Bernoulli **12**(4), 2006, 737–759

Nonparametric inference from the $M/G/1$ workload

MARTIN B. HANSEN[1] and SUSAN M. PITTS[2]

[1]*Department of Mathematical Sciences, Aalborg University, Fredrik Bajers Vej 7G, DK-9220 Aalborg Ø, Denmark. E-mail: mbh@math.aau.dk*
[2]*Statistical Laboratory, Centre for Mathematical Sciences, Wilberforce Road, Cambridge CB3 0WB, UK. E-mail: S.Pitts@statslab.cam.ac.uk*

Consider an $M/G/1$ queue with unknown service-time distribution and unknown traffic intensity ρ. Given systematically sampled observations of the workload, we construct estimators of ρ and of the service-time distribution function, and we study asymptotic properties of these estimators.

Keywords: asymptotic normality; empirical processes; functional central limit theorem; infinite-dimensional delta method; $M/G/1$ queue; regenerative processes; workload

1. Introduction

Throughout this paper we consider an $M/G/1$ queue, where the inter-arrival times are independently and exponentially distributed with mean $1/\lambda$, the service times are independently non-negative, not concentrated at 0 and otherwise generally distributed with distribution function F and finite mean f_1, there is one server and an infinite waiting room. We will assume starting in an empty queue.

For later purposes let $\{(T_n, S_n), n \geqslant 1\}$, denote the sequence of arrival times and service times of the customers. Let S be a generic random variable with the same distribution as S_1.

Let the workload in the system at time t be denoted by V_t, i.e. V_t is the sum of the residual service times of the customer currently being served and the customers awaiting service and is zero if there are no customers in the system. By convention, a workload process $\{V_t, t \geqslant 0\}$ will be taken to be right-continuous with left-hand limits. For the $M/G/1$ queue, the evolution of V_t between two arrivals is described by Lindley's equation,

$$V_t = \begin{cases} 0, & t \in [0, T_1), \\ (V_{T_n-} + S_n - (t - T_n)) \wedge 0, & t \in [T_n, T_{n+1}),\ n \geqslant 1. \end{cases} \tag{1}$$

In what follows a cumulative distribution function (cdf) of a non-negative random variable is denoted by a capital letter, A, say. The kth moment is denoted by a_k, the complementary cdf by $\overline{A}(x) = 1 - A(x)$, $x \geqslant 0$, the stationary excess distribution by $A_e(x) = a_1^{-1} \int_{[0,x]} \overline{A}(y)\,\mathrm{d}y$, $x \geqslant 0$, and the Laplace–Stieltjes transform by $\tilde{A}(\theta) = \int_{[0,\infty)} \mathrm{e}^{-\theta x} A(\mathrm{d}x)$, $\theta \in \mathbb{R}$.

If we let F denote the cdf of the service-time distribution and assume the stability condition $\rho < 1$, then a limiting distribution (in weak convergence as well as total variation)

1350–7265

of V_t exists (Asmussen 2003: Corollary X.3.3). Moreover, if we let G be the cdf of the limiting distribution then the Pollaczeck–Khintchine's formula holds (Asmussen 2003: Theorem X.5.2):

$$G = \Psi(\rho, F), \tag{2}$$

where

$$\Psi(\rho, F) = (1-\rho)\sum_{k=0}^{\infty} \rho^k F_e^{\star k},$$

and $\star k$ denotes k-fold convolution.

We now assume that it is possible to test the performance of the queue by sampling the workload, without loss of generality, at every positive integer time point, as other sampling intervals can be obtained by proper rescaling. This process is denoted by $\{V_i, i \geqslant 1\}$. The main objective of this paper is to infer F_e and F from the sampled workloads. We suggest the empirical cumulative distribution function (ecdf) as an estimator for G:

$$G_n(x) = n^{-1}\sum_{i=1}^{n} 1(V_i \leqslant x), \qquad 0 \leqslant x < \infty. \tag{3}$$

Its nth *empirical process* counterpart is defined as

$$\mathbb{G}_n(x) = n^{1/2}(G_n(x) - G(x)), \qquad 0 \leqslant x < \infty. \tag{4}$$

In the following we will provide sufficient conditions for the empirical process to converge weakly to a Gaussian process.

Assume we want to make make statistical inference about the workload distribution function, then we notice that G has a $1-\rho$ atom at zero. This leads to a plug-in estimator $\hat{\rho}_n$ for ρ given by

$$\hat{\rho}_n = 1 - n^{-1}\sum_{i=1}^{n} 1(V_i = 0).$$

Later on (see (6)) we need to divide by $\hat{\rho}_n$ and $1-\hat{\rho}_n$, so we notice that with probability one $\hat{\rho}_n$ will be in (0, 1) eventually. For a discussion of the potential danger in heavy and light traffic, see Section 7.

Formula (2) can via Laplace transformation be inverted under conditions in the following way:

$$F_e(x) = \frac{1}{\rho}\sum_{k=1}^{\infty}(-1)^{k+1}\left(\frac{1}{1-\rho}\right)^k (G^o)^{\star k}(x), \tag{5}$$

where, for any function h, we write h^o for the function $x \mapsto h(x) - h(0)$ and convergence is in a suitable weighted Banach space; see Proposition 3 below. This leads to the following plug-in estimator of F_e:

$$F_{n,e}(x) = \frac{1}{\hat{\rho}_n}\sum_{k=1}^{\infty}(-1)^{k+1}\left(\frac{1}{1-\hat{\rho}_n}\right)^k (G_n^o)^{\star k}(x). \tag{6}$$

Turning to the estimation of F from $F_{n,\mathrm{e}}$, one notices, from the definition of the stationary excess distribution, that this involves some sort of numerical differentiation of a concave distribution function. One popular method for this is to take the left derivative of the concave majorant of the empirically determined distribution function, normally termed the Grenander estimator.

One motivation for the set-up presented above arises when a call admission controller (CAC) in an ATM network decides whether there are sufficient resources to allow a new connection to be established. This problem was considered by Sharma and Mazumdar (1998) and Sharma (1999). They consider an $M/G/1$ queue which is probed by Poisson traffic from which moments of the service-time distribution are inferred by the well-known moment relations which can be derived from the Pollaczeck–Khintchine formula (Asmussen 2003: Theorem VIII.5.7 (5.6) and (5.7)). In the present paper we attack the even harder problem of inferring the entire distribution function of the service time.

Another motivation arises in an infinite-capacity storage model, where inputs $S_1, S_2, \ldots$ to the storage facility arrive in a Poisson process with rate λ, where $S_1, S_2, \ldots$ are independently and identically distributed (i.i.d.) random variables with distribution function F. The total amount in the storage facility at time t has the same distribution as V_t in the $M/G/1$ queuing model above. Suppose that observations on the sampled total amounts at times $t = 1, 2, \ldots$ are available, and that interest lies in inference for the distribution F of the inputs to the facility. This inference problem is exactly analogous to that described above for the queuing model, and the methods of this paper apply.

In (5), F_{e} is given in terms of ρ and G. We observe that $\rho = 1 - G(0)$, so that F_{e} is determined by G, and we can write F_{e} as a functional of G. The proposed estimator $F_{n,\mathrm{e}}$ is then the result of applying the same functional to the estimator G_n of G. Given an appropriate asymptotic normality result for G_n, the infinite-dimensional delta method can be used to derive asymptotic normality of $F_{n,\mathrm{e}}$, provided that the functional in question satisfies a particular differentiability result. For general descriptions of the infinite-dimensional delta method, see Gill (1989) and van der Vaart (1998).

From (2), we see that G is a compound geometric distribution function, based on ρ and F_{e}, and so the inverse functional that takes G onto F_{e} is a *decompounding* functional that 'decompounds' the compound geometric distribution. As such, this functional is closely related to that in Buchmann and Grübel (2003), where the notion of decompounding is introduced in the context of decompounding for a compound Poisson distribution. The set-up and proofs for the definition and differentiability of our functional follow those in Buchmann and Grübel (2003), making adaptations for the geometric case as necessary. However, our data do not consist of i.i.d. observations, but rather exhibit regenerative structure, so that the asymptotic normality result for our input estimator is obtained using an empirical central limit theorem for regenerative data; see Tsai (1998) and Levental (1988).

The infinite-dimensional delta method has been used for stochastic models in previous work, and this paper follows the set-up and approach developed in, for example, Grübel and Pitts (1993) and elsewhere; see also Bingham and Pitts (1999a, 1999b). The latter two papers study inference for service-time distributions given data on busy and idle periods for the $M/G/1$ and the $M/G/\infty$ queues respectively, and can be regarded as tackling inverse problems, in the same way that inference for decompounding can be regarded as an inverse

problem. Hall and Park (2004) give a different approach to estimation of the density of F from busy period data.

The paper is organized as follows. In Section 2 the regenerative structure of the data is discussed in detail, together with related measurability issues. The asymptotic normality result for G_n is then stated in Section 3, and our main result, giving the asymptotic normality of the proposed estimator $F_{n,\mathrm{e}}$ of F_e, is stated in Theorem 2. Section 4 contains discussion of two pragmatic approaches for dealing with the step from F_e to the service-time distribution function F. Examples of the estimators in action are given in Section 5, and proofs related to the regenerative structure and also proofs of the asymptotic normality results are given in Section 6. Section 7 contains discussion and conclusions.

2. Preliminaries

2.1. Measurability considerations

Let $\tilde{T}_n = T_n - T_{n-1},\ n \geqslant 1$, where $T_0 = 0$. Then the input to the $M/G/1$ queue can be described as $\{(\tilde{T}_n, S_n),\ n \geqslant 1\}$ and viewed as the coordinate projection on the probability space $(\Omega, \mathcal{F}, \mathrm{P})$, where $\Omega = ([0, \infty) \times [0, \infty))^{\mathbb{N}}$, $\mathcal{F} = (\mathcal{B}[0, \infty) \times \mathcal{B}[0, \infty))^{\mathbb{N}}$ and P is the $\mathbb{N}$-fold product of the product of the exponential distribution with mean $1/\lambda$ and the distribution with cdf F.

Furthermore, by Lindley's equation (1) it is straightforward to prove that the workload process $\{V_t,\ t \geqslant 0\} : \Omega \to [0, \infty)^{[0,\infty)}$ is $(\mathcal{F}, \mathcal{B}[0, \infty)^{[0,\infty)})$-measurable, where $\mathcal{B}[0, \infty)^{[0,\infty)}$ is the Baire σ-field (Hoffmann-Jørgensen, 1994b: (9.2.4)).

Now $\{V_i,\ i \geqslant 1\} : [0, \infty)^{[0,\infty)} \to [0, \infty)^{\mathbb{N}}$ is the process $\{V_t,\ t \geqslant 0\}$ sampled at all integer points. Hence $\{V_i,\ i \geqslant 1\}$ is $(\mathcal{B}[0, \infty)^{[0,\infty)}, \mathcal{B}[0, \infty)^{\mathbb{N}})$-measurable.

Finally, let $(D[0, \infty), \|\cdot\|_\infty)$ be the Banach space of cadlag functions f on $[0, \infty)$ such that $\lim_{x\to\infty} f(x)$ is in $\mathbb{R}$, with supremum norm. We will equip this space with the open ball (with respect to $\|\cdot\|_\infty$) σ-field $\mathcal{P}$; see Pollard (1984: 199). We can then consider $\mathbb{G}_n : [0, \infty)^{\mathbb{N}} \to D[0, \infty)$ as a $([0, \infty)^{\mathbb{N}}, \mathcal{P})$-measurable map.

Altogether, the measurable mappings can now be summarized as

$$(\Omega, \mathcal{F}, \mathrm{P}) \overset{\{V_t\}}{\to} ([0, \infty)^{[0,\infty)}, \mathcal{B}[0, \infty)^{[0,\infty)}, \mathrm{P}_{V_t})$$

$$\overset{\{V_i\}}{\to} ([0, \infty)^{\mathbb{N}}, \mathcal{B}[0, \infty)^{\mathbb{N}}, \mathrm{P}_{V_i})$$

$$\overset{\mathbb{G}_n}{\to} (D[0, \infty), \mathcal{P}, \mathrm{P}_{\mathbb{G}})$$

where P_{V_t}, P_{V_i} and $\mathrm{P}_{\mathbb{G}_n}$ are image probability measures of the underlying measure P (Hoffmann-Jørgensen, 1994a: (1.44.1)).

2.2. Regenerative structure of the subsampled workload

A stochastic process is regenerative if there are random times where it starts 'stochastically' anew. More formally, we follow the approach of Levental (1988).

Definition 1. *Let* $\{X_i, i \geqslant 0\}$ *be a discrete parameter stochastic process with state space* $\mathbb{R}$ *equipped with the Borel* σ*-algebra* $\mathcal{B}(\mathbb{R})$ *and* $\{N_i, i \geqslant 1\}$ *a sequence of integer-valued random variables satisfying* $0 \leqslant N_1 < N_2 < \ldots < \infty$ *and* $\lim_{i\to\infty} N_i = \infty$ *almost surely. Both processes are assumed to be supported on the probability space* $(\Omega, \mathcal{F}, \mathrm{P})$*. The stochastic process* $\{X_i, i \geqslant 0\}$ *will be called a regenerative process with regeneration times* $\{N_i, i \geqslant 1\}$ *if*

$$\mathrm{E}[f(X_{N_i}, X_{N_i+1}, \ldots)|\mathcal{F}_{N_i}] = \mathrm{E}[f(X_{N_1}, X_{N_1+1}, \ldots)],$$

for all functions $f : [0, \infty)^{\mathbb{N}} \to \mathbb{R}$ *which are bounded and* $(\mathcal{B}[0, \infty)^{\mathbb{N}}, \mathcal{B}(\mathbb{R}))$*-measurable, and where* $\mathcal{F}_{N_i} = \sigma(\{X_{k\wedge N_i} : k = 1, 2, \ldots\})$.

First, we shall notice that (within the workload process) the subsampled workload is an embedded Markov chain.

Lemma 1. *Let* $\bar{s}_k = (s_1 \ldots, s_k)$, $\bar{t}_k = (t_1 \ldots, t_k)$,

$$u_{n,k}(\bar{s}_k, \bar{t}_k) = \sum_{i=1}^{k} ((n - t_i) \wedge s_i) - 1,$$

$$v_{n,k}(\bar{s}_k, \bar{t}_k) = \sum_{i=1}^{k} (s_i - (n - t_i) \wedge s_i).$$

Then the subsampled workload process $\{V_n, n \geqslant 0\}$ *is a Markov chain embedded within the workload process with transition kernel*

$$\begin{aligned}\mathrm{P}(V_{n+1} \in B|V_n) = {} & \mathrm{e}^{-\lambda} 1((V_n - 1)^+ \in B) \\ & + \sum_{k=1}^{\infty} \mathrm{e}^{-\lambda} \frac{\lambda^k}{k!} \int_{[n,n+1]^k} \int_{[0,\infty)^k} \\ & \times 1\big([(V_n + u_{n+1,k}(\bar{s}_k, \bar{t}_k))^+ + v_{n+1,k}(\bar{s}_k, \bar{t}_k)] \in B\big) \\ & \times \mathrm{d}F^{\otimes k}(\bar{s}_k)\, \mathrm{d}\bar{t}_k, \qquad a.s.,\end{aligned}$$

for any $B \in \mathcal{B}(\mathbb{R}_+)$ *and* $n \in \mathbb{N}_0$.

Secondly, the subsampled workload process is regenerative in the following sense.

Proposition 1. *Let* $N_1 = \inf\{n \geqslant 1|V_n = 0\}$ *and* $N_i = \inf\{n > N_{i-1}|V_n = 0\}$ *for* $i \geqslant 2$. *Then* $\{V_n, n \geqslant 1\}$ *is a regenerative process with regeneration times* $\{N_i, i \geqslant 1\}$. *Moreover,*

N_i is a stopping time with respect to the increasing sequence of σ-algebras $\mathcal{F}_n = \sigma(\{V_1 \ldots, V_n\})$, and if $\mathrm{E}(S^2) < \infty$ then $\mathrm{E}((N_2 - N_1)^2) < \infty$.

Proposition 2. *If $\mathrm{E}(N_2 - N_1) < \infty$, then*

$$\pi(A) = \mathrm{E}\left(\sum_{i=N_1}^{N_2-1} 1(V_i \in A)\right) / \mathrm{E}(N_2 - N_1) \tag{7}$$

is a probability measure on $([0, \infty), \mathcal{B}[0, \infty))$ and its distribution function $F_\pi(t) = \pi([0, t])$ equals G as given in (2).

3. Estimation of the stationary excess distribution F_e

In this section we state our main asymptotic normality results for the various estimators introduced above. The estimator G_n is constructed directly from the observations $V_1 \ldots, V_n$ of the regenerative process $\{V_i, i \geqslant 1\}$, and we first give an asymptotic normality result for G_n based on regenerative data. This is a regenerative empirical central limit theorem for G_n in $D[0, \infty)$. For this space, and throughout the paper, weak convergence in Banach spaces refers to σ-algebras generated by the open balls in the respective norms.

Theorem 1. *If $\mathrm{E}(S^2) < \infty$, then*

$$\mathbb{G}_n \to_{\mathcal{D}} Z \qquad \text{as } n \to \infty$$

in $(D[0, \infty), \|\cdot\|_\infty)$ where Z is a centred Gaussian process with covariance structure

$$\begin{aligned}\mathbb{D}(s, t) &= \mathrm{E}\, Z(s) Z(t) \\ &= \frac{1}{\mathrm{E}(N)} \operatorname{cov}(U_s - G(s)N, U_t - G(t)N),\end{aligned}$$

for all $s, t \geqslant 0$, $U_s = \sum_{i=N_1}^{N_2-1} 1(V_i \leqslant s)$ and $N = N_2 - N_1$.

We follow the approach and methodology of Buchmann and Grübel (2003), and in particular we use the weighted spaces defined there as follows. For τ in $\mathbb{R}$, let $D_\tau[0, \infty)$ be the space of all functions $f : [0, \infty) \to \mathbb{R}$, such that the function $x \mapsto \mathrm{e}^{-\tau x} f(x)$ is in $D[0, \infty)$. For f in $D_\tau[0, \infty)$, let $\|f\|_{\infty,\tau} = \sup_{x \geqslant 0} \mathrm{e}^{-\tau x} |f(x)|$, so that $(D_\tau[0, \infty), \|\cdot\|_{\infty,\tau})$ is a Banach space. We later show that, under conditions on G, the right-hand side of (5) is in $D_\tau[0, \infty)$.

In order to state our next result, let Z be the limiting process in Theorem 1, and let

$$\pi_k(\rho) = \frac{1}{\rho(1 - \rho)^k}, \qquad k \geqslant 1.$$

With this notation, we can now formulate the main result on weak convergence of the inverse estimator of the stationary excess distribution function.

Theorem 2. *Assume* $\mathrm{E}(S^2) < \infty$ *and that* $\tau > 0$ *is such that* $\tilde{F}_{\mathrm{e}}(\tau) < 1/(2\rho)$. *Then*

$$\sqrt{n}(F_{n,\mathrm{e}} - F_{\mathrm{e}}) \to_{\mathcal{D}} A, \qquad \text{as } n \to \infty,$$

in $(D_\tau[0, \infty), \|\cdot\|_{\infty,\tau})$, *where* A *is a centred Gaussian process given by* $A =_{\mathcal{D}} -Z(0)\Gamma + Z^o \star H$, *in which*

$$\Gamma = \sum_{k=1}^{\infty} (-1)^{k+1} \pi_k'(\rho)(G^o)^{\star k}$$

and

$$H = \sum_{k=1}^{\infty} (-1)^{k+1} k \pi_k(\rho)(G^o)^{\star(k-1)}.$$

4. Estimation of service-time distribution function F

Assume that the objective is to estimate the service-time distribution function F based on an estimate of the stationary excess distribution function. First, we notice the relations

$$F(x) = 1 - f_1 F_{\mathrm{e}}'(x),$$

$$f_1 = 1/F_{\mathrm{e}}'(0),$$

which are easily derived from the definition of the stationary excess distribution function. Consequently, the estimation problem is reformulated as a problem of estimating derivatives of F_{e}.

In the present paper we will suggest two pragmatic approaches to this problem, and pinpoint possible difficulties in obtaining rigorous convergence results. The performance of one of the suggestions is illustrated in the following section.

4.1. Kernel smoothing

One possibility that springs to mind is to smooth out the probability mass around a given point x. Specifically, let K be a probability density with mean 0 and variance 1, for instance the standard normal density. A kernel estimate with kernel K is defined as

$$F_{n,\mathrm{e}}'(x) = \int K\left(\frac{x - y}{h}\right) \mathrm{d}F_{n,\mathrm{e}}(y).$$

Here $h > 0$ is a number to be chosen, called the bandwidth of the estimator. If this method is pursued one has to find a way of dealing with the fact that the density of $F_{n,\mathrm{e}}$ has a discontinuity at zero. Various methods have been proposed for solving this; see Wand and Jones (1995: Section 2.11) or the recent and promising boundary-adjusted density estimation method of Chiu (2000).

Carrying on with an asymptotic analysis along the lines of Section 3 seems to be a

difficult matter. It is well known that if $F_{n,\mathrm{e}}$ is formed by i.i.d. distributed random variables, the finite-dimensional marginals converge to a multivariate Gaussian with a diagonal covariance matrix. Such a result is difficult to establish as the marginals of $F'_{n,\mathrm{e}}$ are asymptotically dependent as they arise from smoothing a convolution series and it seems hard to get a handle on the dependency structure. Aiming for a uniform result is not possible in the i.i.d. case because the limiting process is not tight. One might hope that the convolution series imposes tightness, but it is not at all clear how the results should be established. This is left as an open problem.

4.2. The Grenander estimator

Another possibility is to notice that F_{e} has a monotone density and is thus concave. One could then use the Grenander estimator, $\hat{F}_{n,\mathrm{e}}$ (i.e. the least concave majorant of $F_{n,\mathrm{e}}$) of the concave distribution F_{e}. Taking the left derivatives $\hat{F}'_{n,\mathrm{e}}(x)$ of the Grenander estimator yields the following nonparametric estimator of F:

$$F_n(x) = 1 - \hat{F}'_{n,\mathrm{e}}(0)\hat{F}'_{n,\mathrm{e}}(x).$$

This procedure is rather pragmatic in nature, as theoretical results seem hard to obtain. One might be motivated by the i.i.d. case, where the pointwise limit distribution was proved by Prakasa Rao (1983) to be proportional to the distribution of the argmax of the standard Brownian motion process with parabolic drift, or Groeneboom (1989), where the limit distribution is fully characterized. This is also left as an open problem.

5. Examples

To check the performance of the proposed procedure, we applied it to various simulated data sets. Results from two typical cases are summarized in this section.

We carried out all programming and simulation in the freely available computational statistical software package R; see http://www.r-project.org for more details.

Numerical estimates are obtained by discretizing the data and applying Panjer recursion (Panjer 1981: Section 4). Let F be a service-time distribution function. First, choose a discretization level $h > 0$. Secondly, let f_k denote the mass given by F_{e} to the interval $((k-0.5)h, (k+0.5)h]$,

$$f_k = F_{\mathrm{e}}((k+0.5)h) - F_{\mathrm{e}}((k-0.5)h),$$

and let the discrete distribution that gives mass f_k to the point kh, $k = 0, 1, 2, \ldots$, be an approximation to the distribution F_{e}. Finally, consider the following recursively defined approximation of the density of the workload distribution function:

$$g_0 = \frac{1-\rho}{1-\rho f_0}, \qquad g_k = \frac{\rho}{1-\rho f_0}\sum_{j=1}^{k} f_j g_{k-j}.$$

It is straightforward to invert the Panjer recursion,

$$f_0 = \frac{g_0 - 1 + \rho}{\rho g_0}, \qquad f_k = \frac{(1-\rho)g_k}{g_0^2 \rho} - \frac{1}{g_0}\sum_{j=1}^{k-1} f_j g_{k-j},$$

so that, given a discrete approximation (g_k) to G, we can calculate recursively a discrete approximation (f_k) to F_e.

The Grenander estimator $\hat{F}_{n,e}$ is found by forming the convex hull of the set

$$S = \{(x, y) \in \mathbb{R}^2 : x \geqslant 0,\ 0 \leqslant y \leqslant F_{n,e}(x))\}$$

and letting

$$\hat{F}_{n,e} = \sup\{y : (x, y) \in \text{chull}(S)\}.$$

This is easily implemented by the `chull` routine in R. Taking left derivatives yields the nonparametric estimator of F, as described above.

Suppose one is inclined to check whether the service-time distribution has heavy tails (i.e. $1 - F(x) = x^{-\alpha}L(x)$, as $x \to \infty$, where L is slowly varying; see Resnick 1997: (2.2)); it is desirable to develop a Hill-type estimator for the tail index α. This is indeed not a simple task for the problem at hand, as we are not facing an estimate of the distribution function based on an i.i.d. sample; see the comprehensive review by Resnick (1997) for the i.i.d. case. Instead we take a simpler and exploratory approach and plot $\log(1 - F_n(x))$ against x and $\log(x)$ to check for exponentially or polynomially decaying (heavy) tails, respectively.

5.1. Exponentially distributed service times

Figure 1 displays the estimates obtained for two independent samples of size 1000 from a discretely sampled $M/G/1$ queue with traffic intensity $\rho = 0.5$ and exponentially distributed service times with mean 1 (i.e. the arrival intensity of the Poisson distribution is 1/2). In what follows the dashed line shows the underlying theoretical properties of the distribution function. The solid black and grey lines are two independent realizations of the experiment. The upper left-hand panel shows the ecdfs of the subsampled workloads, the upper right-hand panel shows the estimated stationary excess distributions and the lower left-hand panel shows left derivatives of the Grenander estimator. Finally, the lower right-hand panel shows a log plot of the estimated tail probabilities.

5.2. Pareto distributed service times

In this example we consider service times with cdf

$$F(x) = 1 - x^{-\alpha}$$

for $x \geqslant 1$ and $\alpha > 1$. This is actually a Pareto distribution with with parameter α.

Figure 2 displays the estimates obtained for two independent samples of size 1000 from a discretely sampled $M/G/1$ queue with traffic intensity $\rho = 0.5$ and Pareto distributed service times with mean 5/3 (i.e. the arrival intensity of the Poisson distribution is 3/10 and the

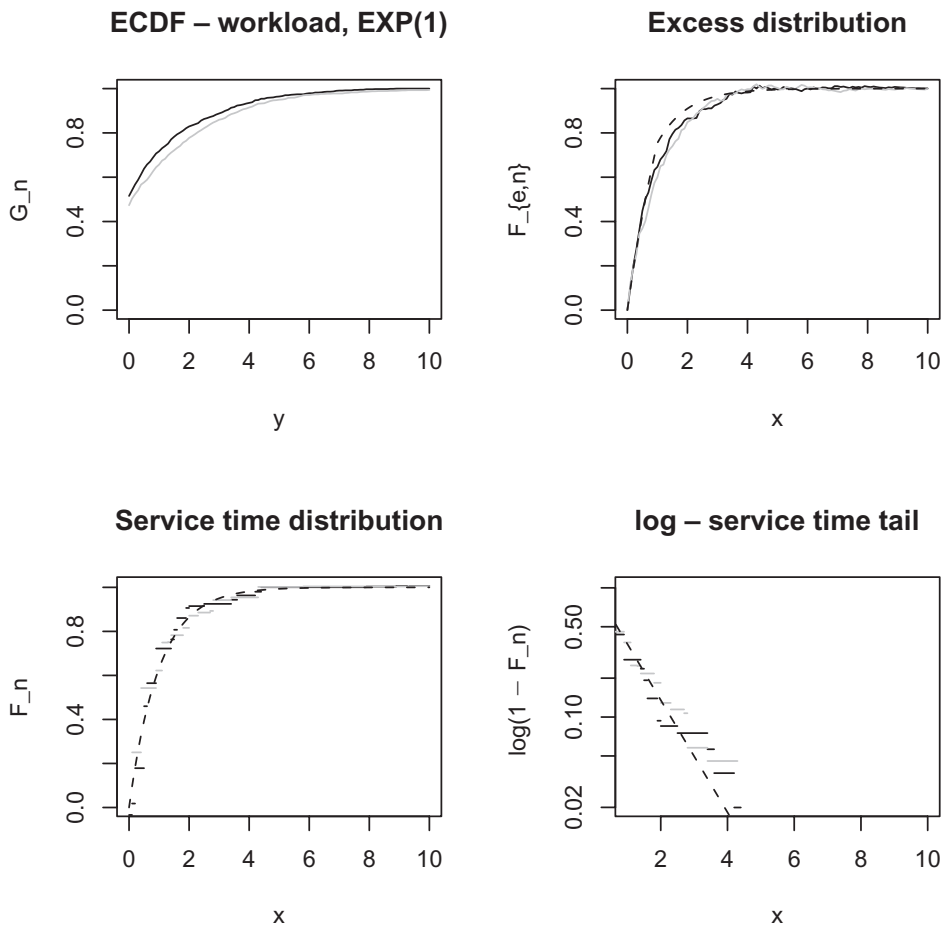

Figure 1. Two estimates for service-time distribution for exponentially distributed service times ($n = 1000$, $\lambda = 1/2$, $f_1 = 1$ and $\rho = 0.5$).

parameter of the Pareto distribution is $\alpha = 2.5$). As above, the dashed line shows the underlying theoretical properties of the distribution function. The upper left-hand panel shows the ecdfs of the subsampled workloads, the upper right-hand panel shows the estimated stationary excess distributions and the lower left-hand panel shows left derivatives of the Grenander estimator. Finally, the lower right-hand panel shows a log plot of the estimated tail probabilities.

Figure 3 shows that increasing the sample size improves the estimate, but Figure 4 shows that decreasing the rate of the Pareto distribution (into a region where the functional central limit theorem is not guaranteed) leads to a deterioration.

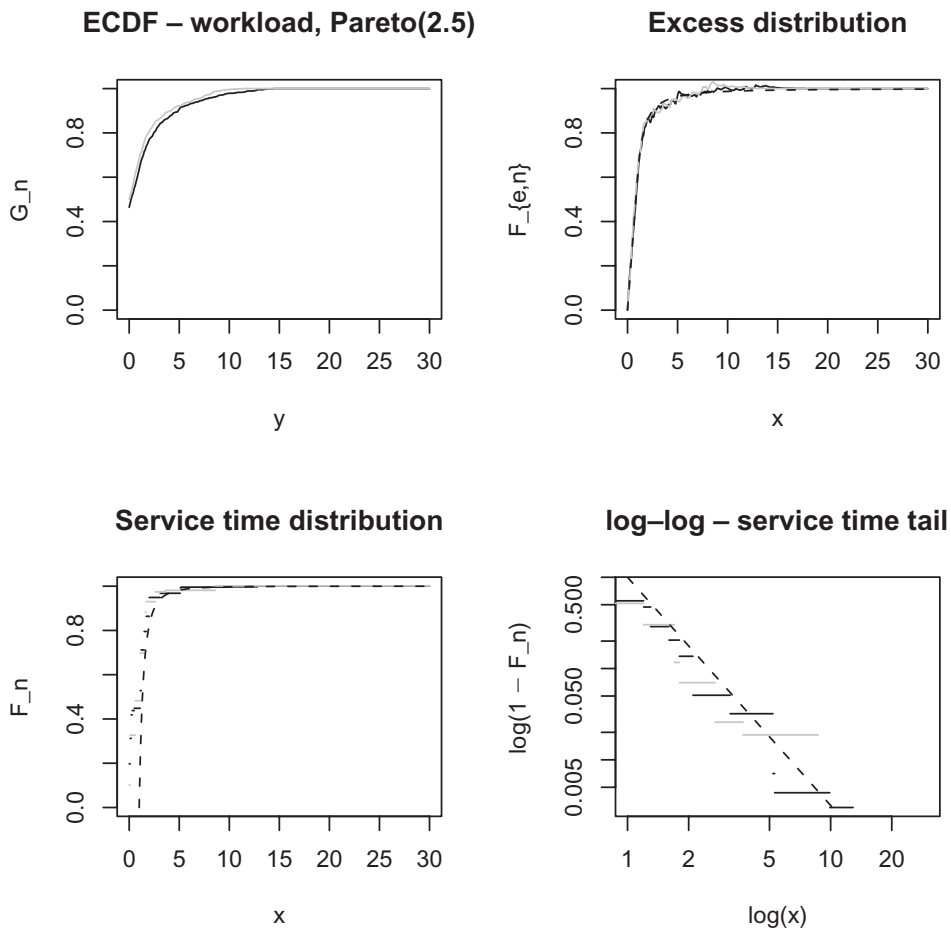

Figure 2. Two estimates for the service-time distribution for Pareto distributed service times ($n = 1000$, $\lambda = 3/10$, $\alpha = 2.5$ and $\rho = 0.5$).

6. Proofs

6.1. Proofs of the regenerative structure

Proof of Lemma 1. As the number of arrivals in $(n, n+1]$ is Poisson distributed and, given the number of arrivals in this interval, the arrival times are uniformly distributed over $(n, n+1]$, the transition kernel is obtained by conditioning. The terms $u_{n+1,k}$ and $v_{n+1,k}$ are

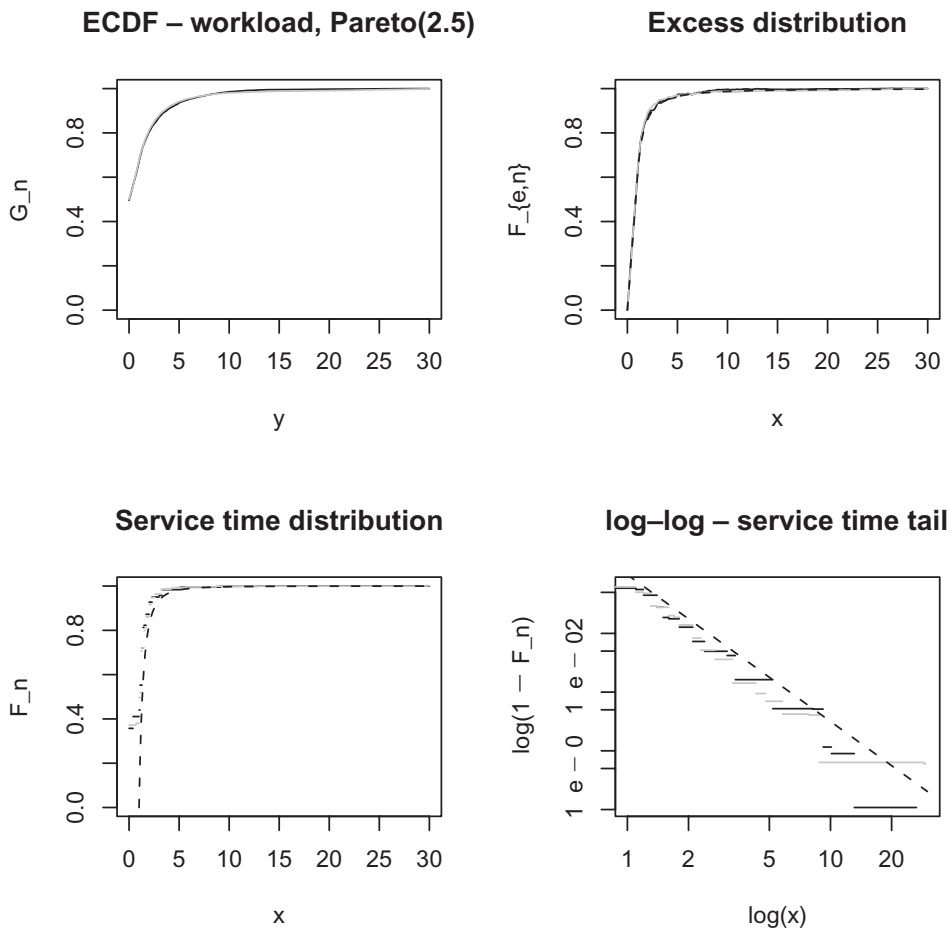

Figure 3. Two estimates for the service-time distribution for Pareto distributed service times ($n = 10\,000$, $\lambda = 3/10$, $\alpha = 2.5$ and $\rho = 0.5$).

obtained by splitting the arrived workload into the part contained in $(n, n+1]$ and the part contained in $(n+1, \infty)$. □

Proof of Proposition 1. The sequence $\{N_i, i \geqslant 1\}$ is integer-valued, strictly increasing and satisfies $\lim_{i\to\infty} N_i = \infty$ a.s. by construction. For finiteness of the N_i consider $N_2 - N_1$. Let $B_1, B_2, \ldots$ be the lengths of the busy periods after N_1 and $I_1, I_2, \ldots$ the lengths of the idle periods after N_1. If $\tau = \inf\{n \geqslant 1 | I_n \geqslant 1\}$, then $N_2 - N_1 \leqslant \sum_{i=1}^{\tau} (B_i + I_i)$. Consequently,

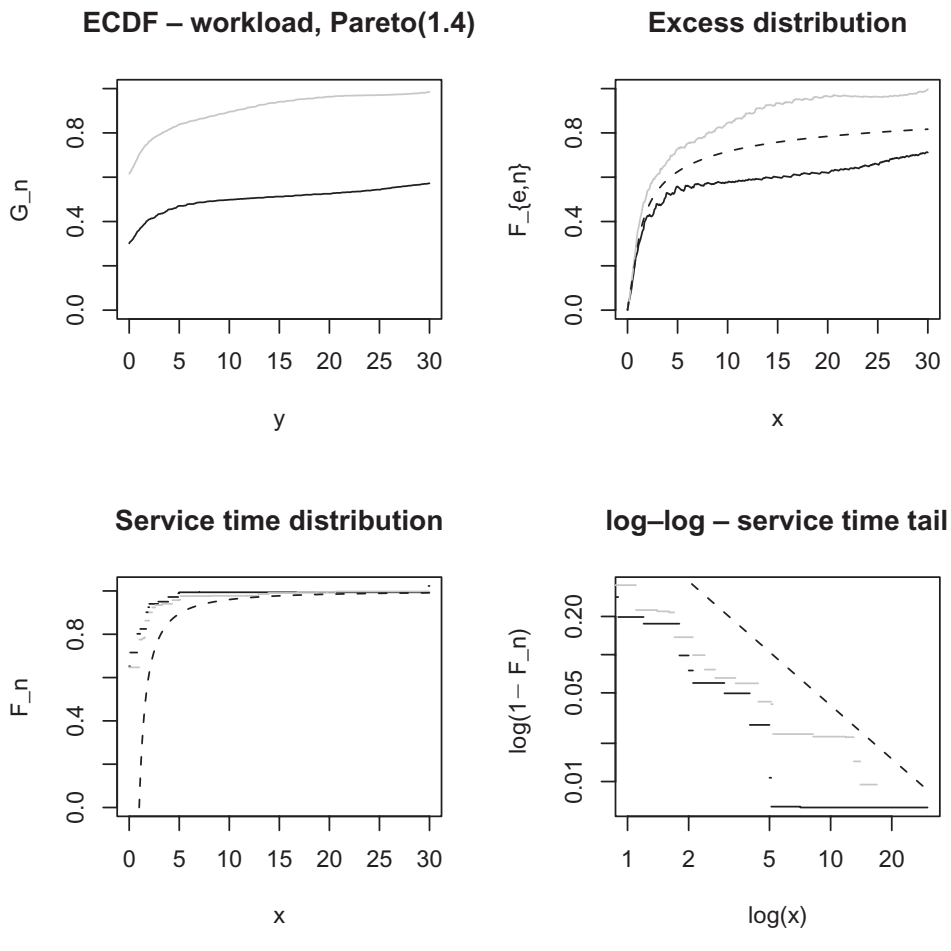

Figure 4. Two estimates for the service-time distribution for Pareto distributed service times ($n = 1000$, $\lambda = 1/7$, $\alpha = 1.4$ and $\rho = 0.5$).

$$\mathrm{P}(N_2 - N_1 < \infty) \geqslant \mathrm{P}\left(\sum_{i=1}^{\tau} (B_i + I_i) < \infty\right)$$

$$\geqslant 1.$$

As N_1 is stochastically dominated by $N_2 - N_1$, we obtain, for $i \geqslant 2$,

$$N_i = N_1 + \sum_{k=2}^{i} (N_k - N_{k-1}) < \infty, \qquad \text{a.s.}$$

First notice by standard arguments (see Hoffmann-Jørgensen 1994a: Section 7.2) that N_i is a stopping time with respect to the increasing sequence of σ-algebras $\mathcal{F}_n$. Then by the strong Markov property (Asmussen 2003: Theorem I.1.1),

$$\begin{aligned}\mathrm{E}[f(V_{N_i}, V_{N_i+1}, \ldots)|\mathcal{F}_{N_i}] &= \mathrm{E}_{V_{N_i}} f(V_0, V_1, \ldots) \\ &= \mathrm{E}_0 f(V_0, V_1, \ldots) \\ &= \mathrm{E} f(V_{N_1}, V_{N_1+1}, \ldots).\end{aligned}$$

By Wald's second moment identity (Asmussen 2003: Proposition A.10.2),

$$\begin{aligned}\mathrm{E}(N_2 - N_1)^2 &\leqslant \mathrm{E}\left(\sum_{i=1}^{\tau} (B_i + I_i)\right)^2 \\ &= \mathrm{var}(I_1 + B_1)\mathrm{E}(\tau) + 3\mu^2\mathrm{E}(\tau^2),\end{aligned}$$

which is finite if and only if $\mathrm{E}(B_1^2) < \infty$, which again is finite if and only if $\mathrm{E}(S^2) < \infty$. The last equivalence follows directly from Kendall's functional equation; see Feller (1971: (4.1)). □

Proof of Proposition 2. It is straightforward to prove that π is a probability measure. It is also obvious that V_t and V_n have the same limiting distribution (both in weak convergence and total variation). As V_n is regenerative it follows that its limiting distribution exists (in weak convergence and total variation) and equals π (Asmussen 2003: Theorem VI.1.2 and Corollary 1.5). As the limiting distributions of V_n and V_t coincide, the stated result follows. □

6.2. Proof of Theorem 1

A standard method for proving empirical central limit theorems is first to prove that the finite-dimensional distributions (fidis) converge to a multivariate normal distribution with the stated covariance structure and secondly to prove a tightness condition (Pollard 1984: Section VII.5). Henceforth, we notice that $\{V_n, n \geqslant 0\}$ is a regenerative process taking values in $\mathbb{R}_+$, with regeneration times $\{N_i, i \geqslant 1\}$ (as defined in Proposition 1). Convergence of the fidis is now proved by noticing that any linear combination

$$a\mathbb{G}_n(s) + b\mathbb{G}_n(t) = \sqrt{n}\left(\frac{1}{n}\sum_{i=1}^{n}(a1(V_i \leqslant s) + b1(V_i \leqslant t)) - (aG(s) + bG(t))\right)$$

is asymptotically normal with mean 0 and variance

$$\sigma^2 = \frac{1}{\mathrm{E}(N)}\mathrm{var}\left(\sum_{i=N_1}^{N_2-1}(a1(V_i \leqslant s) + b1(V_i \leqslant t)) - (aG(s) + bG(t))N\right)$$

$$= \frac{1}{\mathrm{E}(N)}(a^2\mathrm{var}(U_s - G(s)N) + b^2\mathrm{var}(U_t - G(t)N)$$

$$+ 2ab\,\mathrm{cov}(U_s - G(s)N,\, U_t - G(t)N))$$

by the central limit theorem for renewal–reward processes; see Asmussen (2003: Theorem VI.3.2). Hence, $(\mathbb{G}(s), \mathbb{G}(t))$ has a zero-mean Gaussian distribution with the stated covariance by the Cramér–Wold device (Billingsley 1968: Theorem 7.7). The above argument clearly extends to linear combinations of any finite number of fidis.

Secondly, tightness is handled by considering the measure space $(\mathbb{R}_+, \mathcal{B}(\mathbb{R}_+))$ and defining the family of functions $\mathcal{H} = \{1(\cdot \geqslant t) | t \in \mathbb{R}\}$. We notice that $\mathcal{H}$ is indexed by a Borel set of a metric space and thus permissible (Pollard 1984: 16). As $\mathcal{H}$ is a Vapnik–Chervonenkis class (Pollard 1984: 16) the combinatorial condition

$$\int_0^\infty [\log N_2(u, \mathcal{H})]^{1/2}\,\mathrm{d}u < \infty$$

is satisfied. Finally, from Proposition 1 we notice that $\mathrm{E}(N_2 - N_1)^2 < \infty$. Hence the result is a direct application of Tsai (1998: Theorem 4.3). See also Levental (1988: Theorem 4.9) for a slightly weaker result.

6.3. Proof of Theorem 2

We essentially follow Buchmann and Grübel (2003). In (5), our quantity of interest F_e is given in terms of ρ and G, and this inverse representation is made precise in Proposition 3 below. Thus the equilibrium distribution can be regarded as the output that arises when a particular functional, Λ say, is applied to (ρ, G), so that

$$F_\mathrm{e} = \Lambda(\rho, G)(= \Lambda(1 - G(0), G)).$$

The plug-in estimator given in Section 1 is then

$$F_{n,\mathrm{e}} = \Lambda(\hat{\rho}_n, G_n)(= \Lambda(1 - G_n(0), G_n)).$$

Proposition 4 is a suitable differentiability result for Λ, and this leads to the asymptotic normality of $F_{n,\mathrm{e}}$ as stated in Theorem 2. In adapting the proofs of Buchmann and Grübel (2003) to the case in hand, we deal with the obvious technical changes necessary for a compound geometric rather than a compound Poisson, and in addition we incorporate estimation of the parameter ar ρ defining the geometric distribution. We give sufficient details here to make the present paper reasonably self-contained.

First, we note that (5) involves convolutions, and so we use the convolution framework of Buchmann and Grübel (2003). Define $D(\infty) = \cup_{\tau>0} D_\tau[0, \infty)$, and let $D_\mathrm{m}(\infty)$ ($\subseteq D(\infty)$) be the space of functions in $D(\infty)$ that have finite variation on $[0, x]$ for all $x > 0$. Then any function H in $D_\mathrm{m}(\infty)$ can be identified with a (not necessarily finite) signed measure

μ_H on $([0, \infty), \mathcal{B}[0, \infty))$, where $\mathcal{B}[0, \infty)$ denotes the Borel sets of $[0, \infty)$, and this identification is given by $H(x) = \mu_H([0, x])$. Elements of $D_{\mathrm{m}}(\infty)$ can act as integrators in convolution integrals, and for g in $D(\infty)$ and H in $D_{\mathrm{m}}(\infty)$, define

$$g \star H(x) = \int g(x - y) H(\mathrm{d}y), \qquad x \geqslant 0.$$

We note further that elements of $D_{\mathrm{m}}(\infty)$ are identified via their Laplace transforms, given by

$$\tilde{H}(\theta) = \int \mathrm{e}^{-\theta x} H(\mathrm{d}x).$$

If $H \in D_{\mathrm{m}}(\infty)$ is in $D_\tau[0, \infty)$, $\tau > 0$, then the Laplace transform is defined for $\theta > \tau$. Lemma 6(b) of Buchmann and Grübel (2003) provides a useful result linking convolution and the norms $\| \cdot \|_{\infty,\tau}$ and $\| \cdot \|_\infty$, and we quote this result without proof. Let $D_m^+(\infty)$ be the subset of $D_{\mathrm{m}}(\infty)$ consisting of functions H such that the associated measure is non-negative. If H is in $D_m^+(\infty)$, then

$$\| g \star H \|_{\infty,\tau} \leqslant \| g \|_{\infty,\tau} \tilde{\mathrm{H}}(\tau) \qquad \text{for all } \tau > 0. \tag{8}$$

We are now ready to state and prove the following result which can be regarded as specifying the inverse of the compound geometric functional.

Proposition 3. *Let $0 < \rho < 1$ and $\tau > 0$. If G is a distribution function on $[0, \infty)$ with $\tilde{G}^o(\tau) < (1 - \rho)$, then the series*

$$\Lambda(\rho, G) = \sum_{k=1}^{\infty} (-1)^{k+1} \pi_k(\rho) (G^0)^{\star k}$$

converges in $D_\tau[0, \infty)$. Furthermore, if $G = \Psi(\rho, F)$ then $F_{\mathrm{e}} = \Lambda(\rho, G)$.

Proof. Observe that G^o and $(G^o)^{\star k}$ are in $D_\tau[0, \infty)$ for $\tau > 0$, and so the partial sums of the series are in $D_\tau[0, \infty)$. Applying (8) repeatedly, we obtain

$$\begin{aligned} \| (G^o)^{\star \mathrm{k}} \|_{\infty,\tau} &\leqslant \| G^o \|_{\infty,\tau} \tilde{G}^o(\tau)^{k-1} \\ &\leqslant \tilde{G}^o(\tau)^{k-1}, \end{aligned} \tag{9}$$

where the last inequality holds because G^o is the distribution function of a (possibly sub-) probability measure, and $\tau > 0$. By assumption we have $\tilde{G}^o(\tau) < 1 - \rho$, so there exists η, $0 < \eta < 1$, such that $\tilde{G}^o(\tau) < \eta(1 - \rho)$. Then

$$\sum_{k=1}^{\infty} \pi_k(\rho) \big(\tilde{G}^o(\tau) \big)^{k-1} \leqslant \frac{1}{\rho(1 - \rho)} \sum_{k=1}^{\infty} \eta^{k-1},$$

and this is finite. Thus the given series converges in $D_\tau[0, \infty)$, and so $\Lambda(\rho, G)$ is in $D_\tau[0, \infty)$.

This series $\Lambda(\rho, G)$ is the difference of two non-decreasing functions, and so it is $D_{\mathrm{m}}(\infty)$, and hence it is identified by its Laplace transform. For $\theta > \tau$ we have

$$\tilde{\Lambda}(\rho, G)(\theta) = \sum_{k=1}^{\infty} (-1)^{k+1} \pi_k(\rho) \tilde{G}^o(\theta)^k$$

$$= \frac{\tilde{G}^o(\theta)}{\rho(1-\rho)} \left(1 + \frac{\tilde{G}^o(\theta)}{1-\rho}\right)^{-1}. \tag{10}$$

On the other hand, using $G = \Psi(\rho, F)$ from (2), we obtain

$$\tilde{G}(\theta) = \frac{1-\rho}{1-\rho\tilde{F}_e(\theta)},$$

so that

$$\tilde{G}^o(\theta) = \tilde{G}(\theta) - (1-\rho) = \frac{(1-\rho)\rho\tilde{F}_e(\theta)}{1-\rho\tilde{F}_e(\theta)}, \qquad \text{for all } \theta > 0.$$

Rearranging this, we find that

$$\tilde{F}_e(\theta) = \frac{\tilde{G}^o(\theta)}{\rho(1-\rho)} \left(1 + \frac{\tilde{G}^o(\theta)}{1-\rho}\right)^{-1},$$

and this gives $\Lambda(\rho, G) = F_e$ by comparison with (10). □

The main part of the proof of Theorem 2 is to establish the differentiability of the functional that maps G onto F_e, and this entails differentiability of the functional Λ as a map from a subset of $(0, 1) \times D[0, \infty)$ to $D_\tau[0, \infty)$.

Proposition 4. *Let ρ and ρ_n ($n \in \mathbb{N}$) be in $(0, 1)$, and suppose that*

$$|\sqrt{n}(\rho_n - \rho) - \gamma| \to 0 \qquad \text{as } n \to \infty, \tag{11}$$

for some γ in $\mathbb{R}$. Let G and G_n ($n \in \mathbb{N}$) be elements of $D_m^+(\infty) \cap D[0, \infty)$ with $G(0) = 0$ and $G_n(0) = 0$ for all $n \in \mathbb{N}$. Suppose that $\tau > 0$ is such that $\tilde{G}(\tau) < (1-\rho)$, and that

$$\|\sqrt{n}(G_n - G) - h\|_\infty \to 0 \qquad \text{as } n \to \infty \tag{12}$$

for some $h \in D[0, \infty)$. Then, as $n \to \infty$,

$$\|\sqrt{n}(\Lambda(\rho_n, G_n) - \Lambda(\rho, G)) - \gamma\Gamma - h \star H\|_{\infty,\tau} \to 0,$$

where $\Gamma = \sum_{k=1}^{\infty} (-1)^{k+1} \pi_k'(\rho) G^{\star k}$ and $H = \sum_{k=1}^{\infty} (-1)^{k+1} k \pi_k(\rho) G^{\star(k-1)}$.

Proof. Similar methods to those used in Proposition 3 show that Γ and H are in $D_\tau[0, \infty)$ and also that H is in $D_m(\infty)$ (and hence is an appropriate quantity for the right-hand factor in a convolution).

We have

$$\begin{aligned}\|\sqrt{n}(\Lambda(\rho_n, G_n) - \Lambda(\rho, G)) - \gamma\Gamma - h \star H\|_{\infty,\tau} \\ \leqslant \|\sqrt{n}(\Lambda(\rho, G_n) - \Lambda(\rho, G)) - h \star H\|_{\infty,\tau} \\ + \|\sqrt{n}(\Lambda(\rho_n, G_n) - \Lambda(\rho, G_n)) - \gamma\Gamma\|_{\infty,\tau}. \end{aligned} \tag{13}$$

By assumption $G(0) = 0$ and $G_n(0) = 0$, and so we have $G^o = G$, and $G_n = G_n^o$. For the second term on the right-hand side of (13), we notice that

$$\begin{aligned}&\|\sqrt{n}(\Lambda(\rho_n, G_n) - \Lambda(\rho, G_n)) - \gamma\Gamma\|_{\infty,\tau} \\ &= \left\| \sqrt{n}\left(\sum_{k=1}^{\infty} (-1)^{k+1} \pi_k(\rho_n) G_n^{\star k} - \sum_{k=1}^{\infty} (-1)^{k+1} \pi_k(\rho) G_n^{\star k} \right) - \gamma \sum_{k=1}^{\infty} (-1)^{k+1} \pi_k'(\rho) G^{\star k} \right\|_{\infty,\tau} \\ &\leqslant \left\| \sum_{k=1}^{\infty} (-1)^{k+1} \left(\sqrt{n}(\pi_k(\rho_n) - \pi_k(\rho)) - \gamma \pi_k'(\rho) \right) G_n^{\star k} \right\|_{\infty,\tau} \\ &\quad + \left\| \gamma \sum_{k=1}^{\infty} (-1)^{k+1} \pi_k'(\rho) \left(G_n^{\star k} - G^{\star k} \right) \right\|_{\infty,\tau}. \end{aligned} \tag{14}$$

We have $G_n^{\star k} - G^{\star k} = (G_n - G) \star H_{n,k}$, with $H_{n,k} = \sum_{j=0}^{k-1} G_n^{\star j} \star G^{\star(k-1-j)}$ for $k \geqslant 1$. Using (8), we find that $\|(G_n - G) \star H_{n,k}\|_{\infty,\tau} \leqslant \|G_n - G\|_{\infty,\tau} \widetilde{H_{n,k}}(\tau)$. This yields an upper bound for the second term of (14),

$$\begin{aligned}\left\| \gamma \sum_{k=1}^{\infty} (-1)^{k+1} \pi_k(\rho) \left(G_n^{\star k} - G^{\star k} \right) \right\|_{\infty,\tau} &\leqslant |\gamma| \sum_{k=1}^{\infty} |\pi_k'(\rho)| \, \|G_n^{\star k} - G^{\star k}\|_{\infty,\tau} \\ &= |\gamma| \sum_{k=1}^{\infty} |\pi_k(\rho)| \, \|(G_n - G) \star H_{n,k}\|_{\infty,\tau} \\ &\leqslant |\gamma| \sum_{k=1}^{\infty} |\pi_k'(\rho)| \, \|G_n - G\|_{\infty,\tau} \widetilde{H_{n,k}}(\tau) \\ &\leqslant |\gamma| \, \|G_n - G\|_{\infty} \sum_{k=1}^{\infty} |\pi_k'(\rho)| \widetilde{H_{n,k}}(\tau), \end{aligned} \tag{15}$$

on noting in the final step that $\|G_n - G\|_{\infty,\tau} \leqslant \|G_n - G\|_{\infty}$ for $\tau > 0$. We now aim to bound $\widetilde{H_{n,k}}(\tau)$, and this means we must bound $\tilde{G}(\tau)$ and $\widetilde{G_n}(\tau)$. As in the proof of Proposition 3, the assumption $\tilde{G}(\tau) < 1 - \rho$ implies that we can find η, $0 < \eta < 1$, such that $\tilde{G}(\tau) < \eta(1 - \rho)$. Integration by parts gives

$$|\widetilde{G_n}(\tau) - \tilde{G}(\tau)| \leqslant \tau \int_0^\infty \mathrm{e}^{-\tau x} |G_n(x) - G(x)| \, \mathrm{d}x$$

$$\leqslant \tau \|G_n - G\|_\infty \int_0^\infty \mathrm{e}^{-\tau x} \, \mathrm{d}x,$$

and this converges to zero as n tends to infinity, since from (12) we know that $\|G_n - G\|_\infty \to 0$. This means that there exists n_1 such that $n \geqslant n_1$ implies that $\widetilde{G_n}(\tau) < \eta(1-\rho)$. This in turn implies that for $n \geqslant n_1$, we have

$$\widetilde{H_{n,k}}(\tau) \leqslant k\eta^{k-1}(1-\rho)^{k-1}.$$

Plugging this into (15), we obtain

$$\left\| \gamma \sum_{k=1}^\infty (-1)^{k+1} \pi_k'(\rho)(G_n^{\star k} - G^{\star k}) \right\|_{\infty,\tau} \leqslant |\gamma| \, \|G_n - G\|_\infty \sum_{k=1}^\infty |\pi_k'(\rho)| k\eta^{k-1}(1-\rho)^{k-1}$$

$$\leqslant |\gamma| \, \|G_n - G\|_\infty \left(\frac{1}{\rho(1-\rho)^2} \sum_{k=1}^\infty k(k+1)\eta^{k-1} + \frac{1}{\rho^2(1-\rho)^2} \sum_{k=1}^\infty k\eta^{k-1} \right).$$

Since $0 < \eta < 1$, the two series above converge to finite limits, and the fact that $\|G_n - G\|_\infty \to 0$ shows that the second term of (14) converges to zero.

For the first term, note that by (9), and with n_1 as above, for all $n \geqslant n_1$, we have $\|G_n^{\star k}\|_{\infty,\tau} \leqslant \eta^{k-1}(1-\rho)^{k-1}$. Then the first term in (14) is bounded as follows:

$$\left\| \sum_{k=1}^\infty (-1)^{k+1} (\sqrt{n}(\pi_k(\rho_n) - \pi_k(\rho)) - \gamma\pi_k'(\rho)) G_n^{\star k} \right\|_{\infty,\tau}$$

$$\leqslant \sum_{k=1}^\infty |\sqrt{n}(\pi_k(\rho_n) - \pi_k(\rho)) - \gamma\pi_k'(\rho)| (\eta(1-\rho))^{k-1}. \tag{16}$$

By the mean value theorem, we have

$$|\sqrt{n}(\pi_k(\rho_n) - \pi_k(\rho)) - \gamma\pi_k'(\rho)|$$

$$\leqslant |\sqrt{n}(\rho_n - \rho) - \gamma| \, |\pi_k'(\xi_{n,k})| + |\gamma| \, |\pi_k'(\xi_{n,k}) - \pi_k'(\rho)|$$

$$\leqslant |\sqrt{n}(\rho_n - \rho) - \gamma| \, |\pi_k'(\xi_{n,k})| + |\gamma| \, |\pi_k''(\zeta_{n,k})| \, |\xi_{n,k} - \rho|,$$

for some $\xi_{n,k}$ and $\zeta_{n,k}$ in $(\rho_{L,n}, \rho_{U,n})$ $(\subseteq (0, 1))$, where $\rho_{L,n} = \rho \wedge \rho_n$ and $\rho_{U,n} = \rho \vee \rho_n$. Then we have

$$|\pi_k'(\xi_{n,k})| \leqslant \frac{k+2}{\rho_{L,n}^2 (1 - \rho_{U,n})^{k+1}}$$

$$|\pi_k''(\zeta_{n,k})| \leqslant \frac{k^2 + 4k + 5}{\rho_{L,n}^3 (1 - \rho_{U,n})^{k+2}}.$$

Let η, $0 < \eta < 1$, be as above, and let α be such that $\eta < \alpha < 1$. Since ρ_n converges to ρ, we have that $0 < \eta(1-\rho)/(1-\rho_n) < \alpha < 1$ for all n large enough, and so $0 < \eta(1-\rho)/(1-\rho_{U,n}) < \alpha$ for all n large enough. We also have

$$\frac{1}{\rho_{L,n}^2(1-\rho_{U,n})^2} \leqslant \frac{2}{\rho^2(1-\rho)^2},$$

$$\frac{1}{\rho_{L,n}^3(1-\rho_{U,n})^3} \leqslant \frac{2}{\rho^3(1-\rho)^3}$$

for all n large enough, because $\rho_{L,n}$ and $\rho_{U,n}$ both converge to ρ. Thus, for all n large enough, (17) is bounded above by

$$|\sqrt{n}(\rho_n - \rho) - \gamma| \frac{2}{\rho^2(1-\rho)^2} \sum_{k=1}^{\infty} (k+2)\alpha^{k-1}$$

$$+ |\gamma|\,|\rho_n - \rho| \frac{2}{\rho^3(1-\rho)^3} \sum_{k=1}^{\infty} (k^2+4k+5)\alpha^{k-1},$$

which converges to zero as n tends to infinity. Hence we have shown that the right-hand side of (14) converges to zero, i.e. the second term on the right-hand side of (13) converges to zero.

The first term on the right-hand side of (13) is similar to that in Proposition 8 in Buchmann and Grübel (2003). Using similar methods to those used there, with η there replaced by η above, and with minor resulting technical modifications in the details of the argument there, we obtain convergence of the first term on the right-hand side of (13) to zero, and the proposition is proved. □

Proof of Theorem 2. We follow the proof of Theorem 2 in Buchmann and Grübel (2003). Theorem 1 implies that $\sqrt{n}(G_n - G) \to_{\mathcal{D}} Z$ as $n \to \infty$, where Z is a zero-mean Gaussian process in $(D[0, \infty), \|\cdot\|_\infty)$.

The sample paths of Z are bounded and uniformly continuous with respect to $L_2(\pi)$ (Levental 1988: Theorem 4.9), where π is the steady state distribution for the regenerative process as given in (7). From Proposition 2 we can replace the corresponding distribution function F_π with G. Then, given $\varepsilon > 0$, there exists $\delta > 0$ such that $|G(s) - G(t)| < \delta$ implies that $|Z(s) - Z(t)| < \varepsilon$. Hence the sample paths of Z can only have discontinuities at the jump points of G, and so the process Z is concentrated on the separable subspace of $D[0, \infty)$ consisting of those functions that can only jump at those points where G jumps.

Let S be the linear continuous map that takes f in $D[0, \infty)$ onto $(-f(0), f^o)^{\mathrm{T}}$ in $\mathbb{R} \times D[0, \infty)$, where $(x, f)^{\mathrm{T}}$ denotes the transpose of (x, f). The continuous mapping theorem (Pollard 1984: Section IV.12) implies that

$$\sqrt{n}\left(\begin{pmatrix} \hat{\rho}_n \\ G_n^o \end{pmatrix} - \begin{pmatrix} \rho \\ G^o \end{pmatrix} \right) \to_{\mathcal{D}} \begin{pmatrix} -Z(0) \\ Z^o \end{pmatrix}, \qquad \text{as } n \to \infty,$$

with the limiting distribution being concentrated on a separable subset of $\mathbb{R} \times D[0, \infty)$.

The Skorohod representation theorem (Pollard 1984: Section IV.3) then allows us to construct a probability space $(\Omega^\dagger, \mathcal{F}^\dagger, \mathrm{P}^\dagger)$ carrying random quantities $(\rho_n^\dagger, G_n^\dagger)^{\mathrm{T}}$ and $(W^\dagger, Z^\dagger)^{\mathrm{T}}$ such that $(\rho_n^\dagger, G_n^\dagger)^{\mathrm{T}} =_{\mathcal{D}} (\hat{\rho}_n, G_n^o)^{\mathrm{T}}$, $(W^\dagger, Z^\dagger)^{\mathrm{T}} =_{\mathcal{D}} (-Z(0), Z^0)^{\mathrm{T}}$, and such that

$$\sqrt{n}\left(\begin{pmatrix} \rho_n^\dagger \\ G_n^\dagger \end{pmatrix} - \begin{pmatrix} \rho \\ G^o \end{pmatrix}\right) \to \begin{pmatrix} W^\dagger \\ Z^\dagger \end{pmatrix}, \qquad \text{as } n \to \infty, \tag{17}$$

in $\mathbb{R} \times D[0, \infty)$ with $\mathrm{P}^\dagger$-probability one.

From (17) we have

$$\sqrt{n}((G_n^\dagger) - G^o) \to Z^\dagger \qquad \text{as } n \to \infty, \tag{18}$$

in $(D[0, \infty), \|\cdot\|_\infty)$, and

$$\sqrt{n}(\rho_n^\dagger - \rho) \to W^\dagger \qquad \text{as } n \to \infty,$$

with $\mathrm{P}^\dagger$-probability one.

It is assumed in the statement of Theorem 2 that $\widetilde{F_{\mathrm{e}}}(\tau) < 1/(2\rho)$, and so we have $\tilde{G}^o(\tau) < 1 - \rho$. Further, $\rho_n^\dagger$ is in $(0, 1)$ for all n large enough $\mathrm{P}^\dagger$-a.s. Thus the conditions of Proposition 4 are satisfied eventually with $\mathrm{P}^\dagger$-probability one, and so we have $\mathrm{P}^\dagger$-a.s.,

$$\sqrt{n}(\Lambda(\rho_n^\dagger, G_n^\dagger) - \Lambda(\rho, G^o)) \to -W^\dagger \Gamma + Z^\dagger \star H, \tag{19}$$

where

$$\Gamma = \sum_{k=1}^{\infty} (-1)^{k+1} \pi_k'(\rho)(G^o)^{\star k}, \qquad H = \sum_{k=1}^{\infty} (-1)^{k+1} k \pi_k(\rho)(G^o)^{\star(k-1)}.$$

Then we have convergence in distribution of the original sequences,

$$\sqrt{n}(\Lambda(\hat{\rho}_n, G_n^o) - \Lambda(\rho, G^o)) \to_{\mathcal{D}} A,$$

where A is a zero-mean Gaussian process, obtained by applying the linear bounded map that takes f in $D[0, \infty)$ onto $-f(0)\Gamma + f^o \star H$ to the sample paths of Z. Using Proposition 3 for the definition of Λ, this gives

$$\sqrt{n}(F_{n,\mathrm{e}} - F_{\mathrm{e}}) \to_{\mathcal{D}} A,$$

where the process A is as given in the statement of the theorem. □

7. Discussion

Buchmann and Grübel (2003) noted that their basic set-up could be generalized in several ways. In the previous sections we have provided such a generalization towards sampled regenerative processes. In particular, we provide a method for decompounding geometric sums. It is our belief that this brings the basic ideas of Buchmann and Grübel (2003) even closer to real-life applications in such areas as control of queuing systems, infinite-capacity models and insurance mathematics.

Having said this, we have to acknowledge that in heavy and thin traffic we easily run into

danger in practice, as $\hat{\rho}_n$ might be 1 or 0, respectively. This causes problems in (6). A similar discussion arises with respect to the requirement that $\tilde{G}_{\mathrm{n}}^{0}(\tau) < 1 - \rho$ for the convolution sum to be identifiable. Technically this is solved by noticing that things are eventually satisfied for large n with probability one.

Besides the open problems stated in Section 4.1 and 4.2 regarding proving weak convergence results for the proposed estimators of F, one could consider ways of assessing the quality of the proposed estimators. The authors will present methods for obtaining confidence bands by bootstrap methods in a separate paper.

Acknowledgements

We would like thank Søren Asmussen, University of Aarhus, clarifying convergence results for regenerative processes.

References

Asmussen, S. (2003) *Applied Probability and Queues* (2nd edn). New York: Springer-Verlag.

Billingsley, P. (1968) *Convergence of Probability Measures*. New York: Wiley.

Bingham, N.H. and Pitts, S.M. (1999a) Nonparametric estimation for the $M/G/\infty$ queue. *Ann. Inst. Statist. Math.*, **51**, 71–97.

Bingham, N.H. and Pitts, S.M. (1999b) Nonparametric inference from $M/G/1$ busy periods. *Comm. Statist. Stochastic Models*, **15**, 247–272.

Buchmann, B. and Grübel, R. (2003) Decompounding: An estimation problem for Poisson random sums. *Ann. Statist.*, **31**, 1054–1074.

Chiu, S.-T. (2000) Boundary adjusted density estimation and bandwidth selection. *Statist. Sinica*, **10**, 1345–1367.

Feller, W. (1971) *An Introduction to Probability Theory and Its Applications* (2nd edn), Vol. II. New York: Wiley.

Gill, R. (1989) Non- and semi-parametric maximum likelihood estimators and the von Mises method (Part 1). *Scand. J. Statist.*, **16**, 97–128.

Groeneboom, P. (1989) Brownian motion with a parabolic drift and Airy functions. *Probab. Theory Related Fields*, **81**, 79–109.

Grübel, R. and Pitts, S.M. (1993) Non-parametric estimation in renewal theory, I: The empirical renewal function. *Ann. Statist.*, **21**, 1431–1451.

Hall, P. and Park, J. (2004) Nonparametric inference about service time distribution from indirect measurements. *J. Roy. Statist. Soc. Ser. B*, **66**, 861–875.

Hoffmann-Jørgensen, J. (1994a) *Probability with a View Toward Statistics*, Vol. I. New York: Chapman & Hall.

Hoffmann-Jørgensen, J. (1994b) *Probability with a View Toward Statistics*, Vol. II. New York: Chapman & Hall.

Levental, S. (1988) Uniform limit theorems for Harris recurrent Markov chains. *Probab. Theory Related Fields*, **80**, 101–118.

Panjer, H.H. (1981) Recursive evaluation of a family of compound distributions. *ASTIN Bull.*, **12**, 22–26.

Pollard, D. (1984) *Convergence of Stochastic Processes*. New York: Springer-Verlag.
Prakasa Rao, B. (1983) *Nonparametric Functional Estimation*. Orlando, FL: Academic Press.
Resnick, S. (1997) Heavy tail modeling and teletraffic data. *Ann. Statist.*, **25**, 1805–1869.
Sharma, V. (1999) Estimating traffic intensities at different nodes in networks via a probing stream. In *GLOBECOM 1999: Seamless Interconnection for Universal Services*, Vol. 1, pp. 374–380. Piscataway, NJ: IEEE Operations Center.
Sharma, V. and Mazumdar, R. (1998) Estimating traffic parameters in queueing systems with local information. *Perform. Eval.*, **32**, 217–230.
Tsai, T.H. (1998) The uniform CLT and LIL for Markov chains. PhD thesis, University of Wisconsin, Madison.
van der Vaart, A.W. (1998) *Asymptotic Statistics*. Cambridge: Cambridge University Press.
Wand, M.P. and Jones, M. (1995) *Kernel Smoothing*. London: Chapman & Hall.

Received August 2006