Reliable estimation via simulation

Vinod Sharma

Department of Electrical Engineering, Indian Institute of Science, Bangalore 560 012, India

Received 22 December 1993; revised 22 July 1994

Let a and s denote the inter arrival times and service times in a $GI/GI/1$ queue. Let $a^{(n)}$, $s^{(n)}$ be the r.v.s. with distributions as the estimated distributions of a and s from iid samples of a and s of sizes n. Let w be a r.v. with the stationary distribution π of the waiting times of the queue with input (a, s). We consider the problem of estimating $E[w^\alpha]$, $\alpha > 0$ and π via simulations when $(a^{(n)}, s^{(n)})$ are used as input. Conditions for the accuracy of the asymptotic estimate, continuity of the asymptotic variance and uniformity in the rate of convergence to the estimate are obtained. We also obtain rates of convergence for sample moments, the empirical process and the quantile process for the regenerative processes. Robust estimates are also obtained when an outlier contaminated sample of a and s is provided. In the process we obtain consistency, continuity and asymptotic normality of M-estimators for stationary sequences. Some robustness results for Markov processes are included.

Keywords: Continuity; rates of convergence; robust estimation; queueing systems; simulation; regenerative processes.

1. Introduction

Suppose we model a queueing system as a $GI/GI/1$ queue. One of the main performance parameters of interest for the system is the mean waiting time Ew under stationarity and the stationary distribution π. Since closed form expressions for Ew are not available for a general $GI/GI/1$ queue, usually one obtains Ew via simulations. For this, one first needs the distributions of the inter arrival time a and the service time s. Invariably, for any practical system these distributions need to be estimated from raw data. The problem of estimation of distributions from iid data is a classical one. Either one assumes that the unknown distributions belong to a parametric family with the parameter belonging to a subset of a finite dimensional space or one estimates the probability density or cumulative distribution function in a nonparametric way. For the parameter estimation problem see Lehmann [29] for an extensive treatment and references therein and for the non-parametric problem see Bean and Tsokos [8] and Prakasa Rao [34]. Once estimates of the distributions of a and s are obtained, these can be used in the simulation to obtain Ew.

The problem of simulation of queueing systems has received increasing attention recently (see Asmussen [2], Glynn and Whitt [16], Whitt [43] and the review papers by Pawlikowski [32] and Lavenberg and Welch [27]). These papers address the problem of determining the simulation length required for certain accuracy and how to improve the accuracy for a given run length. The problem of parametric estimation for queueing systems is reviewed in Bhat and Rao [10]. Some other recent references are Basawa and Prabhu [6] and Thiruvairaya and Basawa [39] and Whitt [44].

In this paper we address the problem of estimating the stationary distribution and moments of the waiting time when estimated distributions of a and s are used in the simulation. This problem does not seem to have been addressed till now. We point out the requirements on the estimates of distributions of a and s so that we obtain good estimates of the distribution and the moments of waiting times. The main performance indices for estimators are rate of convergence to the asymptotic value, the asymptotic bias and the asymptotic variance (see Lehmann [29] and Glynn and Whitt [16]). Then the question arises when can we make sure that these efficiency parameters for estimators of Ew will not be affected if the distributions of a and s differ some what. Our results indicate that the goodness of fit using Kolmogorov–Smirnov distance may not provide satisfactory estimates for distributions of a and s but the empirical distribution function does satisfy all our requirements. Incidently using empirical distribution functions for a and s in simulation makes our simulations a boot strap method, very popular in the recent statistics literature (see Lepage and Billard [30]).

Next we consider the problem of estimating Ew and π when the distributions of a and s are estimated from a "contaminated" sample of a and s. Contamination of data by "outliers" and round off errors is a much more common problem than is generally believed (see chapter 1 in Hampel et al. [18]). This is the problem of "robust" estimation and has been addressed in the context of parametric estimation from iid sample in Huber [20] and Hampel et al. [18]. We will consider this problem in the context of queueing systems and regenerative processes in general. This problem has not been studied in the literature so far.

Although we address these problems in the context of estimating Ew^{α} and π for a $GI/GI/1$ system, these problems and the solutions we propose are relevant in simulation of other performance indices and for other queueing systems. We also obtain some general results on rates of convergence for regenerative processes and robust estimates for Markov processes.

The paper is planned as follows. In section 2 we obtain the continuity and rates of convergence of estimates from an iid sample of a and s. General results on regenerative processes are also presented. Section 3 provides continuity of estimates from a contaminated sample of a and s. For this we also obtain consistency and asymptotic continuity of M-estimators. Robustness of estimates from a contaminated sample of a Markov process are obtained in section 4, where we also obtain asymptotic normality of M-estimators. These are useful when stationary distributions and moments of w are estimated from a contaminated sample of waiting times.

2. Estimation based on iid sample

Assume we have an iid sample of size n available for the interarrival times a and service times s. We can estimate the means Ea and Es by the sample means and verify for the stability condition $Ea > Es$. If the system is stable we want to estimate Ew^α (if it exists) and π. Let $a^{(n)}$ and $s^{(n)}$ be r.v.s. with the distributions as the estimated distributions of a and s from this sample (the estimated distributions are r.v.s. and hence in the following results the conditions on the distributions of $a^{(n)}$, $s^{(n)}$ and on $E[(a^{(n)})^\alpha]$, $E[(s^{(n)})^\alpha]$ will be in the a.s. sense; but we will not mention it). Then we can use simulation with the inter arrival times and service times generated as iid sequences with the distributions of $a^{(n)}$ and $s^{(n)}$. We denote the estimates so obtained as $Ew^{(n)}$ and $\pi^{(n)}$. The estimates $Ew^{(n)}$ and $\pi^{(n)}$ not only depend upon the distributions of $a^{(n)}$ and $s^{(n)}$ but also on the simulation length. Since we are interested in finding the requirements on the estimation of $a^{(n)}$ and $s^{(n)}$ to obtain satisfactory estimates for Ew and π, we denote by $Ew^{(n)}$ and $\pi^{(n)}$ the asymptotic values of consistent estimates of the mean and the distributions and hence the actual quantities corresponding to input $(a^{(n)}, s^{(n)})$. To obtain accurate estimates we obtain conditions for $E(w^{(n)})^\alpha \to Ew^\alpha$, and $\pi^{(n)} \xrightarrow{w} \pi$ a.s. ($\xrightarrow{w}$ denotes weak convergence) and the effect of estimates $a^{(n)}$ and $s^{(n)}$ on the rates of convergence in the simulation and on the asymptotic variance of the estimates of $Ew^{(n)}$. We shall also consider asymptotic loss under more general loss functions. For reference, we will use the sample mean of waiting times as an estimator of $Ew^{(n)}$. This estimator has also been considered by Whitt [43] and Asmussen [2]. We have in fact obtained the rates of convergence results for general regenerative processes.

Let us denote by $\tau^{(n)}$ the regeneration length for the waiting time process (the regeneration epochs are the arrival instants finding the system empty) when the r.v.s $a^{(n)}$ and $s^{(n)}$ are used in simulation. The following result from Kalashnikov [23, pp. 130–131] will be repeatedly used in this section.

LEMMA 1

Let there exist $\alpha > 0$, $\Delta > 0$, $b < \infty$ and $p > 1$ such that $E[s^{(n)}] < E[a^{(n)}]$,

$$\sup_n E[s^{(n)} - \min(\alpha, a^{(n)})] \le -\Delta$$

and $\sup_n E[(s^{(n)})^p] \le b$. Then $\sup_n E[(\tau^{(n)})^p] < \infty$. □

From Lemma 1, if $E[s] < E[a]$, $s^{(n)} \xrightarrow{w} s$, $a^{(n)} \xrightarrow{w} a$, $\sup_n E[(s^{(n)})^p] < \infty$ and $E[s^p] < \infty$ then $\sup_n E[(\tau^{(n)})^p] < \infty$ and $E[\tau^p] < \infty$.

We will also be interested in exponential moments of τ. From Kalashnikov and Rachev [24, p. 279] if in addition to the conditions of Lemma 1, we also have $\sup_n E[\exp(\beta s^{(n)})] \le c$, for some positive β and c then $\sup_n E[\exp(\gamma\tau^{(n)})] \le$

$\exp(2\alpha\gamma/\Delta)$ where

$$\gamma = \min\left(\left(\frac{\Delta^2\beta^2}{4c}\exp\left(\frac{-\Delta\beta}{2}\right)\right), \beta\right).$$

Since we will be interested in estimating π and $E[w^\alpha]$ from simulations, we also need conditions for $\pi^{(n)} \xrightarrow{w} \pi$ and $E[(w^{(n)})^\alpha] \to E[w^\alpha]$. From Kalashnikov [23, p. 256] we obtain that if $s^{(n)} \xrightarrow{w} s$, $a^{(n)} \xrightarrow{w} a$, $E[s] < E[a]$ and $\{s^{(n)}\}$ are uniformly integrable then $\pi^{(n)} \xrightarrow{w} \pi$. Also, from Kalashnikov [22], we obtain the following lemma.

LEMMA 2

Let $a^{(n)} \xrightarrow{w} a$, $s^{(n)} \xrightarrow{w} s$, $E[s] < E[a]$ and $\sup_n E[(s_1^{(n)})^{1+\alpha'}] < \infty$ for some $\alpha' > \alpha > 0$. Then $E[(w^{(n)})^\alpha] \to E[w^\alpha] < \infty$. □

Now we obtain upper bounds on $|E(w) - E(w')|$ where w and w' denote r.v.s. with the stationary distributions of waiting times corresponding to the iid inputs $\{a_n, s_n\}$ and $\{a'_n, s'_n\}$ respectively. These can provide rates of convergence in the above convergence result. Let $x_n = s_n - a_n$, $x'_n = s'_n - a'_n$ and w_n and w'_n denote the corresponding waiting times and let τ and τ' be the first time after $n = 0$, w_n, w'_n are zero (taking $w_0 = 0 = w'_0$). If $Ex_1 < 0$, $Ex'_1 < 0$, $E[\tau^\alpha] < c_1$ and $E[(\tau')^\alpha] < c_1$ for some $\alpha > 1$ and some $c_1 < \infty$, we can compute a constant c and a r.v. $\hat{\tau}$ such that $\hat{\tau}$ is a common regeneration epoch (when $w_n = 0 = w'_n$) of $\{w_k\}$ and $\{w'_k\}$ and $E[(\hat{\tau})^\alpha] \le c$ (Kalashnikov [23]) (actually one may have to redefine the stochastic processes appropriately). Now we have

$$\begin{aligned} E[\hat{\tau}][E[w] - E[w']] &= E\left[\sum_{k=1}^{\hat{\tau}-1} [w_k - w'_k]\right] \\ &= E\left[\sum_{k=1}^{\hat{\tau}-1} \sum_{n=0}^{k-1} [x_n - x'_n]\right]. \end{aligned}$$

Also, $\{y_k, n \ge 1\}$ is a martingale with respect to the natural filtration of $\{(x_n, x'_n)\}$ where

$$y_k = \left(\sum_{n=1}^{k} [x_n - x'_n]\right) - kE[x_1 - x'_1].$$

Thus, $E[\sum_{k=1}^{\hat{\tau}} y_k] = E[\hat{\tau} y_{\hat{\tau}}] \le E[\hat{\tau}^2]^{1/2} E[(y_{\hat{\tau}})^2]^{1/2} = E[\hat{\tau}^2]^{1/2}(\mathrm{var}(x_1 - x'_1)E\hat{\tau})^{1/2}$. Also,

$$E\left[\sum_{k=1}^{\hat{\tau}-1} \sum_{n=1}^{k} |x_n - x'_n|\right] = (E[(\hat{\tau})^2] - E[\hat{\tau}])E[|x_1 - x'_1|]/2$$

when the second moments of s, s', a, a', τ, τ' are finite.

From Kalashnikov [23, chapter 5], one can explicitly obtain upper bounds on $E[(\tau)^p]$ and then from Kalashnikov [23, chapter 6] on $[E(\hat{\tau})^p]$. One can also obtain upper bound on $|Ew - Ew'|$ using the upper and lower bounds on Ew available in Shanthikumar [37]. Similarly one can obtain upper bounds on $|E(w)^\alpha - E(w')^\alpha|$, $\alpha > 1$. Kalashnikov [21] and Kalashnikov and Rachev [24] provide bounds on $d(w, w')$ where d can be one of the various probability metrics provided there. In fact, from Kalashnikov [23] one can obtain explicit bounds on $\sup_k d(w_k, w_k')$ which can be very useful in our simulation context.

It would be rather unpleasant if using estimators $a^{(n)}$ and $s^{(n)}$ would lead to different rates of convergence in the simulation experiment for different n. Now we obtain sufficient conditions for certain minimal uniform (w.r.t. n) rates of convergence to stationary distributions in the simulation experiments. First, for the rate of convergence of $\pi_k^{(n)}$ (the distribution of $w_k^{(n)}$) to $\pi^{(n)}$, we use Lemma 1 and a result of Kalashnikov [23]: denoting by d the total variation metric, under the conditions of Lemma 1 with some $\alpha > 1$, $\sup_n d(\pi_k^{(n)}, \pi^{(n)}) \leq c_1 k^{1-\alpha}$. But for the estimation purposes, more important are the rates of convergence of the strong laws of large numbers (SLLN). One way to obtain these rates is from the following functional central limit theorems and functional laws of iterated logarithm for regenerative processes that we have obtained in Sharma [38]. Let $\{z_k\}$ be a real valued regenerative sequence, τ its regeneration length, and let $U_1 = \sum_{k=0}^{\tau-1} z_k$ ($k = 0$ is a regeneration epoch).

THEOREM 1

Let $0 < Var(U_1), E[(\sum_{k=0}^{\tau-1} |z_k|)^2] < \infty$ and $E[\tau] < \infty$. Then $\{z_k\}$ satisfies the functional central limit theorem and functional law of iterated logarithm. Further, if $E[|U_1|^{r_1}] < \infty$ and $E[\tau^{r_2}] < \infty$ for some $r_1 > 2, r_2 > 2$ then we can redefine $\{z_k\}$ on a probability space along with a standard Wiener process W such that

$$\sup_{0 \leq n \leq m} \left| \left(\sum_{k=0}^{n} z_k - nEz \right) \Big/ \theta - W(n) \right| = 0(m^\rho) \quad \text{a.s.}$$

where $\theta = var\, U_1/E\tau + (E\tau Ez)^2\, var\, \tau/(E\tau)^3$, ρ is any constant greater than $5/12 + max\,(1/r_1, 1/r_2)/6$ and Ez is the mean of z_1 under stationarity. □

Sharma [38] has proved the above result for a continuous time regenerative process also and the regeneration cycles can be m-dependent for m some nonnegative constant with the constant θ appropriately modified.

Using Theorem 1 (see also Asmussen [1]), we provide conditions for asymptotic normality of our estimates and then show the continuity of the asymptotic variance. These two facts imply that the accuracy of the estimates of $Ew_k^{(n)}$ at finite simulation time will be comparable if n is large. Since $\{w_k^{(n)}\}$ is a regenerative process, from Theorem 1, if $E[(\tau^{(n)})] < \infty$ and $E[(\sum_{k=1}^{\tau^{(n)}} w_k^{(n)})^2 | k = 0$ is a regeneration epoch] $< \infty$, then $(\sum_{k=1}^{m}(w_k^{(n)} - E[w^{(n)}]))/m^{1/2} \xrightarrow{w} N(0,(\sigma^{(n)})^2)$ where $N(0,\sigma^2)$ is a r.v. with normal distribution, mean zero and variance σ^2. Also, taking $k = 0$ as a regeneration epoch (see Asmussen [1]), we obtain

$$(\sigma^{(n)})^2 = var\left[\sum_{k=1}^{\tau^{(n)}} w_k^{(n)}\right] + \left(E\left[\sum_{k=1}^{\tau^{(n)}} w_k^{(n)}\right] E[\tau^{(n)}]^{-1}\right)^2 var\,(\tau^{(n)})$$
$$- 2E\left[\sum_{k=1}^{\tau^{(n)}} w_k^{(n)}\right] cov\left(\tau^{(n)}, \sum_{k=1}^{(n)} w_k^{(n)}\right)(E[\tau^{(n)}])^{-1}.$$

Now,

$$E\left[\left(\sum_{k=1}^{\tau^{(n)}} w_k^{(n)}\right)^{\alpha}\right] \le E\left[(\tau^{(n)})^{\alpha}\left(\max_{1\le k\le \tau^{(n)}} w_k^{(n)}\right)^{\alpha}\right]$$
$$\le E[(\tau^{(n)})^{2\alpha}]^{1/2} E\left[\left(\max_{1\le k\le \tau^{(n)}} w_k^{(n)}\right)^{2\alpha}\right]^{1/2}$$
$$\le E[(\tau^{(n)})^{2\alpha}]^{1/2} E\left[\left(\sum_{k=1}^{\tau^{(n)}} (w_k^{(n)})^{2\alpha}\right)\right]^{1/2}.$$

Using Lemma 1 we find that $E[(\tau^{(n)})^4] < \infty$ if $E[(s^{(n)})^4] < \infty$. Also, it is known that $E[\tau^{(n)}]E[(w^{(n)})^4] = E[\sum_{k=1}^{\tau^{(n)}}(w_k^{(n)})^4] < \infty$, if $E[(s^{(n)})^5] < \infty$ and hence we obtain central limit theorem if $E[(s^{(n)})^5] < \infty$.

Next we obtain the conditions for continuity of $(\sigma^{(n)})^2$. We denote by $\{w_k\}$ the waiting time sequence and by σ^2 its asymptotic variance corresponding to (a, s).

THEOREM 2

Let $a^{(n)} \xrightarrow{w} a$, $s^{(n)} \xrightarrow{w} s$, $Es < Ea$ and $\sup_n E[(s_1^{(n)})^{5+\alpha}] < \infty$ for some $\alpha > 0$. Then $(\sigma^{(n)})^2 \to \sigma^2 < \infty$ as $n \to \infty$.

Proof

We will prove $(\sigma^{(n)})^2 \to \sigma^2$ by providing these limits for each component in (1) separately.

We note that under our conditions

$$(w_1^{(n)}, w_2^{(n)}, \ldots, w_k^{(n)}) \xrightarrow{w} (w_1, w_2, \ldots, w_k)$$

for all $k \geq 1$ if $w_0^{(n)} \to w_0$. This implies that

$$(w_1^{(n)}, w_2^{(n)}, \ldots, w_{\tau^{(n)}}^{(n)}, \tau^{(n)}) \xrightarrow{w} (w_1, w_2, \ldots, w_\tau, \tau)$$

because $\tau^{(n)}$ and τ are the stopping times and hence

$$\sum_{k=1}^{\tau^{(n)}} w_k^{(n)} \xrightarrow{w} \sum_{k=1}^{\tau} w_k. \tag{3}$$

From Lemma 1, since $\tau^{(m)} \xrightarrow{w} \tau$, because of uniform integrability we obtain

$$E[(\tau^{(n)})^\beta] \to E[(\tau)^\beta]$$

for $0 \leq \beta \leq 4$, and hence using Lemma 2

$$E\left[\sum_{k=1}^{\tau^{(n)}} w_k^{(n)}\right] = E[\tau^{(n)}]E[w^{(n)}] \to E[\tau]E[w] = E\left[\sum_{k=1}^{\tau} w_k\right].$$

Again as in (2) we obtain

$$E\left[\left(\sum_{k=1}^{\tau^{(n)}} w_k^{(n)}\right)^{2+\epsilon}\right] \leq E[(\tau^{(n)})^{4+2\epsilon}]^{1/2} E[(w^{(n)})^{4+2\epsilon}]^{1/2} E[\tau^{(n)}]^{1/2} \tag{4}$$

and hence using Lemma 1 and the fact that $E[(w^{(n)})^{4+2\epsilon}] \to E[w^{4+2\epsilon}]$ for $\epsilon < \alpha/2$, we obtain

$$\sup_n E\left[\left(\sum_{k=1}^{\tau^{(n)}} w_k^{(n)}\right)^{2+\epsilon}\right] < \infty$$

under our conditions. Now (3) implies that

$$E\left[\left(\sum_{k=1}^{\tau^{(n)}} w_k^{(n)}\right)^2\right] \to E\left[\left(\sum_{k=1}^{\tau} w_k\right)^2\right].$$

Finally, since

$$E\left[\left((\tau^{(n)}-E(\tau^{(n)}))\cdot\left(\sum_{k=1}^{\tau^{(n)}} w_k^{(n)}-E[\tau^{(n)}]E[w^{(n)}]\right)^{\beta}\right]$$

$$\le E[(\tau^{(n)}-E(\tau^{(n)}))^2]^{1/2}E\left[\left(\sum_{k=1}^{\tau^{(n)}} w_k^{(n)}-E[\tau^{(n)}]E[w^{(n)}]\right)^{2\beta}\right]^{1/2}$$

again using (4) and Lemma 1 we obtain uniform integrability and hence

$$cov\left(\tau^{(n)},\sum_{k=1}^{\tau^{(n)}} w_k^{(n)}\right)\to cov\left(\tau,\sum_{k=1}^{\tau} w_k\right).$$

This proves the result. □

Theorem 2 also provides the continuity of asymptotic variance of other loss functions. Let T_n be an estimator of θ (the parameter to be estimated) based on a sample of size n and let $L(T_n-\theta)$ be the corresponding loss. Also let $(T_n-\theta)\xrightarrow{p}0$, $\sqrt{n}(T_n-\theta)\xrightarrow{w}N(0,\sigma^2)$, $L(0)=0$ and let L be twice continuously differentiable. Then by Taylor series

$$L(T_n-\theta)=L(0)+(T_n-\theta)L'(0)+\frac{(T_n-\theta)^2}{2}L''(y_n)$$

where $|y_n|\le|T_n-\theta|$. Since $|T_n-\theta|\xrightarrow{p}0$, $y_n\xrightarrow{p}0$ and hence $L''(y_n)\xrightarrow{p}L''(0)$. Also $\sqrt{n}(T_n-\theta)\xrightarrow{w}N(0,\sigma^2)$ implies that $\sqrt{n}(T_n-\theta)^2\xrightarrow{p}0$. Therefore $\sqrt{n}L(T_n-\theta)\xrightarrow{w} L'(0)N(0,\sigma^2)$. Thus, under conditions of Theorem 2, we will have continuity of the asymptotic variance of a loss function for the estimators of the waiting times.

Now we provide certain rates of convergence for regenerative processes which can be directly applied to the $GI/GI/1$ queue. Thus we will be able to obtain conditions for estimating a and s such that these rates of convergence are not affected.

We use the following notation in addition to the notation in Theorem 1. Let T_i be the regeneration epoches ($T_0=0$) and $N(n)$ the number of regenerations till time n. Also let $U_0=0$, $U_k=\sum_{n=T_{k-1}}^{T_k-1}z_n$, $k\ge1$ and $\tilde{U}_k=\sum_{n=T_{k-1}}^{T_k-1}|z_n|$. Then we obtain

THEOREM 3

Let there be constants r and α with $r \geq 1$, $\frac{1}{2} < \alpha \leq 1$, $\alpha r \geq 1$ such that $E[(\tilde{U}_1)^r] < \infty$ and $E[\tau^r] < \infty$, then, for any $\epsilon > 0$,

$$\sum_{n=1}^{\infty} n^{\alpha r - 2} P\left[\left|\sum_{k=1}^{n} (z_k - E(z))\right| > n^{\alpha}\epsilon\right] < \infty.$$

Proof

Since, for $U_{\Delta(n)} \triangleq \sum_{k=T_{N(n)}+1}^{n} z_k \leq \tilde{U}_{N(n)+1}$, we obtain

$$\begin{aligned}\left|\sum_{k=1}^{n} (z_k - E(z))\right| &= \left|\sum_{k=1}^{N(n)} U_k + U_{\Delta(n)} - nE[z]\right| \\ &\leq \left|\sum_{k=1}^{N(n)} (U_k - E[\tau]\cdot E[z])\right| + |N(n)E[\tau]E[z] - nE[z]| \\ &\quad + |\tilde{U}_{N(n)+1}|.\end{aligned}$$

Therefore,

$$\begin{aligned}&\sum_{n=1}^{\infty} n^{\alpha r-2} P\left[\left|\sum_{k=1}^{n} (z_k - E[z])\right| > n^{\alpha}\epsilon\right] \\ &\leq \sum_{n=1}^{\infty}\left\{P\left[\left|\sum_{k=1}^{N(n)} (U_k - E[\tau]E[z])\right| > n^{\alpha}\epsilon/3\right]\right. \\ &\quad + P[|N(n)E[\tau]E[z] - nE[z]| > n^{\alpha}\epsilon/3] \\ &\quad \left. + P[|\tilde{U}_{N(n)+1}| > n^{\alpha}\epsilon/3]\right\} n^{\alpha r-2}\end{aligned}$$

and hence it is sufficient to show that the three summations on the right side are finite.

Since $E[\tau^r] < \infty$, from Gut [17, p. 103], we obtain, for all $\delta > 0$,

$$\sum_{n=1}^{\infty} n^{\alpha r-2} P[|N(n) + 1 - n/E[\tau]| > n^{\alpha}\delta] < \infty. \tag{5}$$

Also,

$$\begin{aligned}P[|N(n) - n/E[\tau]| > n^{\alpha}\delta] &\leq P[|N(n) + 1 - n/E[\tau]| + 1 > n^{\alpha}\delta] \\ &\leq P[|N(n) + 1 - n/E[\tau]| > n^{\alpha}\delta/2]\end{aligned}$$

for all n large enough. Thus we obtain

$$\sum_{n=1}^{\infty} n^{\alpha r-2} P[|N(n) - n/E[\tau]| > n^{\alpha}\delta] < \infty$$

for all $\delta > 0$. Also from Gut [17, p. 43], we obtain

$$\sum_{n=1}^{\infty} n^{\alpha r-2} P\left[\left|\sum_{k=1}^{N(n)} (U_k - E[\tau]E[z])\right| > n^{\alpha}\epsilon/3\right] < \infty.$$

Next we show

$$\sum_{n=1}^{\infty} n^{\alpha r-2} P[\tilde{U}_{N(n)+1} > n^{\alpha}\epsilon/3] < \infty.$$

For this, since

$$\sum_{n=1}^{\infty} n^{\alpha r-2} P[\tilde{U}_{N(n)+1} > n^{\alpha}\epsilon/3]$$

$$\leq \sum_{n=1}^{\infty} n^{\alpha r-2} P[\tilde{U}_{N(n)+1} > n^{\alpha}\epsilon/3, |N(n) + 1 - n/E[\tau]| > n^{\alpha}\delta]$$

$$+\sum_{n=1}^{\infty} n^{\alpha r-2} P[\tilde{U}_{N(n)+1} > n^{\alpha}\epsilon/3, |N(n) + 1 - n/E[\tau]| \leq n^{\alpha}\delta]$$

using (5), we only need to prove the finiteness of the second summation on the right. But this summation is not greater than

$$\sum_{n=1}^{\infty} n^{\alpha r-2} P\left[\max_{-n^{\alpha}\delta+n/E\tau \leq k \leq n^{\alpha}\delta+n/E\tau} \tilde{U}_k > n^{\alpha}\epsilon/3\right]$$

$$\leq 2\delta \sum_{n=1}^{\infty} n^{\alpha(r+1)-2} P[\tilde{U}_n > n^{\alpha}\epsilon/3]$$

where the right side is finite when $E[\tilde{U}_1^r] < \infty$. This completes the proof of the theorem. □

Under the conditions of the above theorem we also obtain

$$\sum_{n=1}^{\infty} n^{\alpha r-2} P\left[\sup_{0\leq t\leq 1}\left|\sum_{k=1}^{[nt]} z_k - tEz\right| > n^{\alpha}\epsilon\right] < \infty.$$

For this, note that for each t, the above relation can be obtained from the theorem. Then using monotonicity, the result can be shown uniformly over t, for nonnegative z_k. For general z_k obtain it separately over z_k^+ and z_k^- where $z_k^+ = max\,(0, z_k)$ and $z_k^- = -min\,(0, z_k)$.

We can also obtain exponential rates of convergence in the above theorem for $\alpha = 1$ when $E[e^{t\tilde{U}_1}] < \infty$ and $E[e^{t\tau}] < \infty$ for some $t > 0$ using the following facts:

(i) From Heathcote [19], for $\epsilon > 0$, $P[|\sum_{k=1}^{n}(U_k - E[\tau]E[z])| > \epsilon n]$ converges exponentially.

(ii) From Tiefeng [40], for $\epsilon > 0$, $P[|N(n)/n - 1/E\tau| > \epsilon]$ converges exponentially.

(iii) From (i) and (ii) we can show that $P[|\sum_{k=1}^{N(n)}(U_k - E[\tau]E[z])| > \epsilon N(n)]$ and then $P[|\sum_{k=1}^{N(n)}(U_k - E[\tau]E[z])| > \epsilon n]$ converges exponentially.

(iv) $P[|\Delta U_n| > \epsilon n]$ goes to zero exponentially fast.

Thus we obtain that

$$P\left[\left|\sum_{k=1}^{n}(z_k - Ez)\right| > n\epsilon\right] < e^{-\alpha n},$$

where α can be explicitly calculated. Now of course we also obtain

$$P\left[\sup_{m \le n}\left|\sum_{k=1}^{n}(z_k - Ez)\right| > n\epsilon\right] < e^{-\alpha m}/(1 - e^{-\alpha}).$$

Theorem 3 can be used to obtain estimates of rate of convergence of the sample moments of w_k to $E[w^\alpha]$. Lemma 1 provides conditions for $\sup_n E[(\tau^{(n)})^\alpha] < \infty$ and sufficient conditions for $E[(\tilde{U}_1)^\alpha] < \infty$ are obtained from (2).

The following lemma is useful for exponential convergence.

LEMMA 3

Let $\{x_n\}$ be an iid real valued sequence, τ a positive integer valued stopping time w.r.t. $\{x_n\}$ and let $E[e^{tx_1}] \triangleq c < \infty$ and $E[c^\tau] < \infty$. Then $E[\exp(t/2\sum_{k=1}^{\tau} x_k)] < \infty$.

Proof

Since $\{\exp(t\sum_{k=1}^{n} x_k)/c^n, n \ge 1\}$ is a martingale, by optional sampling theorem $\{\exp(t\sum_{k=1}^{\tau\wedge n} x_k)/c^{\tau\wedge n}\}$ is also a martingale. Therefore, $E[\exp(t\sum_{k=1}^{\tau\wedge n} x_k)/c^{\tau\wedge n}] = 1$. Since $\lim_{n\to\infty}\exp(t\sum_{k=1}^{\tau\wedge n} x_k)/c^{n\wedge\tau} = \exp(t\sum_{k=1}^{\tau} x_k)/c^\tau$ a.s., Fatou's lemma implies $E[\exp(t\sum_{k=1}^{\tau} x_k)/c^{\tau}] \le 1$. Now

$$E\left[\exp\left(\left(\sum_{k=1}^{\tau} x_k\right)t/2\right)\right] \le E\left[\exp\left(\left(\sum_{k=1}^{\tau} x_k\right)t\right)/c^{\tau}\right]^{1/2}(E[c^{\tau}])^{1/2}$$

provides the result □

Now we obtain rates of convergence for empirical and quantile processes. For this using results for iid r.v.s., as in Theorem 3 is not suitable. To be able to use the results for stationary sequences, we will use the following lemma. We denote

$$\alpha(n) = \sup_{m\ge 1}\{|P(AB) - P(A)P(B)|, A \in \sigma(z_1,\dots,z_m), B \in \sigma(z_{m+n},\dots)\}.$$

Sequence $\{z_k\}$ is called strongly mixing if $\alpha(n) \to 0$.

LEMMA 4

Let $\{z_k\}$ be a zero delayed regenerative sequence and $\{\tilde{z}_k\}$ its stationary version if it exists. Then the $\alpha(n)$ for $\{z_k\}$ and $\{\tilde{z}_k\}$ satisfy $\alpha(n) \le c/n^r$ (c may be different for $\{z_k\}$ and $\{\tilde{z}_k\}$) whenever $E[\tau^{r+1}] < \infty$, $r > 0$.

Proof

Let $A \in \sigma(z_1,\dots,z_m)$, $B \in \sigma(z_{m+n}, z_{m+n+1},\dots)$ and $C = \{$a regeneration occurs in time $m,\dots,m+n-1\}$. Then $P(AB|C) = P(A|C)P(B|C)$. Therefore,

$$\begin{aligned}|P(AB) - P(A)P(B)| &= |P(A|C)P(B|C)P(C) + P(ABC^c)\\ &\quad - P(A|C)P(C)P(B) - P(AC^c)P(B)|\\ &\le P(A|C)P(C)|P(B|C) - P(B)|\\ &\quad + |P(ABC^c) - P(AC^c)P(B)|\\ &\le |P(B|C) - P(B|C)P(C)| + P(BC^c) + P(C^c) \le 3P(C^c).\end{aligned}$$

When $E[\tau^{r+1}] < \infty$, then from Gut [17, p. 59], $\sup_m E[R(m)^r] < \infty$, where $R(m)$ is the time to next regeneration at time m. Thus

$$P(C^c) \le \sup_m E[R(m)^r]/n^r.$$

In case of $\{\tilde{z}_k\}$, $\{R(m)\}$ is stationary and $E[(R(m))^r] < \infty$ when $E[\tau^{r+1}] < \infty$. □

The above lemma directly generalizes to continuous time R^d valued regenerative processes with paths in $D[0,\infty)$.

The examples of countable state ergodic Markov chains which do not converge to stationary distributions exponentially show that we may not obtain for $\{z_k\}$ or $\{\tilde{z}_k\}$ the ϕ-mixing or ψ-mixing property where

$$\phi(n) = \sup_{m \geq 1} \{ |P(B|A) - P(B)|, A \in \sigma(z_1, \ldots, z_m), B \in \sigma(z_{m+n} \ldots)\}$$

$$\psi(n) = \sup_{m \geq 1} \left\{ \left| \frac{P(AB) - P(A)P(B)}{P(A)P(B)} \right|, A \in \sigma(z_1, \ldots, z_m), B \in \sigma(z_{m+n} \ldots) \right\}$$

when $E[\tau^{r+1}] < \infty$ for some $r > 0$.

The above lemma extends to the case (for $\{\tilde{z}_k\}$) when the regeneration cycles are m-dependent (general Harris ergodic Markov chains are 1-dependent) if we define the set $C = \{$there are $m+1$ regenerations in time $l, \ldots, l+n-1\}$. Now to obtain $P(C^c) \leq C_1/n^r$ when $E[\tau^{r+1}]$, we can use Janson [21, Theorem 3.1].

Next we use Lemma 4 to obtain an approximation result for the empirical process of $\{z_k\}$ and $\{\tilde{z}_k\}$.

THEOREM 4

Let F be the marginal distribution of $\tilde{z}_1$. Assume $E[\tau^{r+1}] < \infty$ for some $r > 4$, if F is continuous, otherwise let $r > 6$. Then we can redefine (if necessary) on a probability space $\{\tilde{z}_k\}$ and $\{B_k\}$ where B_k is a Brownian bridge such that, for some $\lambda > 0$ (depending only on r)

$$\sup_s \left| \sum_{k=1}^{n} (1\{\tilde{z}_k \leq s\} - F(s)) - \sum_{k=1}^{n} B_k(F(s)) \right| = O(n^{1/2}(\log n)^{-\lambda}) \quad \text{a.s.} \tag{6}$$

Also, if τ is aperiodic then we can define on a probability space $\{z_k\}$, $\{\tilde{z}_k\}$ and $\{B_k\}$ such that (6) and

$$\sup_s \left| \sum_{k=1}^{n} (1\{z_k \leq s\} - F(s)) - \sum_{k=1}^{n} B_k(F(s)) \right| = O(n^{1/2}(\log n)^{-\lambda}) \quad \text{a.s.} \tag{7}$$

hold.

Proof

From Lemma 4, $\alpha(n) = O(n^{-r})$. Now using results for strongly mixing stationary sequences (Philipp [33, p. 254]) we can construct $\{\tilde{z}_k\}$ and $\{B_k\}$ satisfying (6). When τ is aperiodic and $E[\tau^r] < \infty$, we can construct on a probability space sequences $\{z'_k\}$ and $\{\tilde{z}'_k\}$ with $\{z'_k\} \overset{\text{dist}}{=} \{z_k\}$, $\{\tilde{z}'_k\} \overset{\text{dist}}{=} \{\tilde{z}_k\}$ and$\{z'_k\}$ and $\{\tilde{z}'_k\}$ have a coupling epoch $\hat{\tau}$ with $E[\hat{\tau}^r] < \infty$ (Kalashnikov [23], Chapter 6). Then for any $\epsilon > 0$, denoting

$$A_n = \left\{ \sup_s \frac{|\sum_{k=1}^n 1\{z_k' \le s\} - \sum_{k=1}^n 1\{\tilde{z}_k' \le s\}|}{n^{1/2}(\log n)^{-\lambda}} > \epsilon \right\},$$

$$P\left[\bigcup_{n \ge N} A_n\right] = P\left[\bigcup_{n \ge N} A_n; \hat{\tau} \le m\right] + P\left[\bigcup_{n \ge N} A_n; \hat{\tau} > m\right].$$

Taking $m = [\epsilon N^{1/2}(\log N)^{-\lambda}]$, $P[A_n; \hat{\tau} \le m] = 0$ for each $n \ge N$ and hence

$$P\left[\bigcup_{n \ge N} A_n\right] \le P[\hat{\tau} > m] \le E[(\hat{\tau})^r]/m^r \to 0$$

as $N \to \infty$. Thus we have

$$\sup_s \left| \sum_{k=1}^n 1\{z_k' \le s\} - \sum_{k=1}^n 1\{\tilde{z}_k' \le s\} \right| / n^{1/2}(\log n)^{-\lambda} \to 0 \quad \text{a.s.}$$

Now since $\{z_k'\}$, $\{\tilde{z}_k'\}$ and $\{B_k\}$ are r.v.s. with values in Polish spaces, we can simultaneously define $(\{z_k''\}, \{\tilde{z}_k''\}, \{B_k''\})$ on a probability space such that $(\{z_k''\}, \{\tilde{z}_k''\}) \stackrel{\text{dist}}{=} (\{z_k', \tilde{z}_k'\})$ and $(\{\tilde{z}_k''\}, \{B_k''\}) \stackrel{\text{dist}}{=} (\{\tilde{z}_k\}, \{B_k\})$ and hence this construction satisfies (6) and (7). □

The above result extends to the case when $\{z_k\}$ are R^d valued if we take $r > 3 + d$ for F continuous and $r > 4 + 2d$ for the general case in the above theorem (see Philipp [33]). Also, from Phillip [33, p. 258], we note that Theorem 4 remains valid if in (6) and (7) we replace the empirical process by quantile process, which at times can be more interesting to estimate than F.

The results of Theorems 1–4 indicate that for estimating distributions of a and s it is not sufficient that $a^{(n)} \xrightarrow{w} a$ and $s^{(n)} \xrightarrow{w} s$ but rather certain moments of s should also converge. Then we can not only estimate Ew^α and π accurately but the rates of convergence will also be comparable to the case when the distributions of a and s were available. There are more demands on the estimates of s than on that of a which in fact is fortunate because in many practical systems, such as all communication networks designed – Arpanet, Ethernet, ATM etc., the packet lengths (and hence service times) are bounded and hence all the moment conditions are satisfied. In the general case if we take the empirical distribution function of s then all the assumptions are satisfied. But if we take for example, minimum distance estimates from a family of distributions (e.g. a parametric family) where the distance chosen is Prohorov or Kolmogorov–Smirnov distance, then the moments of the estimates will not converge and the estimates $Ew^{(n)}$ and $\pi^{(n)}$ obtained can be arbitrarily bad (see for instance Donoho and Liu [14]).

3. Estimation with contaminated IID sample

Usually an available sample of a and s is contaminated with outliers or round off errors in the data (see Huber [20] and Hampel et al. [18]). The amount of contamination can be quantified by the Prohorov metric d_p defined as

$$d_p(P, Q) = \inf\{\epsilon > 0 \,|\, P(F) \leq Q(F^\epsilon) + \epsilon \text{ for all } F \text{ closed sets}\}$$

where P and Q are probability measures on a Polish space (X, d) and

$$F^\epsilon = \{x \in X \,|\, d(x, F) < \epsilon\}.$$

The round-off errors in the data are taken care of by F^ϵ and the outliers by $+\epsilon$ term in the definition of d_p. One also notices that even a small amount of contamination by outliers can make the moments of the distribution arbitrarily large. Thus the Prohorov distance d_p (which metrizes weak convergence) is ideally suited to model contamination by outliers and round off errors but any other metric which also implies convergence of some moments is entirely unsuitable (for more motivation and discussion see Hampel et al. [18] and Huber [20]).

This is in direct contradiction to the requirements needed in Theorems 1–4 because with contaminated data Es cannot be accurately estimated even when the contamination is small. Therefore, in this section we suggest some ways to take care of this very real problem. Theorem 5 and Lemma 4 of this section are of interest for the general robust estimation problem also.

Let F_a and F_s denote the distributions of a and s respectively while let the contaminated sample of a and s be iid with distributions G_a and G_s. The most basic requirement for a $GI/GI/1$ queue is the stability condition $Ea > Es$. Now we suggest how we can check this stability condition and other conditions of section 2.

In many realistic situations, the service times are upper bounded by a known constant. Then the contaminated sample of service times can be truncated at that upper bound. Thus if $d_p(G_s, F_s)$ is small (G_s denoting the distribution of truncated contaminated service times), then moments of G_s are close to that of F_s. Let $a_1, a_2, \ldots$ be an iid sample with the distribution G_a. We estimate $E[min\,(c, a)]$ for a suitably large c and compare with the estimate of Es. Other conditions of section 2 are also satisfied.

Next we consider the unbounded service times. If s has an exponential distribution with its rate μ unknown then an MLE (maximum likelihood estimator) from an uncontaminated sample satisfies our requirements. For a contaminated sample one may use M-estimators in Huber [20] which unlike MLE are d_p-continuous.

For s having an Erlang distribution, a similar solution suffices (now Theorem 5 below provides d_p-continuity). Further assume that the distribution of s is from the family of (finite) mixtures of Erlang distributions. Estimation of parameters in mixture families has been extensively studied; for a review see Redner and Walker [35], who however, do not discuss mixtures of Erlang distributions. But we can apply the general results of Lebroux [28] to the family of finite mixtures of Erlang distributions

$$\sum_{i=1}^{k} \alpha_i \delta^{n(i)} \frac{x^{n(i)-1}}{(n(i)-1)!} e^{-\delta x},$$

where k is a finite known constant and the parameters to be estimated are $0 < \delta < \infty$, $0 \leq \alpha_i \leq 1$, $\sum_{i=1}^{k} \alpha_i = 1$, and $0 < n(i) < \infty$, $i = 1, \ldots, k$ where $n(i)$ are integers. If we upper bound $n(i)$, then the conditions, in Lebroux [28] are satisfied and we obtain (without constraining k to be finite) the existence and a.s. convergence of MLE and maximum penalized likelihood estimators even if we upriori arbitrarily restrict the value of k.

These estimators are not robust. For robustness we may use the M-estimators obtained by solving $\inf \sum \phi(X_i, \theta)$ or $\sum \psi(X_i, \theta) = 0$ where ϕ and ψ are appropriate bounded continuous functions, $\{X_i\}$ is the sample and θ is the parameter to be estimated. But the consistency, asymptotic normality and continuity conditions for M-estimators provided by Huber [20] or Hampel et al. [18] do not cover this case. Therefore, we now show consistency and d_p continuity in greater generality (asymptotic normality for Markov $\{X_i\}$ is provided in the next section).

For the next theorem we need the following notation and definitions. Let Θ be the parameter space which is a closed subset of a metric space with metric d. Let P_θ be the distribution on the sample space X (considered a separable complete metric space) corresponding to parameter $\theta \in \Theta$. Let $\mathcal{P} = \{P_\theta, \theta \in \Theta\}$, and let $\mathcal{M}$ denote the family of all probability distributions on X.

Let $\phi: X \times \Theta \to R$ and $\psi: X \times \Theta \to X$ be measurable functions. We denote

$$f(\theta, P) = \int \phi(x, \theta)\, dP, \quad g(\theta, P) = \int \psi(x, \theta)\, dP$$

$$A(P) = \{\theta_0: f(\theta_0, P) = \inf_\theta f(\theta, P)\},$$

$$A_\epsilon(P) = \{\theta_0: f(\theta_0, P) < \inf_\theta f(\theta, P) + \epsilon\},$$

$$Z(P) = \{\theta: g(\theta, P) = 0\},$$

$$Z_\epsilon(P) = \{\theta: \|g(\theta, P)\| < \epsilon\},$$

where $\epsilon > 0$ and for $Z_\epsilon(P)$ we assume X to be a Banach space with norm $\|\cdot\|$. We use the following definitions from Bank et al. [4].

DEFINITION

A point to set valued map $P \mapsto B(P)$ is called usc-B if for any open set $U \supset B(P)$ there exists a $\delta > 0$ such that $B(P') \subset U$ for all P' with $d_p(P', P) < \delta$. □

DEFINITION

A point to set valued map $P \mapsto B(P)$ is called lsc-B if for each open set U s.t. $U \cap B(P) \neq \phi$, there exists a $\delta > 0$ such that $U \cap B(P') \neq \phi$ for all P' with $d_p(P, P') < \delta$. □

If a set-valued map is actually point-valued (i.e. image of each point is a singleton), then both the above definitions are equivalent to continuity of the function. Usc-B and lsc-B of the above sets of optimal points and zero sets will imply the qualitative robustness of the corresponding estimators.

THEOREM 5

Let Θ be a closed subset of a metric space.

(a) Let $\int \phi(x, \theta_n)\, dP_n \to \int \phi(x, \theta)\, dP$ whenever $\theta_n \to \theta$ and $P_n \xrightarrow{w} P$. Then
  (i) if $\hat{\theta}_n \in A(P_n)$, $\hat{\theta}_n \to \theta_0$ and $P_n \xrightarrow{w} P$, then $\theta_0 \in A(P)$.
  (ii) if $P_n \xrightarrow{w} P$ and there is a compact set $K \subset \Theta$ such that $A(P_n) \cap K \neq \phi$ then $\inf_\theta \int \phi(x, \theta)\, dP_n \to \inf_\theta \int \phi(x, \theta)\, dP$.
  (iii) $A(P)$ is usc-B if $A(Q) \subset K$ for all $Q \in \mathcal{M}$, for some compact K.
  (iv) $A_\epsilon(P)$, for any ϵ, is lsc-B under conditions of part (ii).

(b) If $\| \int \psi(x, \theta_n)\, dP_n - \int \psi(x, \theta)\, dP \| \to 0$ whenever $\theta_n \to \theta$ and $P_n \xrightarrow{w} P$, then (i)–(iv) hold under the same conditions with $A(P)$, $A(P_n)$ and $A_\epsilon(P)$ replaced by $Z(P)$, $Z(P_n)$ and $Z_\epsilon(P)$ respectively.

Proof

(a) Suppose $\theta_0 \notin A(P)$. Then there is an $\epsilon > 0$ s.t. $\int \phi(x, \theta_0)\, dP > \inf_\theta \int \phi(x, \theta)\, dP + \epsilon$. Let $\theta' \in \Theta$ such that $\int \phi(x, \theta')\, dP < \inf_\theta \int \phi(x, \theta)\, dP + \epsilon/2$. But $\int \phi(x, \theta')\, dP_n \to \int \phi(x, \theta')\, dP$ and $\int \phi(x, \hat{\theta}_n)\, dP_n \to \int \phi(x, \theta_0)\, dP_n > \int \phi(x, \theta'))\, dP_n$ which is a contradiction because $\hat{\theta}_n \in A(P_n)$. Parts (ii) and (iii) follow now from theorem 4.2.1 and (iv) follows from theorem 4.2.4 of Bank et al. [4].

(b) Notice that if $Z(P)$ is nonempty then $Z(P) = \operatorname{argmin}_\theta \| \int \psi(x, \theta)\, dP \|$. Also $\| \int \psi(x, \theta_n) dP_n \| \to \| \int \psi(x, \theta) dP \|$ whenever $\| \int \psi(x, \theta_n) dP_n - \int \psi(x, \theta) dP \| \to 0$. Now use part (a). □

A sufficient set of conditions for $\int \phi(x, \theta_n)\, dP_n \to \int \phi(x, \theta)\, dP$ whenever $P_n \xrightarrow{w} P$ and $\theta_n \to \theta$ is

(i) $\phi(x, \theta_n) \to \phi(x, \theta)$ uniformly (for x) over compact sets and

(ii) there exist measurable functions h_n and h such that $h_n(x) \to h(x)$ uniformly over compact sets, $|\phi(x, \theta_n)| \le h_n(x)$ for all x and $\int h_n \, dP_n \to \int h \, dP$.

If we replace $P_n \xrightarrow{w} P$ by the somewhat stronger condition: $P_n(E) \to P(E)$ for all measurable sets E then the condition of uniform convergence over compact sets in the above assumptions can be replaced by pointwise convergence (see Serfozo [36]).

Results related to Theorem 5 have been recently obtained by Wets [42] and King and Rockafellar [26]. They assume X to be finite dimensional and do not study $A_\epsilon(P)$ and $Z_\epsilon(P)$ (see next paragraph). They also obtain asymptotic distributions for iid sample which are applicable to our mixtures of Erlang model. Under the conditions of Wets [42], we can show that Theorem 5(a) (iii), (iv) also hold if we further assume the existence of a compact set K mentioned in Theorem 5(a) (iii) and (iv). If we assume that $\int \phi(x, \theta) \, dP$ is a quasiconvex function of θ and $\Theta \subseteq R^n$ is convex and closed then we do not require the existence of compact K in Theorem 5.

As against $A_\epsilon(P)$ and $Z_\epsilon(P)$, lsc-B of $A(P)$ and $Z(P)$ holds under rather stringent conditions. One of the advantages of lsc-B over usc-B is that if in addition, $A_\epsilon(P)$ and $Z_\epsilon(P)$ are convex for each P (a sufficient condition for this is that ϕ and ψ are convex functions of θ), Θ is a subset of a Banach space and $\mathcal{M}$ is a countable union of compact sets then we can have a continuous selection of $A_\epsilon(P)$ and $Z_\epsilon(P)$ (see Aubin [3, p. 355]). Using a continuous selection in the estimation algorithm, we can obtain unique robust estimates.

In addition to asymptotic robustness, Theorem 5 provides the consistency of the M-estimators under general conditions. Let $X_1, X_2, \ldots, X_n$ be a stationary ergodic sample with marginal distribution π and let P_n be the empirical measure obtained from this sample. Then $P_n \xrightarrow{w} \pi$ a.s. and hence results of Theorem 5 hold a.s. (measurability conditions can be taken care of as in the above references). In addition, the following holds. Let $\hat{P}_n$ be the empirical measure obtained from an iid sample of size n with distribution P_n and let $P_n \xrightarrow{w} P$ then $d_p(\hat{P}_n, P_n) \to 0$ in probability (see Beran et al. [9]) and hence $d_p(\hat{P}_n, P) \to 0$ in probability and our results still apply.

If in addition to the assumptions of Theorem 5, $A(P_\theta) = \{\theta\}$ and $Z(P_\theta) = \{\theta\}$ for all $\theta \in \Theta$ (Fisher consistency) then $A(P_n) \to \{\theta\}$ and $Z(P_n) \to \{\theta\}$ and the estimators are strongly consistent. But it would be nice if $A_\epsilon(P)$ and $Z_\epsilon(P)$ are in a nbd of θ when $d_p(P, P_\theta) < \epsilon$ for ϵ small and the Fisher consistency is satisfied. Lsc-B of $A_\epsilon(P)$ and $Z_\epsilon(P)$ guarantees that in any δ-nbd of θ, $A_\epsilon(P)$ and $Z_\epsilon(P)$ will have a point. If we further, assume that for any $\delta > 0$ there exist an $\epsilon_1 > 0$ s.t. $d(\theta, \theta_0) > \delta$ implies that $|f(\theta, P_{\theta_0})| > \epsilon_1 (g(\theta, P_{\theta_0}) > \epsilon_1)$ then we can ensure that $A_\epsilon(P)(Z_\epsilon(P))$ will be in a small nbd of θ_0 when ϵ and $d_p(P, P_{\theta_0})$ are small enough.

Consider a further generalization. Suppose s has a general distribution F_s on R^+ and let $\int x^m \, dF_s(x) < \infty$ for some $m > 0$. Then there are distributions $F^{(n)}$ in the family of mixtures of Erlangs such that $F^{(n)} \xrightarrow{w} F_s$ and $\int x^m \, dF^{(n)} \to \int x^m \, dF_s$ (Asmussen [1]). We can also define a metric d_s such that the above convergence is equivalent to $d_s(F^{(n)}, F_s) \to 0$. Suppose for a given $\epsilon > 0$, we know a family of finite mixtures of Erlangs such that there is a G in this family with $d_s(F_s, G) < \epsilon$. Then the following lemma provides conditions for robust estimation. Let $T: \mathcal{M} \to$, $\Theta(\mathcal{M}, \Theta, \mathcal{P}$ are as in Theorem 5 corresponding to the mixtures of Erlangs) be a mapping which will denote our estimator. By $T(F)$ we will also denote a distribution corresponding to the parameter $T(F)$.

LEMMA 5

Assume that

(1) for all $\theta_n, \theta \in \Theta$, $d_s(P_{\theta_n}, P_\theta) \to 0$ if $d(\theta_n, \theta) \to 0$,
(2) (robustness) $d(T(P_n), T(P)) \to 0$ if $d_p(P_n, P) \to 0$ and
(3) if $d_s(F, \mathcal{P}) < \epsilon$ then $d_s(F, T(F)) < \epsilon + \epsilon_2$, for some $\epsilon_2 > 0$.

Let $F \in \mathcal{M}$ be such that $d_s(F, \mathcal{P}) < \epsilon$, then for any $\epsilon_1 > \epsilon_2$ there exists a $\delta > 0$ such that $G \in \mathcal{M}$, $d_p(F, G) < \delta$ implies that $d_s(F, T(G)) < \epsilon + \epsilon_1$.

Proof

From assumption (3),

$$d_s(F, T(G)) \le d_s(F, T(F)) + d_s(T(F), T(G))$$
$$\le \epsilon + \epsilon_2 + d_s(T(F), T(G)).$$

From assumption (1), given $\epsilon_1 > 0$ we can find $\epsilon_3 > 0$ s.t. if $d(T(F), T(G)) < \epsilon_3$ then $d_s(T(F), T(G)) < \epsilon_1 - \epsilon_2$. Again by assumption (2) we can find a $\delta > 0$ s.t. if $d_p(F, G) < \delta$ then $d(T(F), T(G)) < \epsilon_3$ and hence we obtain the result. □

The class of minimum distance (w.r.t. d_s but not w.r.t. d_p) estimators (see Donoho and Lin [13] for robustness of these estimators) satisfies (3) in Lemma 5. Also, as seen from Theorem 5, the M-estimators satisfy (2) and (3) if $\mathcal{P}$ is taken as a finite family of mixtures of Erlangs mentioned above. This family also satisfies (1).

Now suppose we have a iid sample of size n with distribution G. Let $G^{(n)}$ be the empirical distribution. Then $\sup_x |G^{(n)}(x) - G(x)| \to 0$ a.s. and hence if $d_p(F_s, G) < \delta$ then $d_s(T(G^{(n)}), F_s) < \epsilon + \epsilon_1$ a.s. for n large enough.

Since F_s is unknown, for a given family $\mathcal{P}$, $d_s(F_s, \mathcal{P})$ is not known. But choosing d_s such that $d_s(F, G) \ge d_p(F, G)$ for all F, G (see for example page 76, Asmussen [1]), by (3) in Lemma 5,

$$\begin{aligned} d_s((F, \mathcal{P}) &\ge d_s(F, T(F)) - \epsilon_2 \\ &\ge d_s(F, T(G)) - d_s(T(G), T(F)) - \epsilon_2 \\ &\ge d_p(F, T(G)) - d_s(T(G), T(F)) - \epsilon_2 \\ &\ge d_p(G, T(G)) - d_p(F, G) - d_s(T(G), T(F)) - \epsilon_2. \end{aligned}$$

Therefore, by (1) and (2) in Lemma 5, if $d_p(F, G)$ is small, then a large $d_p(G, T(G))$ will imply that $d_s(F, \mathcal{P})$ is large and hence the class $\mathcal{P}$ does not approximate F well. Then we will have to enlarge $\mathcal{P}$. Of course it is also possible that $d_s(F_s, \mathcal{P})$ is large but $d_p(G, T(G))$ and $d_p(F_s, G)$ are small in which case we can consider estimating the parameters in the full family of mixtures of Erlang distributions. This is an infinite dimensional non locally compact parameter space. Then Theorem 5 can still be applied but the important assumption (1) in Lemma 5 may not be satisfied.

4. Estimation from contaminated regenerative processes

This section is concerned about estimation of parameters when instead of an iid sample of a and s, a contaminated sample of waiting times is available. We consider it in the general setting of a regenerative sample.

Let $\{X_k\}$, $\{Y_k\}$, $\{Z_k\}$ and $\{R_k\}$ be jointly stationary ergodic real valued (or regenerative with finite mean regeneration period) processes where Z_k is $\{0, 1\}$ valued and let

$$Y_k = X_k(1 - Z_k) + Z_k R_k.$$

We want to estimate the stationary distribution, stationary moments or certain parameters of $\{X_k\}$ based on the observation of $\{Y_k\}$. In this model $\{R_k\}$ represents the outliers. For motivation, other examples and results for this model, see Martin and Yohai [31]. Let $P[Z_k = 1] = \epsilon$. Let us denote the stationary measure of the various processes with $P_\pi[\ldots]$. Then estimating $P_\pi(X_1 = i)$ by $n^{-1}\sum_{k=1}^n 1\{Y_k = i\}$ we obtain by strong law of large numbers (SLLN),

$$\frac{1}{n}\sum_{k=1}^n 1\{Y_k = i\} \to P_\pi[Y_k = i] = P_\pi[X_k = i, Z_k = 0] + P_\pi[R_k = i, Z_k = 1] \quad \text{a.s.}$$

and hence

$$\sum_i |P_\pi[Y_k = i] - P_\pi[X_k = i]| \le \epsilon, \tag{8}$$

which is the robustness property of the estimator. Similarly, the influence functional defined in Martin and Yohai [31], can be shown to be bounded for the above estimator. This result of course holds if the state space of the processes is a Polish space.

Now if $\{X_k\}$ is also a countable state Markov chain then denoting by $P(i, j)$ the transition matrix, the maximum likelihood estimate (MLE) of $P(i, j)$ from a sample of $\{X_k\}$ is given by $\sum_{k=1}^{n-1} 1\{X_k = i, X_{k+1} = j\}/n_i$ where $n_i = \sum_{k=1}^n 1\{X_k = i\}$. For a contaminated sample $\{Y_k\}$

$$\left(\sum_{k=1}^n 1\{y_k = i\}\right)^{-1} \sum_{k=1}^{n-1} 1\{Y_k = i, Y_{k+1} = j\} \to P_\pi[Y_2 = j, Y_1 = i]/P_\pi[Y_1 = i].$$

Also, we can easily show that

$$\sum_{i,j} |P_\pi[Y_2 = j, Y_1 = i] - P_\pi[X_2 = j, X_1 = i]| \leq 2\epsilon. \tag{9}$$

From (8) and (9) we obtain the continuity of the estimate of $P(i, j)$ and an explicit upper bound. Similarly we can show that the influence functional of the estimator is bounded. These estimates hold for Polish space valued Markov chains and imply robustness. Of course, even for arbitrarily small ϵ, these results do not guarantee the ergodicity or positive recurrence of chain for the estimated P in general unless P satisfies uniform ergodicity (Kartashov [25]).

Of course as in the iid case, sample mean of $\{Y_k\}$ as estimator of $E_\pi X_1$ is not robust but one can resort to other estimators (e.g. M-estimators) which are continuous functions of above perturbations in the parametric situation (use Theorem 5).

The asymptotic normality (and other asymptotic behaviour) of M-estimators in Wets [42] and King and Rockafellar [26] is obtained for an iid sample while Basawa [5] obtains it for time series. We use the method of Basawa [5] (which is similar to the iid case) to obtain asymptotic normality of M-estimators for Markov chains. Consider a sample $X_i, \ldots, X_n$ of a countable state Markov chain with transition probability $P(i, j, \theta)$, $\theta \in \Theta$ where Θ, an open subset of R^m, is a parameter set from which the parameter has to be estimated. We assume $P(\theta) = (P(i, j, \theta))$ to be ergodic, aperiodic and irreducible for each $\theta \in \Theta$ with regeneration length $\tau(\theta)$. The M-estimator θ_n is obtained as a solution of equations $S_n(\theta) = \sum_{k=1}^{n-1} \psi(P(X_k, X_{k+1}; \theta)) = 0$ where the true parameter for the sample $X_1, \ldots, X_n$ is θ_0 and the corresponding stationary distribution is π. Assume $\psi(x, \theta)$ has continuous partial derivatives w.r.t. θ for almost all x. Then under various sufficient conditions we can show that for θ_n close to θ_0 (this holds if $\{\theta_n\}$ forms a consistent sequence of estimators and n is large enough)

$$n^{-1/2}(S_n(\theta_n) - S_n(\theta_0)) = n^{-1/2}(S_n'(\theta_0)(\theta_n - \theta_0)) + o_p(1) \tag{10}$$

where $S_n(\theta_n) = 0$ and S_n' denotes the matrix of derivatives of S_n w.r.t. θ. If $\|\theta_n - \theta_0\| \xrightarrow{p} 0$, $n^{-1/2} S_n(\theta_0) \xrightarrow{w} N(0, I(\theta_0))$ and $S_n'(\theta_0)/n \xrightarrow{p} J(\theta_0)$ where $I(\theta_0)$ and $J(\theta_0)$ are deterministic matrices, $I(\theta_0)$ symmetric positive definite and $J(\theta_0)$ nonsingular then $(\theta_n - \theta_0)n^{1/2} \xrightarrow{w} N(0, (J(\theta_0)^{-1})^T I(\theta_0) J(\theta_0)^{-1})$.

Now we obtain sufficient conditions for the above asymptotic normality result. Notice that $\{(x_k, x_{k+1}), k \geq 0\}$ is also an irreducible, aperiodic ergodic Markov chain on the state space $\{(i, j)\}: P(i, j, \theta) \neq 0\}$ (assumed independent of θ) with regeneration length $\tau(\theta)$. Then $\|\theta_n - \theta_0\| \xrightarrow{p} 0$ is obtained from Theorem 5 by taking $\psi((x_k, x_{k+1}), \theta) = \psi(P(x_k, x_{k+1}, \theta))$. Also, if

$$E_\pi\left[\left|\frac{\partial \psi_i(P(x_1, x_2, \theta))}{\partial \theta_j}\right|\right] < \infty, \quad \psi = (\psi_1, \psi_2, \ldots, \psi_m), \quad \theta = (\theta_1, \ldots, \theta_m),$$

then SLLN provides $S_n'(\theta_0)/n \xrightarrow{p} J(\theta_0)$ and if $E[(\tau(\theta_0))^4] < \infty$ and $E_\pi[(\psi_i(P(x_1, x_2, \theta))^4] < \infty$, $i = 1, \ldots, m$ we obtain $n^{-1/2}S_n(\theta_0) \xrightarrow{w} N(0, I(\theta_0))$ (for $J(\theta_0)$ and $I(\theta_0)$ appropriately defined) (Sharma [38]).

One sufficient condition for $o_p(1)$ in (10) is that S_n has continuous second order partial derivatives S_n'' and $n^{-1/2}S_n''(\theta)$ are bounded in probability in a neighbourhood of θ_0 (or have a uniformly bounded (in a nbd of θ_0) moment).

The above conclusions hold if $\{X_k\}$ is Harris ergodic with general state space. The results can be extended to a stationary process $\{X_k\}$ satisfying some mixing conditions.

Finally we consider continuous time processes $\{Y_t\}$, $\{X_t\}$, $\{Z_t\}$ and $\{R_t\}$. Under the above mentioned assumptions we can again show the robustness of $\int_0^t 1\{Y_t = i\}/t$ for $P_\pi\{X_t = i\}$ with the error bound (8).

If $\{X_t\}$ is a countable state Markov process with infinitesimal generator $Q = (q(i,j))$ then given observations $\{X_t, t, \leq T\}$, MLE of $q(i,j)$ is (see Basawa and Prakasa Rao [7]) $N_T(i,j)/A_T(i)$ where $N_T(i,j)$ is the number of transitions of $\{X_t\}$ from i to j in time $[0, T]$ and $A_T(i) = \int_0^T 1\{X_t = i\}\,dt$. For finite state processes this MLE is shown to be strongly consistent and asymptotically normal in the above reference. But the following example shows that it is not a robust estimator even for a finite state Markov process.

Example

Let $\{X_t\}$ be a Markov process with state space $\{0, 1\}$ and Q matrix $\begin{bmatrix} -1 & 1 \\ 1 & -1 \end{bmatrix}$. Also let $\{Z_t\}$ and $\{R_t\}$ be independent of $\{X_t\}$ and defined on $\Omega = [0, 1]$ as follows. $Z_t(0) = 0$ when $t \in [k, k+1]$ where $k = 0, 2, \ldots,$ and $Z_t(0)$ alternates between 0, 1, N times in the time intervals $[k, k+1]$, $k = 1, 3, \ldots,$. Also, $Z_t(w) = Z_{t+w}(0)$ where $w \in [0, 1]$ and $R_t \equiv 1$. Then $N_T(0,1)/A_T(0) \to 1$ a.s. Let $\tilde{N}_T(0,1)$ and $\tilde{A}_T(0)$ denote the corresponding quantities for $\{Y_t\}$ while $N'(t)$ denotes the number of arrivals in time t in a Poisson process with rate 1. Then $|\tilde{N}_T(0,1) - (N'(T/4 + (1-2\epsilon)/2)) + NT/4)|/T \to 0$ a.s. and hence $\tilde{N}_T(0,1)/\tilde{A}_T(0) \to (1 + N/4)$ a.s. where $P_\pi[Z_t = 1] = \epsilon$. Since N can be taken arbitrarily large, for any $\epsilon > 0$, $\tilde{N}_T(0,1)/\tilde{A}_T(0)$ can provide arbitrarily bad estimate. □

Thus for a contaminated process, we suggest the following procedure. Consider the observations of $\{Y_t\}$ at discrete steps kh, $k = 0, 1, 2, \ldots,$ where $h > 0$ is fixed. Since $\{X_{kh}\}$ is a Markov process with transition probability matrix (say) P^h, denoting its estimators from n observations of $\{Y_k\}$ as $\tilde{P}^{h,n}$ from (8) and (9) we obtain a robust estimate of $P^h(i,j)$ with explicit bounds. We also know that $|P^h_{[t/h]}(i,j) - P_t(i,j)| \to 0$ as $h \to 0$, uniformly over t in compact sets, where $\{P_t\}$ is the transition probability function of $\{X_t\}$, $[x]$ is integer part of x and P_n^h is n step transition probability matrix of $\{X_{kh}\}$. Also if $\{\tilde{X}^h(t)\}$ is a process defined from $\{X_{kh}\}$ as $\tilde{X}^h(t) = X_{kh}$ where $k \leq t/h < k+1$, and $\{X_t\}$ has a finite state space then we can form versions of $\tilde{X}^h(t)$ and $\{X_t\}$ with $\{\tilde{X}_t^h\} \xrightarrow{w} \{X_t\}$ in $D[0,\infty)$ (see section 5, Van Dijk [41]). In addition, if $Q^h = (q^h(i,j))$ is a generator corresponding to $\{\tilde{X}_t^h\}$ then $q^h(i,j) \to q(i,j)$ for all (i,j) as $h \to 0$ (Ethier and Kurtz [15, Chapter 1, Theorem 6.1]). Thus we can obtain a robust estimate of $q(i,j)$.

Acknowledgements

Professor Rakesh Bansal's comments on robustness issues are gratefully acknowledged. I am also thankful to the referee for mentioning refs. [23] and [24] and to Dr. Whitt for promptly sending several useful references. Professor Prabhu's editorial comments have improved the presentation. This work is financially supported by the Government of India MHRD Telematics project.

References

[1] S. Asmussen, *Applied Probability and Queues* (Wiley, New York, 1987).

[2] S. Asmussen, Queueing simulation in heavy traffic, Math. Oper. Res. 17 (1992) 84–111.

[3] J.-P. Aubin, *Set Valued Analysis* (Birkhauser, Boston, 1990).

[4] B. Bank, J. Guddat, D. Klatte, B. Kummer and K. Tammer, *Nonlinear Parametric Optimization* (Academic Verlag, Berlin, 1982).

[5] I.V. Basawa, Large sample statistics for stochastic processes: some recent developments, in: *Probability, Statistics and Design of Experiments, Proc. of R.C. Bose Symp.* (Wiley Eastern, New Delhi, 1988).

[6] I.V. Basawa and N.U. Prabhu, Large sample inference from single server queues, Queueing Systems 3 (1988) 298–304.

[7] I.V. Basawa and B.L.S. Prakasa Rao, *Statistical Inference for Stochastic Processes* (Academic Press, New York, 1980).

[8] S.J. Bean and C.P. Tsokos, Developments in nonparametric density estimation, Int. Stat. Rev. 48 (1980) 267–287.

[9] R.J. Beran, L. LeCam and P.W. Millar, Convergence of stochastic empirical measures, J. Multivariate Anal. 23 (1987) 159–168.

[10] U.N. Bhat and S.S. Rao, Statistical analysis of queueing systems, Queueing Systems 1 (1987) 217–247.

[11] M. Czorgo, P. Deheuvels and L. Horvath, An application of stopped sums with applications in queueing theory, Adv. Appl. Prob. 19 (1987) 674–690.

[12] M. Czorgo and Revesz, *Strong Approximations in Probability and Statistics* (Academic Press, New York, 1981).

[13] D.L. Donoho and R.C. Liu, The automatic robustness of minimum distance estimators, Ann. Stat. 16 (1988) 552–586.

[14] D.L. Donoho and R.C. Liu, Pathologies of minimum distance estimators, Ann. Stat. 16 (1988) 587–608.

[15] S.N. Ethier and T.G. Kurtz, *Markov Processes, Characterisation and Convergence* (Wiley, New York, 1986).

[16] P.W. Glynn and W. Whitt, The asymptotic efficiency of simulation estimators, Oper. Res. 40 (1992) 505–520.

[17] A. Gut, *Stopped Random Walks, Limit Theorems and Applications* (Springer, New York, 1988).

[18] F.R. Hampel, P.J. Rousseeuw, E.M. Ronchetti and W.A. Stahel, *Robust Statistics* (Wiley, New York, 1986).
[19] C.R. Heathcote, Complete exponential convergence and some related topics, J. Appl. Prob. 4 (1967) 217–256.
[20] P.J. Huber, *Robust Statistics* (Wiley, New York, 1981).
[21] S. Janson, Renewal theory for *m*-dependent variables, Ann. Prob. 11 (1983) 558–568.
[22] V.V. Kalashnikov, Assessing the sensitivity of queueing system, Avtomat. i Telemekh. (May 1981).
[23] V.V. Kalashnikov, *Mathematical Methods in Queueing Theory* (Kluwer, Dordrecht, 1994).
[24] V.V. Kalashnikov and Rachev, *Mathematical Methods for Construction of Queueing Models* (Wadsworth and Brooks, Pacific Grove, 1990).
[25] N.V. Kartashov, Inequalities in theorems of ergodicity and stability for Markov chains with common phase space I, Theory Prob. Appl. 30 (1985) 247–259.
[26] A.J. King and R.T. Rockafellar, Asymptotic theory for solutions in statistical estimation and stochastic progress, Math. Oper. Res. 18 (1993) 148–162.
[27] S.S. Lavenberg and P.D. Welch, A perspective on the use of control variables to increase the efficiency of Monto Carlo simulations, Manag. Sci. 28 (1981) 322–335.
[28] Lebroux, Consistent estimation of a mixing distribution, Ann. Stat. 20 (1992) 1350–1360.
[29] E.L. Lehmann, *Theory of Point Estimation* (Wiley, New York, 1983).
[30] R. Lepage and L. Billard, *Exploring Limits of Boot Strap* (Wiley, New York, 1992).
[31] R.D. Martin and V.J. Yohai, Influence functionals for time series, Ann. Stat. 14 (1986) 781–818.
[32] K. Pawlikowski, Steady-state simulation of queueing processes: a survey of problems and solutions, ACM Comp. Surv. 22 (1990) 123–170.
[33] W. Philipp, Invariance principles for independent and weakly dependent random variables, in: *Dependence in Probability and Stat.*, eds. E. Eberlain and M.S. Taqqu (Obevolfach, 1985).
[34] B.L.S. Prakasa Rao, *Asymptotic Theory of Statistical Inference* (Wiley, New York, 1987).
[35] R. Redner and H.F. Walker, Mixture densities, maximum likelihood and the EM algorithm, SIAM Rev. 26 (1984) 195–239.
[36] R. Serfozo, Convergence of Lebesgue integrals with varying measures, Sankhya, Series A 44 (1982) 380–402.
[37] J.G. Shanthikumar, Bounds and an approximation for single server queues, J. Oper. Res. Soc. Jpn. 26 (1983) 118–134.
[38] V. Sharma, Invariance principles for regenerative and Markov processes with applications to queueing networks, Technical Report ISRO-IISc STC (1993).
[39] D. Thiruvaiyaru and I.V. Basawa, Empirical Bayes estimation for queueing systems and networks, Queueing Systems 11 (1992) 179–202.
[40] J. Tiefeng, Large deviations for renewal processes, Stoch. Proc. Appl. 50 (1994) 57–71.
[41] N.M. Van Dijk, Controlled Markov processes, time discretization, CWI Tracts (1984).
[42] J.B. Wets Roger, Constrained estimation: consistency and asymptotics, Appl. Stoch. Meth. Data Anal. 7 (1991) 17–32.
[43] W. Whitt, Planning queueing simulations, Manag. Sci. 35 (1989) 1341–1365.
[44] W. Whitt, Asymptotic formulas for Markov processes with applications to simulation, Oper. Res. 40 (1992) 279–291.