLT above with Q_t, the quantile estimator based on the entire sample path. We now establish the Bahadur-Ghosh representation

$$Q_t = q - \frac{F(q,t) - F(q)}{F'(q)} + R(t), \tag{19}$$

where $t^{1/2}R(t) \Rightarrow 0$ as $t \to \infty$. Applying this representation to each batch, $i = 1, 2, \ldots, b$ yields

$$Q_i(t) = q - \frac{F_i(q,t) - F(q)}{F'(q)} + R_i(t),$$

and averaging gives

$$\begin{aligned} \bar{Q}(t) &= q - \frac{F(q,t) - F(q)}{F'(q)} + \frac{1}{b}\sum_{i=1}^{b} R_i(t) \\ &= Q_t - R(t) + \frac{1}{b}\sum_{i=1}^{b} R_i(t) \end{aligned}$$

which gives the desired result. It therefore remains to prove the Bahadur-Ghosh representation. We first state a lemma due to [42], and then prove the representation.

Lemma 9 ([42])

Let $(\nu_t : t \geq 0)$ and $(\xi_t : t \geq 0)$ be two stochastic processes satisfying the following conditions.

1. *The process $(\xi_t : t \geq 0)$ is tight, i.e., for all $\delta > 0$ there exists $M > 0$ such that $\mathbb{P}(|\xi_t| > M) \leq \delta$.*
2. *For all $y \in \mathbb{R}$ and $h > 0$,*

$$\lim_{t\to\infty} \mathbb{P}(\nu_t \leq y, \xi_t \geq y + h) = \lim_{t\to\infty} \mathbb{P}(\nu_t \geq y + h, \xi_t \leq y) = 0.$$

Then $\nu_t - \xi_t \Rightarrow 0$ as $t \to \infty$.

Proposition 2 *Suppose that F is differentiable at q with $F'(q) > 0$ and $\mathbb{E}\tau^2 < \infty$. Then the Bahadur-Ghosh representation (19) is valid.*

Proof The essential elements of our proof are similar to those in [42] for the i.i.d. case. Let $y \in \mathbb{R}$ be arbitrary. As in the proof of Theorem 1, the events $\{t^{1/2}(Q_t - q) \leq y\}$ and

$$\{-t^{1/2}(F(q + t^{-1/2}y, t) - F(q + t^{-1/2}y)) \leq t^{1/2}(F(q + t^{-1/2}y) - p)\}$$

are identical.

The differentiability of F at q ensures that $t^{1/2}(F(q+t^{-1/2}y)-p) = F'(q)y+o(1)$ as $t \to \infty$. Furthermore,

$$t^{1/2}(F(q+t^{-1/2}y,t) - F(q+t^{-1/2}y)) = t^{1/2}(F(q,t) - F(q)) + V(t),$$

where the remainder term $V(t)$ is given by

$$t^{1/2}(F(q+t^{-1/2}y,t) - F(q+t^{-1/2}y) - F(q,t) + F(q)).$$

The proof will be complete if we show that $V(t) \Rightarrow 0$ as $t \to \infty$. (To see why, take $\nu_t = t^{1/2}(Q_t - q)$ and $\xi_t = t^{1/2}(F(q,t) - F(q))/F'(q) + V(t)/F'(q)$ in Lemma 9 above.)

From Lemma 5, we can write

$$V(t) = t^{-1/2} \sum_{i=1}^{\lfloor \lambda t \rfloor} W_i(t) + R(t),$$

where the (mean-zero) cycle-term $W_i(t) = Z_i(q+t^{-1/2}y) - Z_i(q)$ and $R(t) \Rightarrow 0$ as $t \to \infty$. Chebyshev's inequality then gives that for arbitrary $\epsilon > 0$,

$$\mathbb{P}\left(\left|t^{-1/2}\sum_{i=1}^{\lfloor \lambda t \rfloor} W_i(t)\right| > \epsilon\right) \le \frac{1}{\epsilon^2 t}\left(\lfloor \lambda t \rfloor \mathbb{E}W_1^2(t) + 2(\lfloor \lambda t \rfloor - 1)\mathbb{E}[W_1(t)W_2(t)]\right). \tag{20}$$

Now, exactly as in Lemma 6, for any fixed i, $W_i(t) \to 0$ as $t \to \infty$ a.s., and $|W_i(t)| \le \tau_i$, and so dominated convergence ensures that the right-hand side of (20) converges to 0 as $t \to \infty$, thereby completing the proof. □

Remark 2 The Bahadur-Ghosh representation immediately provides a weak law of large numbers for the quantile estimator Q_t as well as the means to prove a CLT for Q_t based on the empirical CDF. It is natural to ask why we did not use this representation earlier in our development. An inspection of the proof of Proposition 2 shows that the essential elements of the proof are the same as those we developed in earlier sections, so it does not appear that there is anything to gain from doing so.

6 Sufficient Conditions

The assumptions of Theorems 3 and 4 are difficult to verify as stated. In this section we provide sufficient conditions for some of those assumptions that are often more easily verified in applications. Where possible, we try to give a unified treatment of both discrete-time and continuous-time Harris processes. Let $(X_t : t \ge 0)$ be a Markov process in discrete or continuous time as defined in Section 4. (Recall that in continuous time we assume that the process is non-explosive, strong Markov, and has sample paths that are right continuous with left limits.) We begin with the condition that the regenerative cycle lengths have finite second moment, i.e., that $E\tau_1^2 < \infty$, which can be verified through the use of drift criteria.

Definition 3 Let $X = (X_t : t \geq 0)$ be a Markov process on a complete, separable metric space in discrete or continuous time. Let $u : S \to \mathbb{R}$ and suppose that there exists $h : S \to \mathbb{R}$ such that $M = (M_t : t \geq 0)$ is a $\mathbb{P}_x$-local martingale for all $x \in S$, where

$$M_t = u(X_t) - u(X_0) - \int_0^t h(X_s)\,ds, \tag{21}$$

and t is restricted to discrete or continuous time as appropriate. We then say that u is contained in the domain $\mathcal{D}(\mathcal{A})$ of the generator $\mathcal{A}$ of X, and $\mathcal{A}u = h$.

Suppose that in addition to A2 for discrete chains, or A2 with $m = 1$ for the embedded chain for continuous-time processes, we also have the following, where the set C is as in A2 for the sampled chain.

A4 There exists $g_1 : S \to [0, \infty)$ such that for all $x \in S$, and some $b_1 > 0$,

$$\mathcal{A}g_1(x) \leq -1 + b_1 \mathbb{1}(x \in C).$$

A5 There exists $g_2 : S \to [0, \infty)$ such that for all $x \in S$ and some $b_2 > 0$,

$$\mathcal{A}g_2(x) \leq -g_1(x) + b_2 \mathbb{1}(x \in C).$$

Assumption A4 implies that X is positive-Harris recurrent; see [36, Theorem 14.0.1] for the discrete case and [43] for the continuous case. Assumptions A4 and A5 imply a finite second moment of the regeneration times, i.e., that $\mathbb{E}_\varphi \tau_1^2 < \infty$. We prove the continuous-time result; the discrete-time result follows essentially the same proof with a modest modification since m in A2 cannot be assumed to equal 1.

Lemma 10 *Suppose that A4 holds for the continuous-time process X. Let $\tilde{X}$ be the sampled chain. Then, for all x,*

$$\mathbb{E}_x g_1(\tilde{X}_1) - g_1(x) \leq -1 + b_1 \mathbb{P}_x(\tilde{X}_1 \in C). \tag{22}$$

Proof Since g_1 lies in the domain of the generator, (21) with $u = g_1$ is a $\mathbb{P}_x$ local martingale for all $x \in S$. It follows from the observations on p. 311 of [44] that

$$e^{-t} g_1(X_t) - g_1(X_0) + \int_0^t e^{-s}(g_1(X_s) - \mathcal{A}g_1(X_s))\,ds$$

is also a $\mathbb{P}_x$ local martingale for all $x \in S$. Thus, since $g_1 \geq 0$, for a sequence of stopping times $O_n \to \infty$ as $n \to \infty$ $\mathbb{P}_x$ a.s.,

$$\mathbb{E}_x[e^{-t\wedge O_n} g_1(X_{t\wedge O_n})] + \mathbb{E}_x \int_0^{t\wedge O_n} e^{-s}(g_1(X_s) - \mathcal{A}g_1(X_s))\,ds = g_1(x). \tag{23}$$

Now, A4 implies that $g_1(x) + 1 \leq g_1(x) - \mathcal{A}g_1(x) + b_1 I(x \in C)$ for all $x \in S$. Hence

$$
\begin{aligned}
\mathbb{E}_x \int_0^{t\wedge O_n} e^{-s}(g_1(X_s)+1)\,ds &\le \mathbb{E}_x \int_0^{t\wedge O_n} e^{-s}(g_1(X_s)-\mathcal{A}g_1(X_s))\,ds \\
&\quad + b_1\mathbb{E}_x \int_0^{t\wedge O_n} e^{-s} I(X_s\in C)\,ds \\
&\le g_1(x) + b_1\mathbb{E}_x \int_0^{\infty} e^{-s} I(X_s\in C)\,ds,
\end{aligned}
$$

where, in the second inequality we used (23). Taking $n\to\infty$ and then $t\to\infty$, monotone convergence gives

$$
\mathbb{E}_x \int_0^{\infty} e^{-s} g_1(X_s)\,ds + 1 \le g_1(x) + b_1\mathbb{E}_x \int_0^{\infty} e^{-s} I(X_s\in C)\,ds,
$$

i.e., that $\mathbb{E}_x g_1(\tilde{X}_1) - g_1(x) \le -1 + b_1\mathbb{P}_x(\tilde{X}_1\in C)$. □

Proposition 3 *Suppose that A4 and A5 hold. Then* $\mathbb{E}_\varphi \tau_1^2 < \infty$.

Proof Recall that we have enlarged the path space of the Markov process X to include an independent unit-rate Poisson process $(N(t) : t\ge 0)$ with event times $(\Lambda(n) : n\ge 0)$ with $\Lambda(0)=0$ and an i.i.d. sequence of Bernoulli random variables $(I_n : n\ge 1)$ with $P(I_1=1)=\gamma$.

Let $\mathbb{E}_\varphi$ and $\mathbb{P}_\varphi$ denote the expectation and probability on the enlarged probability space when the chain X has initial distribution φ, so that a regeneration occurs at time 0. For convenience, write τ for τ_1. For $n\ge 0$, let $M(n)=\sum_{j=0}^{n} I(\tilde{X}_j\in C)$ be the number of attempted regenerations by time n. Define the discrete-time stopping time $\tilde{\tau}=\inf\{n\ge 0 : I_{M(n)}=1\}$. Under $\mathbb{P}_\varphi$, the regeneration time $\tau=\Lambda(\tilde{\tau}+1)$.

From (22) and the comparison theorem [36, Theorem 14.2.2],

$$
\mathbb{E}_x\tilde{\tau} \le g_1(x) + b_1\mathbb{E}_x \sum_{j=0}^{\tilde{\tau}-1} h(\tilde{X}_j),
$$

where $h(x)=\mathbb{P}_x(\tilde{X}_1\in C)$. Since $I(\tilde{\tau}>j)$ is measurable with respect to $\mathcal{G}_j=\sigma(\tilde{X}_0,\ldots,\tilde{X}_j,I_1,\ldots,I_{M(j)})$, it follows that

$$
\mathbb{E}_x \sum_{j=0}^{\tilde{\tau}-1} h(\tilde{X}_j) = \sum_{j=0}^{\infty} \mathbb{P}_x(\tilde{\tau}>j,\tilde{X}_{j+1}\in C) = \mathbb{E}_x \sum_{j=1}^{\tilde{\tau}} I(\tilde{X}_j\in C).
$$

Now, each time j that $\tilde{X}_j\in C$, we regenerate with probability γ, so that $\sum_{j=1}^{\tilde{\tau}} I(\tilde{X}_j\in C)$ is geometrically distributed with success probability γ and thus has mean γ^{-1}. We conclude that $\mathbb{E}_x\tilde{\tau}\le g_1(x)+b_1/\gamma$.

With that result in hand,

$$\begin{aligned}
\mathbb{E}_\varphi \tilde{\tau}^2 &\leq 2\mathbb{E}_\varphi \sum_{j=0}^{\tilde{\tau}-1} (\tilde{\tau} - j) \\
&= 2\mathbb{E}_\varphi \sum_{j=0}^{\infty} \mathbb{E}_\varphi [(\tilde{\tau} - j) I(\tilde{\tau} > j) | \mathcal{G}_j] \\
&\leq 2\mathbb{E}_\varphi \sum_{j=0}^{\tilde{\tau}-1} (g_1(\tilde{X}_j) + b_1/\gamma) \\
&= 2\mathbb{E}_\varphi \sum_{j=0}^{\tilde{\tau}-1} g_1(\tilde{X}_j) + 2b_1 \mathbb{E}_\varphi \tilde{\tau}/\gamma \\
&= 2(\mathbb{E}_\varphi \tilde{\tau})\,(\mathbb{E}_\pi g_1(X_0)) + 2b_1(\mathbb{E}_\varphi \tilde{\tau})/\gamma < \infty,
\end{aligned}$$

where $\mathbb{E}_\pi g_1(X_0)$ is finite by virtue of A5 and [43, Theorem 4.2]. Since $\tau = \Lambda(\tilde{\tau} + 1)$ under $\mathbb{P}_\varphi$, Wald's second moment identity then implies that $\mathbb{E}_\varphi \tau^2 < \infty$. □

The hypotheses A4 and A5 simplify when the chain X is V-uniformly ergodic as is assumed in [24]. In fact, A4 and A5 are implied by A6 below; see, e.g., [36, Lemma 17.5.1] and [45].

A6 For the set C defined in A2 there exist constants $b, \beta > 0$, and a function $V : S \to [1, \infty)$ such that for all $x \in S$,

$$\mathcal{A}V(x) \leq -\beta V(x) + bI(x \in C).$$

For the other hypotheses of Theorem 3 it is not clear exactly what form "easily verifiable" conditions should take. Indeed, it appears that one may need to tailor the conditions to a given application. It is difficult to imagine a practical application where the condition $\sigma^2(q) > 0$ would be violated, so we content ourselves with an example sufficient condition for the hypothesis that F is differentiable at q with $F'(q) > 0$. Recall that A2 and A4 imply that the chain X is positive Harris recurrent, and therefore possesses a stationary distribution, so that $F(y) = \mathbb{P}_\pi(f(X_0) \leq y)$ is well-defined. In what follows we assume that X is positive Harris recurrent.

Proposition 4 *Suppose there exists a $t > 0$ such that for all y in an open neighbourhood N of q and all $x \in S$,*

$$\mathbb{P}(f(X_t) \in dy | X_0 = x) = p(x, y)dy.$$

Suppose further that for each fixed $x \in S$, $p(x, \cdot)$ is Lipschitz continuous in $y \in N$ with Lipschitz constant $L(x)$, where $L(\cdot)$ is π-integrable. Then F is differentiable in N. If, in addition, $p(x, q) > 0$ for x in some set of positive π measure, then $F'(q) > 0$.

Proof The proof is very similar to that of Proposition 2 in [46]. Let $B = (a, b] \subseteq N$. Then

$$\begin{aligned}F(b) - F(a) &= \mathbb{P}_\pi(f(X_t) \in B)\\ &= \int_S \pi(dx)\mathbb{P}(f(X_t) \in B|X_0 = x)\\ &= \int_S \int_B \pi(dx)p(x, y)dy\\ &= \int_B \int_S \pi(dx)p(x, y)dy.\end{aligned}$$

It follows immediately that F has a density ψ in N, where

$$\psi(y) = \int_S \pi(dx)p(x, y) \tag{24}$$

at $y \in N$. Now, for h such that both y and $y + h \in N$,

$$|\psi(y + h) - \psi(y)| = \left|\int_S \pi(dx)(p(x, y + h) - p(x, y))\right| \leq h \int_S \pi(dx)L(x). \tag{25}$$

Since L is π integrable, it follows that ψ is Lipschitz continuous in N. Since F has a continuous density in N, we may conclude that it is differentiable (in fact, continuously differentiable) in N with derivative ψ.

Finally, observe from (24) that the condition that $p(x, q) > 0$ for all x in a set of positive π measure implies that $\psi(q) = F'(q)$ is positive at q. □

Proposition 4 basically requires that the t-step probabilities $\mathbb{P}(f(X_t) \in dy|X_0 = x)$ have a density with respect to Lebesgue measure for all x. Typically this condition will be easiest to verify for Harris processes in discrete time in the case where $t = 1$.

Example: Consider the problem of computing quantiles of the steady-state waiting time distribution in the GI/G/1 queue. It is well-known that the sequence $X = (X_n : n \geq 0)$ of customer waiting times in the FIFO single-server queue is a Markov chain on state space $S = [0, \infty)$. In particular, X satisfies the Lindley recursion [35, p. 23] $X_{n+1} = [X_n + Y_{n+1}]^+$, where $[x]^+ = \max(x, 0)$, $Y = (Y_n : n \geq 1)$ is an i.i.d. sequence with $Y_{n+1} = V_n - U_{n+1}$, V_n is the service time of the nth customer, and U_{n+1} is the interarrival time between the nth and $(n + 1)$st customer. Take $f(x) = x$, so that we are interested in computing the quantiles of the steady-state waiting time distribution. We now verify the key conditions of Theorem 3.

As in [35, p. 23], it is straightforward to show that if $\mathbb{E}Y_1^2 < \infty$ and $\mu = \mathbb{E}Y_1 < 0$, then A4 and A5 are satisfied for the Markov chain X with $g_1(x) = 2x/|\mu|$ and $g_2(x) = 2x^2/\mu^2$. Now, for $y > 0$, we have that $P(x, dy) = \mathbb{P}(Y_1 \in d(y - x))$. So if Y_1 has a Lipschitz continuous density with respect to Lebesgue measure and $q > 0$, then Proposition 4 implies that the distribution function F is differentiable in a neighbourhood of q. It remains to establish that $F'(q) > 0$.

First, $\pi(\{0\}) = 1 - \mathbb{E}V_1/\mathbb{E}U_1 > 0$, since $\mathbb{E}Y_1 < 0$. Furthermore, since Y_1 has a continuous density and negative mean, $\mathbb{P}(Y_1 > 0) > 0$ then implies that for each $0 \leq a < b < \infty$, there exists an $m = m(a, b)$ such that $P^m(0, (a, b)) > 0$. Therefore, $\pi((a, b)) \geq \pi(\{0\})P^m(0, (a, b)) > 0$. Proposition 4 then implies that $F'(q) > 0$.

Acknowledgements We have benefited enormously from our association with Pierre L'Ecuyer over many years. We are grateful for Pierre's scholarship, leadership and friendship.

This work was partially supported by National Science Foundation grant CMMI-2035086.

References

1. Donald L. Iglehart. Simulating stable stochastic systems, VI: quantile estimation. *Journal of the Association for Computing Machinery*, 23:347–360, 1976.
2. Andrew F. Seila. A batching approach to quantile estimation in regenerative simulations. *Management Science*, 28:573–581, 1982.
3. Charles R Doss, James M Flegal, Galin L Jones, and Ronald C Neath. Markov chain Monte Carlo estimation of quantiles. *Electronic Journal of Statistics*, 8(2):2448–2478, 2014.
4. P. W. Glynn. Some topics in regenerative steady-state simulation. *Acta Applicandae Mathematicae*, 34:225–236, 1994.
5. W. R. Gilks, S. Richardson, and D. J. Spiegelhalter, editors. *Markov Chain Monte Carlo in Practice*. Chapman & Hall, London, 1996.
6. George S. Fishman. *Monte Carlo: Concepts, Algorithms and Applications*. Springer Series in Operations Research. Springer, New York, 1996.
7. S. Asmussen and P. W. Glynn. *Stochastic Simulation: Algorithms and Analysis*, volume 57 of *Stochastic Modeling and Applied Probability*. Springer, New York, 2007.
8. Athanassios N. Avramidis and James R. Wilson. Correlation-induction techniques for estimating quantiles in simulation experiments. *Operations Research*, 46:574–592, 1998.
9. Robert J. Serfling. *Approximation Theorems of Mathematical Statistics*. Wiley, New York, 1980.
10. X. Jin, M. C. Fu, and X. Xiong. Probabilistic error bounds for simulation quantile estimators. *Management Science*, 49:230–246, 2003.
11. J. C. Hsu and B. L. Nelson. Control variates for quantile estimation. *Management Science*, 36:835–851, 1990.
12. Timothy C. Hesterberg and Barry L. Nelson. Control variates for probability and quantile estimation. *Management Science*, 44(9):1295–1312, 1998.
13. P. W. Glynn. Importance sampling for Monte Carlo estimation of quantiles. *Proceedings of the Second International Workshop on Mathematical Methods in Stochastic Simulation and Experimental Design*, pages 180–185, 1996.
14. Paul Glasserman, Philip Heidelberger, and Perwez Shahabuddin. Variance reduction techniques for estimating value-at-risk. *Management Science*, 46:1349–1364, 2000.
15. Marvin K. Nakayama. Asymptotically valid confidence intervals for quantiles and values-at-risk when applying Latin hypercube sampling. *International Journal on Advances in Systems and Measurements*, 4:86–94, 2011.
16. L. Sun and L. J. Hong. Asymptotic representations for importance-sampling estimators of value-at-risk and conditional value-at-risk. *Operations Research Letters*, 38:246–251, 2010.
17. Fang Chu and Marvin K. Nakayama. Confidence intervals for quantiles when applying variance-reduction techniques. *ACM Transactions on Modeling and Computer Simulation*, 22(2):Article 10, 2012.
18. Marvin K. Nakayama. Using sectioning to construct confidence intervals for quantiles when applying importance sampling. In C. Laroque, J. Himmelspach, R. Pasupathy, O. Rose, and A. M. Uhrmacher, editors, *Proceedings of the 2012 Winter Simulation Conference*. IEEE, 2012.
19. Christos Alexopoulos, David Goldsman, Anup C. Mokashi, Kai-Wen Tien, and James R. Wilson. Sequest: A sequential procedure for estimating quantiles in steady-state simulations. *Operations Research*, 67(4):1162–1183, 2019.
20. P. Heidelberger and P. A. W. Lewis. Quantile estimation in dependent sequences. *Operations Research*, 32:185–209, 1984.

21. E. Jack Chen and W. David Kelton. Quantile and tolerance-interval estimation in simulation. *European Journal of Operational Research*, 168:520–540, 2006.
22. J. E. Bekki, G. Mackulak, J. W. Fowler, and B. L. Nelson. Indirect cycle time quantile estimation using the Cornish-Fisher expansion. *IIE Transactions*, 42:31–44, 2009.
23. Demet C. Wood and Bruce W. Schmeiser. Overlapping batch quantiles. In C. Alexopoulos, K. Kang, W. R. Lilegdon, and D. Goldsman, editors, *Proceedings of the 1995 Winter Simulation Conference*, pages 303–308, Piscataway, New Jersey, 1995. IEEE.
24. David F. Muñoz. On the validity of the batch quantile method for Markov chains. *Operations Research Letters*, 38:223–226, 2010.
25. David F. Muñoz and Adán Ramírez-López. A note on bias and mean squared error in steady-state quantile estimation. *Operations Research Letters*, 43:374–377, 2015.
26. Shane G. Henderson and Peter W. Glynn. Regenerative steady-state simulation of discrete event systems. *ACM Transactions on Modeling and Computer Simulation*, 11:313–345, 2001.
27. P. K. Sen. On the Bahadur representation of sample quantiles for sequences of ϕ-mixing random variables. *Journal of Multivariate Analysis*, 2:77–95, 1972.
28. G. J. Babu and K. Singh. On deviations between empirical and quantile processes for mixing random variables. *Journal of Multivariate Analysis*, 8:532–549, 1978.
29. Ju. A. Davydov. Mixing conditions for Markov chains. *Teor. Verojatnost. i Primenen.*, 18:321–338, 1973.
30. K. B. Athreya and S. G. Pantula. Mixing properties of Harris chains and autoregressive processes. *Journal of Applied Probability*, 23:880–892, 1986.
31. Wei Biao Wu. On the Bahadur representation of sample quantiles for dependent sequences. *Annals of Statistics*, 33(4):1934–1963, 2005.
32. Kemal Dinçer Dingeç, Christos Alexopoulos, David Goldsman, Athanasios Lolos, and James R. Wilson. Geometric-moment contraction of G/G/1 waiting times. Manuscript, 2022.
33. R. N. Bhattacharya and R. Ranga Rao. *Normal Approximation and Asymptotic Expansions*. Wiley, New York, 1976.
34. K. L. Chung. *A Course in Probability Theory*, volume 21 of *Probability and Mathematical Statistics*. Academic Press, San Diego, 2nd edition, 1974.
35. Soren Asmussen. *Applied Probability and Queues*, volume 51 of *Applications of Mathematics: Stochastic Modeling and Applied Probability*. Springer, New York, 2nd edition, 2003.
36. S. P. Meyn and R. L. Tweedie. *Markov Chains and Stochastic Stability*. Springer-Verlag, London, 1993.
37. Karl Sigman. One-dependent regenerative processes and queues in continuous time. *Mathematics of Operations Research*, 15(1):175–189, 1990.
38. P. W. Glynn and W. Whitt. Necessary conditions in limit theorems for cumulative processes. *Stochastic Processes and Their Applications*, 98:199–209, 2002.
39. P. W. Glynn and W. Whitt. Limit theorems for cumulative processes. *Stochastic Processes and Their Applications*, 47:299–314, 1993.
40. M.P. Wand and M.C. Jones. *Kernel Smoothing*. Chapman & Hall, London, 1995.
41. Christos Alexopoulos, David Goldsman, and James R. Wilson. A new perspective on batched quantile estimation. In C. Laroque, J. Himmelspach, R. Pasupathy, O. Rose, and A. M. Uhrmacher, editors, *Proceedings of the 2012 Winter Simulation Conference*. IEEE, 2012.
42. J. K. Ghosh. A new proof of the Bahadur representation of quantiles and an application. *Annals of Mathematical Statistics*, 42(6):1957–1961, 1971.
43. S. P. Meyn and R. L. Tweedie. Stability of Markovian processes III: Foster-Lyapunov criteria for continuous-time processes. *Advances in Applied Probability*, 25:518–548, 1993.
44. S. Karlin and H. M. Taylor. *A Second Course in Stochastic Processes*. Academic Press, Boston, 1981.
45. P. W. Glynn and S. P. Meyn. A Liapounov bound for solutions of the Poisson equation. *Annals of Probability*, 24:916–931, 1996.
46. S. G. Henderson and P. W. Glynn. Computing densities for Markov chains via simulation. *Mathematics of Operations Research*, 26:375–400, 2001.