Operations Research Letters 43 (2015) 374–377

Contents lists available at ScienceDirect

Operations Research Letters

journal homepage: www.elsevier.com/locate/orl

A note on bias and mean squared error in steady-state quantile estimation

David F. Muñoz *, Adán Ramírez-López

Departamento de Ingeniería Industrial y Operaciones, Instituto Tecnológico Autónomo de México, Río Hondo # 1, 01080 Mexico City, Mexico

ARTICLE INFO

Article history:
Received 18 November 2013
Received in revised form
6 May 2015
Accepted 6 May 2015
Available online 14 May 2015

Keywords:
Quantile estimation
Steady-state simulation
Simulation output analysis
Batch means

ABSTRACT

We show that, under reasonable assumptions, the performance of the jackknife, classical and batch means estimators for the estimation of quantiles of the steady-state distribution exhibit similar properties as in the case of the estimation of a nonlinear function of a steady-state mean. We present some experimental results from the simulation of the waiting time in queue for an M/M/1 system to confirm our theoretical results.

1. Introduction

Most of the literature in simulation output analysis has been devoted to the estimation of performance measures that are expressed in terms of expected values or long-run averages. In many applications, however, a quantile can be a performance measure of primary interest (e.g., a reorder point for a given service level in inventory management).

Let $\{X_i : i = 0, 1, \ldots\}$ be a stochastic process (with state space $E \subseteq \Re$) representing the output of a stochastic simulation. We assume that the process has a steady-state distribution, i.e.,

$$X_n \Rightarrow X, \tag{1}$$

where X is a random variable with cumulative distribution function (c.d.f.) F, and "$\Rightarrow$" denotes weak convergence (as $n \to \infty$ unless specified). For $0 < p < 1$, we are interested in the estimation of the p-quantile of X defined by

$$\alpha = T_p(F) = \inf\{x : F(x) \geq p\}, \tag{2}$$

from the output $X_1, \ldots, X_n$ of a stochastic simulation. The point estimators that we discuss in this paper require that the simulation output be divided into b non-overlapping batches of size m (we assume that $n = bm$), so that we consider three different point estimators for α.

Classical estimator:

$$\hat{\theta}_1(n) = T_p(F_n), \tag{3}$$

where F_n is the empirical distribution function defined by $F_n(x) \stackrel{def}{=} \frac{1}{n}\sum_{i=1}^{n} I[X_i \leq x], x \in \Re$, T_p is defined in (2) and I denotes the indicator function.

Batch means estimator:

$$\hat{\theta}_2(n) = \frac{1}{b}\sum_{j=1}^{b} \hat{\alpha}_j(n), \tag{4}$$

where, for $j = 1, 2, \ldots, b$, $\hat{\alpha}_j(n) = T_p(F_{jn})$, and $F_{jn}(x) \stackrel{def}{=} \frac{1}{m}\sum_{i=(j-1)m+1}^{jm} I[X_i \leq x]$ is the empirical distribution function corresponding to batch j.

Jackknife estimator:

$$\hat{\theta}_3(n) = \frac{1}{b}\sum_{j=1}^{b} J_j^b, \tag{5}$$

where, for $j = 1, 2, \ldots, b$, $J_j^b = b\hat{\theta}_1(n) - (b-1)\hat{\alpha}_j^J(n)$, $\hat{\alpha}_j^J(n) = T_p\left(F_{jn}^J\right)$, $F_{jn}^J(x) \stackrel{def}{=} \frac{1}{(b-1)m}\sum_{i\in A_j} I[X_i \leq x]$, and $A_j = \{1, \ldots, n\} - \{m(j-1)+1, \ldots, m(j-1)+m\}$.

The variability of these point estimators can be assessed by computing the sample standard deviation

$$S_b(n) = \sqrt{\frac{1}{b-1}\sum_{j=1}^{b}\left[\hat{\alpha}_j(n) - \hat{\theta}_2(n)\right]^2}, \tag{6}$$

* Corresponding author.
E-mail address: davidm@itam.mx (D.F. Muñoz).

http://dx.doi.org/10.1016/j.orl.2015.05.003

where $\hat{\alpha}_j(n)$ and $\hat{\theta}_2(n)$ are defined in (4). An approximate $100(1-\alpha)\%$ confidence interval (CI) for α is given by

$$\left[\hat{\theta}(n) - t_{(b-1,\alpha)} b^{-1/2} S_b(n),\ \hat{\theta}(n) + t_{(b-1,\alpha)} b^{-1/2} S_b(n)\right], \qquad (7)$$

where $t_{(b-1,\alpha)}$ is the $(1-\alpha/2)$-quantile of a Student-t distribution with $(b-1)$ degrees of freedom, $S_b(n)$ is defined in (6), and $\hat{\theta}(n)$ can be any of the point estimators defined in (3)–(5). In the steady-state estimation context, the asymptotic validity of a CI in the form of (7) usually requires a stronger version of a Central Limit Theorem (CLT) for $\hat{\theta}(n)$ (see Assumption 1), and the corresponding CLT implies that $n^{1/2} S_b(n) \Rightarrow 0$, so that the half width of (7) tends to zero at a rate $o\left(n^{-1/2}\right)$.

If we assume that F is differentiable at α, with $F'(\alpha) > 0$, the asymptotic validity of the CI defined in (7) can be verified using Bahadur's representation for quantiles (see, e.g., [1,7]) for each batch

$$\hat{\alpha}_j(n) - \alpha = \frac{p - F_{jn}(\alpha)}{F'(\alpha)} + R_j(n), \qquad (8)$$

for $j = 1, \ldots, b$, where F_{jn} and $\hat{\alpha}_j(n)$ are defined in (4).

As shown in [5], for the cases $\hat{\theta}(n) = \hat{\theta}_1(n)$ or $\hat{\theta}(n) = \hat{\theta}_2(n)$, the following two assumptions are sufficient to verify the asymptotic validity of the CI defined in (7). Set

$$Z_n(t) = \frac{n^{-1/2}}{F'(\alpha)} \sum_{j=1}^{\lfloor nt \rfloor} \left(p - I\left[X_j \le \alpha\right]\right), \qquad (9)$$

for $0 \le t \le 1$.

Assumption 1. There exists a constant $\sigma_0 > 0$ such that for any initial distribution υ_0, $Z_n \Rightarrow \sigma_0 B$, in $D[0,\ 1]$, the space of $\Re$-valued functions on $[0,\ 1]$ that are right continuous and have left limits, where Z_n is defined in (9) and B denotes a standard Brownian motion on $[0,\ 1]$.

Assumption 2. The residuals $R_j(n)$ defined in Eq. (8) satisfy $m^{1/2} R_j(n) \Rightarrow 0, j = 1, \ldots, b$ (where $n = mb$).

As shown in [4], for the case $\hat{\theta}(n) = \hat{\theta}_3(n)$ the asymptotic validity of the CI defined in (7) can be verified by adding the next Assumption. Set

$$\hat{\alpha}_j^J(n) - \alpha = \frac{p - F_{jn}^J(\alpha)}{F'(\alpha)} + R_j^J(n), \qquad (10)$$

Assumption 3. The residuals $R_j^J(n)$defined in (10) satisfy $(n-m)^{1/2} R_j^J(n) \Rightarrow 0, j = 1, \ldots, b$.

Assumption 1 is called a Functional Central Limit Theorem (FCLT) and is a stronger assumption than a Central Limit Theorem. Conditions under which a FCLT holds for a Markov chain (MC) are provided in Theorem 4.1 of [3]. Also, conditions under which Assumption 2 holds for a MC are provided in [5], and the same conditions can be used to verify Assumption 3.

We remark that the CI defined in (7) is a "cancellation procedure" in the sense that we are not trying to (consistently) estimate the asymptotic variance of the estimators, since it may be a difficult task even for the steady-state mean estimation problem (see, e.g., [2]). In order to prove the validity of the CI defined in (7) we cancel out the variance of the estimators in a Central Limit Theorem (see [4,5] for details). Also, we are not discussing data-driven procedures to choose the batch size m, since we believe the methods discussed in [8] can be extended to the case of steady-state quantile estimation. We remark that, in the mentioned paper, the authors propose a sample size large enough to ensure that lack of normality and independence among batch estimators are negligible, and the same ideas can be used for the case of quantile estimation.

In the next section we impose suitable assumptions on the convergence rate of the remainder terms of (8) and (10), and show that, for a sufficiently large run length, the bias of the jackknife estimator is smaller than the bias of the classical estimator, and the bias of the batch means estimator is larger than the bias of the classical estimator. In Section 3 we present experimental results that confirm our theoretical results. Our experiments also show that all three point estimators exhibit a similar performance from the point of view of their mean squared error (mse).

2. Main results

In order to state our results, we will consider the following assumption.

Assumption 4. There exist constants $K \neq 0$ and β $(1/2 < \beta < 1)$ such that the residuals $R_j(n)$ and $R_j^J(n)$ defined in (8) and (10), respectively, satisfy $m^\beta \log(m) R_j(n) \Rightarrow K$, and $(n-m)^\beta \log(n-m) R_j^J(n) \Rightarrow K, j = 1, \ldots, b$.

Note that Assumption 4 implies Assumptions 2 and 3, and is a more precise statement on the rate of convergence of the remainder terms. Conditions under which Assumption 4 is satisfied are provided in [10] (see also [11,9,7]). Let us denote

$$B(n, \alpha) = \left(\frac{p - E\left[F_n(\alpha)\right]}{F'(\alpha)}\right). \qquad (11)$$

Proposition 1. *Suppose that F is differentiable at p, and the stochastic process $\{X_i : i = 0, 1, \ldots\}$ satisfies Assumptions 3 and 4. Let the number of batches b be fixed with $n = mb$. If there exists $n_0 > 0$ such that $\left\{m^\beta \log(m) R_{jn} : n \ge n_0\right\}$ and $\left\{(n-m)^\beta \log(n-m) R_{jn}^J : n \ge n_0\right\}$ are uniformly integrable, then*

$$Bias\left[\hat{\theta}_1(n)\right] = B(n, \alpha) + \frac{K}{n^\beta \log(n)} + o\left(\left(n^\beta \log(n)\right)^{-1}\right), \qquad (12)$$

$$Bias\left[\hat{\theta}_2(n)\right] = B(n, \alpha) + \frac{b^\beta K}{n^\beta \log(m)} + o\left(\left(n^\beta \log(n)\right)^{-1}\right), \qquad (13)$$

and

$$Bias\left[\hat{\theta}_3(n)\right] = B(n, \alpha) + \frac{a(b, \beta) K}{n^\beta \log(n)} + o\left(\left(n^\beta \log(n)\right)^{-1}\right), \qquad (14)$$

as $n \to \infty$, where $a(b, \beta) = b\left[1 - (1 - 1/b)^{1-\beta}\right]$ and $B(n, \alpha)$ is defined in (11).

The proof of Proposition 1 is provided in the Appendix. Note that $I[X_i \le \alpha]$ may be biased due to an initial transient. However, if the process $\{X_i : i = 0, 1, \ldots\}$ is stationary and the initial distribution is the steady-state distribution, we have $E[I[X_i \le \alpha]] = p$, so that $B(n, \alpha) = 0$. In a typical situation, the first term of Eqs. (11)–(13) converges to zero faster than the second term. For example, when $X_i = g(Y_i)$, where $Y = \{Y_i : i = 0, 1, \ldots\}$ is a Markov chain (MC) with state-space $E \subseteq \Re^d$, a sufficient condition (see [5]) for the validity of the CI defined in (7) is that the MC be geometrically ergodic, i.e., there exist $\rho < 1$ and a real-valued function h such that:

$$d_k(y) \le h(y) \rho^j, \qquad (15)$$