Ann. Inst. Statist. Math.
Vol. 51, No. 1, 71–97 (1999)

NON-PARAMETRIC ESTIMATION FOR THE $M/G/\infty$ QUEUE

N. H. BINGHAM[1] AND SUSAN M. PITTS[2]

[1] *Department of Mathematics and Statistics, Birkbeck College, University of London, Malet Street, London, WC1E 7HX, U.K.*
[2] *Department of Pure Mathematics and Mathematical Statistics, University of Cambridge, 16 Mill Lane, Cambridge, CB2 1SB, U.K.*

(Received November 7, 1996; revised October 17, 1997)

Abstract. Given an $M/G/\infty$ queue with input rate λ and service-time distribution G, we consider the problem of estimating λ and G from data on the queue-length process $Q = (Q_t)$. Our motivation is to study departures of G from exponentiality, following recent work of Bingham and Dunham (1997, *Ann. Inst. Statist. Math.*, **49**, 667–679).

Key words and phrases: Infinite-server queue, infinite-dimensional delta-method, empirical process, Little's formula, Reynolds' formula.

1. Introduction

We study inference for the $M/G/\infty$ queueing model, where the input stream is a Poisson point process $Ppp(\lambda)$ with intensity λ and the service-time distribution G is general; we write α for its mean. This is a semi-parametric problem with (λ, G) the object of study; the parametric sub-problem with (λ, α) the parameter of interest will also be studied. For reasons motivated by the applied background to the problem, we take as our data the queue-length process $Q = (Q_t)_{t\geq 0}$. Our problem splits naturally into three parts, depending on whether or not we use (Q_t) itself—*count* data, in which we count customers in the queueing system—or $(I(Q_t = 0))$—*indicator* data, in which we observe only the idle and busy periods. We deal successively with Problems I–III:

I. *Parametric estimation based on counts.* Here we estimate (λ, α), using *Little's formula* ($\mu := EQ_t = \lambda\alpha$);

II. *Non-parametric estimation based on counts.* Here we estimate (λ, G) using *Reynolds' formula* (which identifies the covariance structure of Q in terms of the integrated tail of G);

III. *Non-parametric estimation based on indicators.* Here we use the methods of Grübel and Pitts (1992, 1993) to estimate (λ, G) from the indicator process $(I(Q_t = 0))$—that is, from the idle-busy cycles.

In each case, we obtain appropriate central limit theorems—one-dimensional for I, finite-dimensional for II, functional for III.

Our work complements that of Bingham and Dunham (1997), who consider two related problems:

A. *Parametric estimation for $M/M/\infty$ based on counts*. Here for the special case G exponential, λ, α are estimated from Q_t by Markov-process methods;

B. *Parametric estimation for $M/G/\infty$ based on indicators*. Here λ, α are estimated from data on idle and busy periods, using results on regenerative phenomena.

The work of Bingham and Dunham (1997) was motivated by a problem in statistical mechanics. In this setting, the service-time law G is known to be approximately exponential. It was a desire to probe the accuracy of this exponential approximation by means of the powerful machinery of Grübel and Pitts (1992, 1993) that motivated this study.

We devote Section 2 to theoretical preliminaries and Section 3 to the applied background and discussion of the links with the $M/M/\infty$ model. We deal with Problem I in Section 4 (Theorem 4.1) and Problem II in Section 5 (Theorem 5.1). The more difficult Problem III then follows, using the methods of Grübel and Pitts (1992, 1993): results are in Section 6 (Theorems 6.1, 6.2 and 6.3—strong consistency, functional central limit theorem, and the bootstrap), and the—rather lengthy—proofs in Section 7. We conclude in Section 8 by illustrating our results with some simulation studies and computer graphics.

We work throughout in a continuous-time setting. For a recent study of the corresponding problem in discrete time, see Pickands and Stine (1997).

2. Theoretical preliminaries

Queueing models typically involve distributions of interest arising as functionals of other distributions—perhaps involved in the specification of the model—from which data are observed. One can then use estimators of such distributions, together with properties of the functional, to obtain estimators of the distributions of interest. The method is well exemplified by the $GI/G/1$ queue. Here, the stationary waiting-time law μ_W is regarded (in the stable case, with traffic intensity $\rho < 1$) as a functional of the laws μ_S, μ_T of service and inter-arrival times. The functional approach is developed in Grübel and Pitts (1992, 1993), and non-parametric estimators for the stationary waiting-time law are obtained in Pitts (1994*a*). One passes from properties of one estimator (consistency, asymptotic normality, etc.) to those of the estimator obtained by applying the functional by using local properties of the functional (continuity, differentiability, etc.). For instance, for asymptotic normality we use von Mises' method—the infinite-dimensional version of the familiar 'delta-method'; for background, see e.g. Gill (1989), Gill and van der Vaart (1993), Andersen *et al.* (1993), II.8, van der Vaart and Wellner (1996), §3.9.

Here we take a similar approach to the $M/G/\infty$ queue, specified as above by (λ, G). There are infinitely many servers (so there is no queueing—all customers present are being served). We focus on the 'queue-size process' (or queue-length process) $Q = (Q_t)$, where Q_t is the number of customers present (being served) at time t. Write A_t, D_t for the number of customers arriving and departing in $[0, t]$; thus A_t, D_t are the number of upward and downward jumps of Q in $(0, t]$.

We restrict attention throughout to the case when the system is in equilibrium. We note the distributional properties of the queue-length process Q, of *counts*, which we need for Problems I and II (Sections 4 and 5).

PROPOSITION 2.1. (i) *The distribution of the queue length Q_t in equilibrium is Poisson $P(\mu)$ with parameter $\mu := \lambda\alpha$.*

(ii) *The finite-dimensional distributions of Q are multivariate Poisson (in particular, are infinitely divisible).*

(iii) *The process Q has linear regression.*

Part (i) of this result is *Erlang's formula*; see e.g. Takács (1969) for proof and references (e.g. to Erlang's work of 1917). Note that the limiting distribution of Q involves G only through its mean α, an example of the phenomenon of *insensitivity*; cf. Schassberger (1978), Baccelli and Brémaud (1994), §3.3. We defer discussion of parts (ii) and (iii) to Section 5.

We turn now to the other main approach we shall adopt, in which instead of the process of *counts* $Q = (Q_t)$ we deal with the process $(I(Q_t = 0))$ of *indicators*. The time-axis is decomposed into alternate idle and busy periods (also called *spacings* and *clumps* in the coverage-process literature—see e.g. Hall (1988)) according as $Q_t = 0$ or $Q_t > 0$. By the lack-of-memory property of the exponential law, the spacings have the same law $E(\lambda)$ (exponential with parameter λ) as the inter-arrival time law μ_T. The busy-period (or clump) distribution C depends on λ and G through the following result (Hall (1988), Theorem 2.2), which we need for Problem II (Section 5).

PROPOSITION 2.2. (i) *The mean clump-length γ is given by Smoluchowski's formula*

$$\gamma = EC = (e^{\alpha\lambda} - 1)/\lambda.$$

(ii) *The Laplace-Stieltjes transform $\hat{C}$ of C is given by*

$$(*) \qquad \hat{C}(s) := \int_0^\infty e^{-sx} dC(x)$$

$$= 1 + \frac{s}{\lambda} - \left(\lambda \int_0^\infty e^{-st} \exp\left\{-\lambda \int_0^t (1 - G(x))dx\right\} dt\right)^{-1}.$$

(iii) *The busy-period variance is finite if and only if the service-time variance is finite, and then*

$$\operatorname{var} C = 2e^{\alpha\lambda}\lambda^{-1} \int_0^\infty \left(\exp\left\{\lambda \int_t^\infty (1 - G(x))dx\right\} - 1\right) dt - (e^{\alpha\lambda} - 1)^2/\lambda^2.$$

There remains Problem III, the hardest. Here we are to estimate G given only the *indicator process* of Q, namely $(I(Q_t = 0))$—that is, given only the busy and idle periods—and use the results of Proposition 2.2. Thus we are to study the functional

$$(\lambda, C) \to G$$

of (*); we do this in Section 5 by an approach modelled on that of Grübel and Pitts (1992, 1993), Pitts (1994*a*) for renewal theory and the $GI/G/1$ queue.

We pause to introduce some notation. Recall that

$$\alpha := \int_0^\infty x dG(x) = \int_0^\infty (1 - G(x))dx$$

is the mean of G. Write

$$H(x) := \int_0^x (1 - G(u))du$$

for the integrated tail of G (it is H, rather than G, that appears in Proposition 2.2). Thus the normalised integrated tail of G,

$$G^*(x) := \frac{1}{\alpha} H(x) = \frac{1}{\alpha} \int_0^x (1 - G(u))du,$$

is a probability distribution (the stationary lifetime distribution of G, in the language of renewal theory). It turns out that it is convenient and natural to focus on H or G^* rather than G itself. We thus focus on functionals such as

$$(\lambda, G) \to (\alpha, H), \qquad (\lambda, C) \to (\alpha, G^*).$$

Note that $G^* = G$ if and only if G is exponential, the $M/M/\infty$ case.

We note in passing that the intensity λ may be estimated easily from the idle periods, since these are exponential $E(\lambda)$; the service-time law G, our main object of interest, is much harder to estimate.

Part of the background to this work is the extent to which our $M/G/\infty$ model may be approximated by an $M/M/\infty$ model with the same means. Here G is exponential with mean α, $G = E(1/\alpha)$, and Q is a birth-and-death process, so Markov—the *first-order equivalent* birth-and-death process, in the language of Baccelli and Brémaud (1994), §4.1. Under this simplifying approximation, the parameters α, λ can (Problem A of Section 1) be estimated by standard maximum-likelihood methods for Markov processes (Billingsley (1961), Example 7.2) as

$$\hat{\lambda}(t) = A_t/t, \qquad \hat{\alpha}(t) = \frac{1}{D_t} \int_0^t Q_u du,$$

the relevant *occurrence-exposure ratios*, for which see e.g. Andersen *et al.* (1993), Chapter VI. For details, and the relevance of the $M/M/\infty$ model, we refer to Bingham and Dunham (1997).

Working with counts rather than indicators corresponds to observing the queue-size process—and so, all arrival and departure epochs—but *not* observing *which* customer leaves at a departure epoch. Of course, if we keep track of which departure epoch corresponds to which arrival epoch, we can observe the service times directly, and then estimate G directly—without any use of the queueing model—by empirical-process methods (see e.g. Shorack and Wellner (1986), van der Vaart and Wellner (1996)). However, in experimental settings such as those that motivated this work (discussed in Section 3 below), it may be much harder to *track* an individual particle over time than to *count* the number of particles present as a function of time. It is thus of prime experimental importance that we pay no heed to the individuality of different particles, but merely count them. The same viewpoint is adopted in earlier work on this subject by Brown (1970), in the setting of light traffic problems in the theory of road traffic. Another similar setting is that of Nozari and Whitt (1988), who consider an industrial production setting. Here it may be easy to count jobs in hand (WIP, or 'work in progress'), rather than keep track of the starting and finishing times of individual jobs.

3. Applied background and the $M/M/\infty$ model

The problem that motivated our work arises in statistical mechanics; see Bingham and Dunham (1997) for a full description. Particles in suspension move according to the Ornstein-Uhlenbeck dynamics

$$\text{(OU)} \qquad dV_t = -\beta V_t dt + c dW_t;$$

here $V = (V_t)$ is the velocity process of a particle, $W = (W_t)$ is standard Brownian motion, $1/\beta$ is the *relaxation time*, and $D := \frac{1}{2}c^2/\beta^2$ is the *diffusion coefficient*. The limiting velocity distribution is then the familiar *Maxwell-Boltzmann distribution* of statistical mechanics, $N(0, \beta D)$, and we can make V stationary by starting it in this distribution. Integrating, one obtains the Ornstein-Uhlenbeck displacement process $X = (X_t)$, $X_t := x_0 + \int_0^t V_u du$. If $I = [a, b]$ is an interval on the line, the distribution of the *occupation-time* T between entry of X into I and first subsequent exit from I (we need to average the velocity of entry over the limiting Maxwell-Boltzmann distribution to make this law well-defined) has been much studied. This law is not known explicitly, but various asymptotic properties are known; see e.g. Doering *et al.* (1989*a*, 1989*b*), Hesse (1991). One question of particular interest about this law is the extent to which it is approximately exponential.

The above relates to the dynamics of an individual particle. In the setting of statistical physics, however, we will have a population of similar particles to observe. As mentioned earlier, it is much easier experimentally to *count* particles than to keep track of an individual particle over time. Accordingly, one may seek to extract information on particle characteristics from observations on the counting process (Q_t)—known as a *Smoluchowski process* in this context—a technique known as *number fluctuation spectroscopy*. A classical instance of this was the *Einstein-Smoluchowski theory of diffusion*, where *Avogadro's number* was studied experimentally by counting numbers Q_t of particles in suspension present in some small region of observation at time t. The key parameter of interest here is $\alpha = ET$, as $1/\alpha$ is a measure of the *mobility* of the particles. A similar approach is used in studies of spermatazoa, leukocytes and the like. For details and references, see e.g. Bingham and Dunham (1997).

Now the occupation-time law G corresponds to the service-time law in the $M/G/\infty$ queueing model, or the segment-length law in the coverage-process model.

There are good probabilistic reasons for thinking that G might be approximately exponential; see e.g. Bingham and Dunham (1997), §6.3. This is confirmed by the analytic approximations obtained by Doering *et al.* (1989*a*, 1989*b*), Hesse (1991). Additionally, modelling (Q_t) as a birth-and-death process, with birth and death rates

$$\lambda_n \equiv \lambda, \quad \mu_n = n/\alpha \quad \text{if } n > 0, \qquad 0 \quad \text{if } n = 0,$$

has obvious intuitive appeal ($\lambda_n \equiv \lambda$ reflects the Poisson input stream, $\mu_n = n/\alpha$ reflects the propensity of each particle present to leave at rate $1/\alpha$, the particle mobility). This again leads to G exponential, and reduces $M/G/\infty$ to $M/M/\infty$. This $M/M/\infty$ model is now fully solved (Bingham and Dunham (1997)). A desire to assess the relevance of these conclusions to the general case—that is, of the dependence of the $M/G/\infty$ model on G, particularly when G is close to exponential—motivates our treatment of Problems II and III.

4. Count data: parametric approach via Little's formula

Little's formula is one of the most important general principles of queueing theory. It holds under very general conditions, and is often stated acronymically as '$L = \lambda W$'—mean queue-length is the product of the input intensity and the mean waiting time. For an excellent recent textbook treatment, see Baccelli and Brémaud (1994), §3.1, and for a survey and further references, see also Whitt (1991). Extensions of the formula—ordinary and functional central limit theorems, etc.—have been given by Glynn and Whitt (1986, 1988, 1989), who also give applications to estimation of parameters in queueing models. In the $M/G/\infty$ case, Little's formula says that

$$\mu := EQ_t = \lambda\alpha,$$

which is part of Proposition 2.1(i) (for the more general Campbell-Little-Mecke formula, see Baccelli and Brémaud (1994), §3.2.1).

We proceed as follows:

(a) estimate λ by

$$\hat{\lambda}_t := A_t/t,$$

the occurrence-exposure ratio based on the arrivals by time t;

(b) estimate μ by the sample mean of Q,

$$\hat{\mu}_t := \frac{1}{t}\int_0^t Q_u du;$$

(c) estimate α by

$$\hat{\alpha}_t := \hat{\mu}_t/\hat{\lambda}_t = \frac{1}{A_t}\int_0^t Q_u du.$$

THEOREM 4.1. *The estimator $\hat{\alpha}_t$ is strongly consistent*:

$$\hat{\alpha}_t \to \alpha \quad (t \to \infty) \qquad a.s.$$

If the service-time law G has finite variance σ^2, one has asymptotic normality:

$$\sqrt{t}(\hat{\alpha}_t - \alpha) \to_d N(0, \sigma^2/\lambda).$$

PROOF. Strong consistency of $\hat{\lambda}_t$ follows by the strong law for renewal theory (see e.g. Billingsley (1979), 23.12), and of $\hat{\mu}_t$ by Birkhoff's ergodic theorem (see e.g. Krengel (1985), §1.2); that for $\hat{\alpha}_t$ follows from this.

Asymptotic normality follows from Theorem 1 of Glynn and Whitt (1988) by the delta method (first-order Taylor expansion: Billingsley (1979), §29; Rao (1973), §6a.2). First, the independence of the inter-arrival times A_n (which are exponential $E(\lambda)$, so with mean $1/\lambda$ and variance $1/\lambda^2$) and service times W_n (which have law G, with mean α and variance σ^2) gives the joint central limit theorem required by the condition (1.1) of Glynn and Whitt (1988) for their Theorem 1. The second and eighth components of this result give the joint central limit theorem

$$t^{-1/2}\left(A_t - \lambda t, \int_0^t Q_u du - \lambda\alpha t\right) \to_d (\lambda^{3/2}U, \lambda^{1/2}(W - \lambda\alpha U),$$

where U, W are independent, $N(0, 1/\lambda^2)$ and $N(0, \sigma^2)$ respectively. A simple application of the delta method (Billingsley (1979), Example 29.1 with $f(x, y) := y/x$) gives

$$\sqrt{t}\left(\frac{1}{A_t}\int_0^t Q_u du - \alpha\right) \to_d N(0, \sigma^2/\lambda). \qquad \square$$

Remark 1. We can replace arrivals A_t by departures D_t in the above, on using the third component of the Glynn-Whitt result instead of the second. Then α is estimated by $\int_0^t Q_u du/D_t$, an exposure-occurrence ratio; see Section 3 and Bingham and Dunham (1997).

Remark 2. This approach via Little's formula is considered in some detail by Glynn and Whitt (1986, 1988, 1989). Their setting is variance reduction in queueing simulation, motivated by the fact that there λ is often known, and using λ rather than estimates of it may greatly increase efficiency. By contrast, they show that when λ is unknown, as here, there is no gain or loss of efficiency in estimating α directly, from service times, or indirectly, from Little's law as here. This is interesting, as it tells us that, when estimating α, there is no loss of efficiency in not keeping track of individual customers, but observing only the queue-length process Q.

5. Count data: non-parametric approach via Reynolds' formula

We turn now to the estimation of the non-parametric component G of the $M/G/\infty$ model, using count data (Q_t).

Recall from Proposition 2.1(i) that the one-dimensional distributions of the Smoluchowski process Q are Poisson $P(\mu)$. The finite-dimensional distributions are, from (ii), multivariate Poisson; for details, see Lindley (1956) §2, Vere-Jones

(1968), Milne (1970), §2, Bartlett (1978), §3.4. The regression of Q_{t+h} on Q_t is, from (iii), linear in Q_t; see Reynolds (1972), §2, Bartlett (1978), §6.31.

We saw that the approximation of $M/G/\infty$ by $M/M/\infty$ amounted to the use of the first-order equivalent birth-and-death process. It turns out that for our purposes, *second-order* information—that is, covariance or correlation structure—suffices. The key fact is that the second-order distributions of Q—its correlation structure—encode the service-time distribution we wish to estimate by the following result, *Reynolds' formula*: if

$$\gamma(h) := \operatorname{cov}(Q_{t+h}, Q_t)$$

is the covariance function,

$$\rho(h) := \gamma(h)/\gamma(0)$$

the correlation function, then

$$\rho(h) \equiv 1 - G^*(h)$$

(Reynolds (1968), (1972), §2, (1975), §2, following Riordan (1951), Beneš (1957)). This identifies the autocorrelation function ρ as the tail of the normalized integrated tail-function of G. Note in particular that the autocorrelation function ρ is non-negative.

Since Q is stationary, Birkhoff's ergodic theorem gives

$$\frac{1}{T}\int_0^T Q_{t+h}Q_t dt \to \rho(h) \quad (T \to \infty) \qquad \text{a.s.}$$

(Krengel (1985), §1.2: the independence of the service- and inter-arrival times, and Kolmogorov's zero-one law, show that the tail σ-field is trivial). Thus for fixed $h > 0$, $\rho(h)$, so $G^*(h)$, may be estimated with arbitrary accuracy from a sufficiently long segment $\{Q_t : 0 \le t \le T\}$ of a realization of Q.

Again, we have a central limit theorem. We present this in discrete time, partly for mathematical convenience, partly because, in practice, our output will be a graph interpolating sample values of the correlation function r—estimating the population correlation function $\rho = 1 - G^*$—at a discrete set of chosen points. We may take these equally spaced—at intervals $h > 0$. For the purposes of this section, we use suffix notation for discrete arguments, bracket notation for continuous ones: thus $Q_i := Q(hi)$, etc.

THEOREM 5.1. *When the service-time law G has finite variance σ^2, the sample correlations*

$$r_j := r(jh) = \frac{\frac{1}{n}\sum_1^{n-j}(Q_i - \bar{Q})(Q_{i+j} - \bar{Q})}{\frac{1}{n}\sum_1^n (Q_i - \bar{Q})^2}, \qquad \bar{Q} = \frac{1}{n}\sum_{i=1}^n Q_i$$

are jointly asymptotically normal, with means $\rho_j := \rho(jh)$:

$$\sqrt{n}(r_j - \rho_j)_{j=1}^s \to_d N(0, W),$$

where the covariance matrix $W = (w_{ij})_{i,j=1}^{s}$ *is given by*

$$w_{ij} = \sum_{r=1}^{\infty}\{\rho_{r+i} + \rho_{r-i} - 2\rho_r\rho_i\}\{\rho_{r+j} + \rho_{r-j} - 2\rho_r\rho_j\}.$$

PROOF. We have

$$\int_0^\infty \rho(t)dt = \int_0^\infty (1 - G^*(t))dt = \frac{1}{\alpha}\int_0^\infty dt \int_t^\infty [1 - G(x)]dx = \frac{1}{2\alpha}\int_0^\infty u^2 dG(u),$$

by Fubini's theorem. Thus G has finite variance if and only if $\int_0^\infty \rho(t)dt < \infty$: $\rho \in \ell_1$. Since ρ is bounded, being a correlation, this gives $\rho \in \ell_2$ also. A similar calculation using sums instead of integrals shows that G has finite variance if and only if $\sum_0^\infty \rho_n < \infty$, i.e. $(\rho_n) \in \ell_1$. Since $(\rho_n) \in \ell_\infty$, this gives $(\rho_n) \in \ell_2$. By Parseval's formula, (ρ_n) is thus the sequence of Fourier coefficients of a function $f(\lambda)$, which is in L_2, and

$$\sum |\rho_n|^2 = \frac{1}{2\pi}\int_0^{2\pi} |f(\lambda)|^2 d\lambda.$$

In the language of Doob (1953), X.8, $f(\lambda)$ is the spectral density of the stationary process (Q_n); $f(\lambda) = |c(\lambda)|^2$, where $c(\lambda)$ is the sum-function of the Fourier series of the coefficients c_n in the moving-average representation

$$Q_n = \sum c_j \xi_{n+j} \qquad \Big((\xi)_n \text{ orthogonal, } \sum |c_n|^2 < \infty\Big)$$

of (Q_n). Asymptotic normality with the stated covariance matrix (whose form results from the moving-average representation) now follows from Theorem 2 of Hannan (1976), the necessary and sufficient condition for which is that the spectral density be square-integrable. □

For further background, see e.g. Hannan (1970), Hannan and Heyde (1972), Hall and Heyde (1980), §6.4. The condition of Hannan and Heyde (1972)—that the best linear predictor is the best predictor—holds here since Q has linear regression (Proposition 2.1(iii)). Results of this form stem from work of Bartlett in 1946; see e.g. Bartlett (1955/78), §9.1.

Theorem 5.1 tells us that, if we wish to obtain a plot of the graph of $\rho = 1 - G^*$, we may choose an interval $h > 0$ between points and a number s of points to plot, then use data $Q_n := Q(rh)$, $0 \le r \le n$ to calculate sample correlations $(r_j)_{j=1}^{s}$; these estimate the population correlations $\rho_j := \rho(jh) = 1 - G^*(jh)$ with the usual rate $\sqrt{n}$ and normal limits.

From a theoretical point of view, it would be desirable to supplement this with explicit confidence intervals, particularly for the maximum discrepancy between sample and population correlations over any chosen range. One might seek a functional central limit theorem for this purpose, but the covariance structure of the limiting Gaussian process is complicated, and we know of no such results.

However, an almost-sure rate of convergence is known in this context: see Hannan and Kavalieris (1983).

Since G^* is monotone, so is ρ by Reynolds' formula. One may thus seek to improve the accuracy of our plot of sample estimates of ρ by using monotone regression; see Chapter 1 of Barlow *et al.* (1972) for background and details. This procedure is easily programmed and is conveniently packaged in, e.g., Genstat 5 Release 3 (Genstat 5 Committee (1993)).

Since G is exponential if and only if G^* is, the accuracy of the $M/M/\infty$ approximation to $M/G/\infty$ may be measured from the closeness of $\rho = 1 - G^*$ to exponential. We may thus use the closeness of our plot of $-\log r$ to linearity to assess the closeness of this approximation (which, as we mentioned in Section 2, was the original motivation for this study).

6. Indicator data: results

We now consider the non-parametric estimation of G, discarding all the information in the queue-length process Q except whether or not any customers are present—that is, using only the idle and busy periods. It is remarkable that satisfactory results—including central limit theorems—can still be obtained.

For Problem III, where we have $(I(Q_t = 0))$ instead of (Q_t), we take our data to be independent random samples of the busy and idle periods. Thus let $Y_1, Y_2, \ldots$ be independent identically distributed positive random variables, the busy periods, with distribution function C, and let $Z_1, Z_2, \ldots$ be independent exponentially distributed random variables, the idle periods, with mean λ^{-1}, independent of $\{Y_i\}$. Our aim is to estimate $H(x) := \int_0^x (1 - G(t))dt$ and G using these data.

We take a functional view and express the quantities of interest H and G in terms of λ and C. Equation $(*)$ expresses a relationship between H and (λ, C) which, when rearranged, is

$$(6.1) \qquad \lambda \int_0^\infty e^{-st} \exp\{-\lambda H(t)\} dt = (1 + s/\lambda)^{-1}(1 - \hat{C}(s)(1 + s/\lambda)^{-1})^{-1}.$$

Write $F \star G(t)$ for $\int_{[0,t]} F(t - x)dG(x)$, and observe that $\hat{C}(s)(1 + s/\lambda)^{-1}$ is the Laplace-Stieltjes transform of $F := E_\lambda \star C$, the distribution function of the sum of a busy period and an idle period. Let $U = \sum_{k=0}^{\infty} F^{\star k}$ be the renewal function associated with F, where $F^{\star 0}$ is the indicator function $I_{[0,\infty)}$ of the set $[0, \infty)$, and for $k \geq 1$, $F^{\star k} = F \star F^{\star (k-1)}$. Then (6.1) yields $\lambda \exp\{-\lambda H(t)\} = e_\lambda \star U(t)$, or

$$(6.2) \qquad H(t) = -\frac{1}{\lambda} \log_e \left(\frac{e_\lambda \star U(t)}{\lambda} \right).$$

Hence H is determined by λ and C and we write $H = \Phi(\lambda, C)$.

Since F has a density $f = e_\lambda \star C$, it follows that $U - I_{[0,\infty)}$ has a density u, called the renewal density. Differentiating (6.2), and using $(e_\lambda \star U)'(t) = \lambda(u - e_\lambda \star U)(t)$, we find

$$1 - G(t) = 1 - \frac{u(t)}{e_\lambda \star U(t)},$$

Lebesgue-almost everywhere. Let $\tilde{G}$ and Ψ be defined by

$$\tilde{G} = \Psi(\lambda, C) = \frac{u}{e_\lambda \star U},$$

then $\tilde{G} = G$ almost everywhere. (Since G is right-continuous we can recover G from $\tilde{G}$; it is more convenient for current purposes to work with $\tilde{G}$. We later take a specific version of u.)

Given the data, we define plug-in estimators $\hat{H}_n$ and $\hat{\tilde{G}}_n$ of H and $\tilde{G}$ respectively, by

$$\hat{H}_n = \Phi(\hat{\lambda}_n, \hat{C}_n) \qquad \hat{\tilde{G}}_n = \Psi(\hat{\lambda}_n, \hat{C}_n),$$

where $\hat{\lambda}_n = (n^{-1}\sum_{i=1}^n Z_i)^{-1}$, and $\hat{C}_n = n^{-1}\sum_{i=1}^n I_{[Y_i,\infty)}$, the empirical distribution function based on $Y_1, \ldots, Y_n$.

Statistical properties, such as strong consistency and asymptotic normality, of $\hat{\lambda}_n$ and $\hat{C}_n$ as estimators of λ and C respectively, are known. In Section 7 we establish continuity and an appropriate differentiability property for Φ and Ψ. These local properties of the functionals ensure that strong consistency and asymptotic normality carry over from the input estimators $\hat{\lambda}_n$ and $\hat{C}_n$ to the output estimators $\hat{H}_n$ and $\hat{\tilde{G}}_n$; for asymptotic normality, this is the delta or von Mises method, see Gill (1989).

Before stating strong consistency and asymptotic normality results for $\hat{H}_n$ and $\hat{\tilde{G}}_n$, we define D_∞ to be the space of real-valued right-continuous functions f on $[0,\infty]$, with left-hand limits that are left-continuous at infinity. A real-valued function on $[0,\infty)$ that is right-continuous with left-hand limits, and a finite limit at infinity may be extended to an element of D_∞. Write $\|\cdot\|_\infty$ for the supremum norm. The theorem below gives strong consistency of our estimators.

THEOREM 6.1. *Assume $\int x^2 dC(x) < \infty$. Then, with probability one, as* $n \to \infty$,

$$\text{(i)}\ \|\hat{H}_n - H\|_\infty \to 0, \qquad \text{(ii)}\ \|\hat{\tilde{G}}_n - \tilde{G}\|_\infty \to 0.$$

The next theorem gives asymptotic normality of the estimators in terms of convergence in distribution to a Gaussian process in D_∞. Here we follow Pollard (1984), Chapter IV, for convergence in distribution in a metric space, giving D_∞ its open ball σ-field. We write "$\to_d$" for convergence in distribution.

THEOREM 6.2. *Assume $\int x^{2\gamma} dC(x) < \infty$ for some $\gamma > 2$. Then, in D_∞, as* $n \to \infty$,

$$\text{(i)}\ \sqrt{n}(\hat{H}_n - H) \to_d Z_H, \qquad \text{(ii)}\ \sqrt{n}(\hat{\tilde{G}}_n - \tilde{G}) \to_d Z_G,$$

where Z_H and Z_G are zero mean Gaussian processes.

A natural next step is to find confidence bands for the unknown functions. In this discussion, we follow that of Grübel and Pitts (1993). Let

$$R_n(t) = P(\sqrt{n}\|\hat{H}_n - H\|_\infty \le t) \quad \text{and} \quad R(t) = P(\|Z_H\|_\infty \le t).$$

For $0 < \alpha < 1$, if $q_n(\alpha)$ is such that $R_n(q_n(\alpha)) = \alpha$, then $\hat{H}_n \pm n^{-1/2}q_n(\alpha)$ is an exact $100\alpha\%$ confidence band for H. However R_n is not known. Theorem 6.2 implies that $R_n(t) \to R(t)$ for all continuity points t of R, and so asymptotic confidence bands could be obtained using quantiles of R. Unfortunately, R turns out to have a complicated dependence on the unknown λ and C. We take an alternative approach, and use the bootstrap estimator $\hat{R}_n$ for R_n. This is constructed so that $\hat{R}_n$ depends on $\hat{C}_n$ and $\hat{\lambda}_n$ in exactly the same way as R_n depends on C and λ.

First, we give an explicit representation for R_n in terms of C and λ. Let $\mathbb{F}_n : \mathbb{R}^n \to D_\infty$ be defined by $\mathbb{F}_n(\boldsymbol{x}) = n^{-1}\sum_{i=1}^n I_{[x_i,\infty)}$ for $\boldsymbol{x} = (x_1, \dots, x_n)$ in $\mathbb{R}^n$. Write $C^{\otimes n}$ for the n-th measure-theoretic power of C. Then

$$R_n(t) = \int_{\mathbb{R}^n}\int_{\mathbb{R}^n} I_{[0,t]}\{\sqrt{n}\|\Phi(\Phi_0(\mathbb{F}_n(\boldsymbol{z})), \mathbb{F}_n(\boldsymbol{y})) - \Phi(\lambda, C)\|_\infty\}dC^{\otimes n}(\boldsymbol{y})dE_\lambda^{\otimes n}(\boldsymbol{z}),$$

where $\Phi_0(F) = (\int x dF(x))^{-1}$, so that $\Phi_0(\mathbb{F}_n(\boldsymbol{z})) = (n^{-1}\sum_{i=1}^n z_i)^{-1}$. The bootstrap estimator $\hat{R}_n$ of R_n is defined by replacing C and λ in the above expression by $\hat{C}_n$ and $\hat{\lambda}_n$ respectively to get

$$\hat{R}_n(t) = n^{-n}\sum_{\boldsymbol{i}\in\mathcal{I}_n} \cdot \int_{\mathbb{R}^n} I_{[0,t]}\{\sqrt{n}\|\Phi(\Phi_0(\mathbb{F}_n(\boldsymbol{z})), \mathbb{F}_n(Y_{i_1}, \dots, Y_{i_n})) - \Phi(\hat{\lambda}_n, \hat{C}_n)\|_\infty\}dE_{\hat{\lambda}_n}^{\otimes n}(\boldsymbol{z}),$$

where $\boldsymbol{i} = (i_1, \dots, i_n)$ and $\mathcal{I}_n = \{1, \dots, n\}^n$. Here we have a combined 'non-parametric' and 'parametric' bootstrap, in contrast to Grübel and Pitts (1993) and Pitts (1994*a*), where only the non-parametric bootstrap is involved. Let $\hat{q}_{H,n}(\alpha)$ be the α-quantile of $\hat{R}_n$. Let S_n, S and $\hat{S}_n$ be the quantities corresponding to R_n, R and $\hat{R}_n$ when Φ is replaced by Ψ and H by $\tilde{G}$, and let $\hat{q}_{G,n}(\alpha)$ be the α-quantile of $\hat{S}_n$. Then our final theorem shows that 'the bootstrap works.'

THEOREM 6.3. *Suppose that* $0 < \alpha < 1$. *Assume that* $\int x^{2\gamma}dC(x) < \infty$ *for some* $\gamma > 2$. *Then, as* $n \to \infty$,

$$\text{(i)}\quad P(\|\sqrt{n}(\hat{H}_n - H)\|_\infty \le \hat{q}_{H,n}(\alpha)) \to \alpha,$$

$$\text{(ii)}\quad P(\|\sqrt{n}(\hat{\tilde{G}}_n - \tilde{G})\|_\infty \le \hat{q}_{G,n}(\alpha)) \to \alpha.$$

Moment conditions. In Theorem 6.1, our moment condition is finite variance, which is natural and that used in Theorems 4.1 and 5.1. By contrast, Theorems 6.2 and 6.3 require finite $(4+\epsilon)$-th moments for some $\epsilon > 0$. This condition derives from results of Grübel and Pitts (1993), Proposition 3.15 and Theorems 2.2 and 2.3; we will not pursue the question of weakening it here. However, we suspect that the condition can be dropped if we are content to restrict ourselves to estimation of H, G^* on a compact set (as in Grübel and Pitts (1993), §4.3)—as we have already done in Theorem 5.1.

Extensions and further work. The confidence bands (confidence regions in D_∞) obtained here are constant width. One possible way to have non-constant width bands is to establish conditions for the functional to map into a D_β-space, and to rework the analysis of the functional in this case. A related direction for further research is to investigate weight functions for the output processes using the approach of Csörgő and Zitikis (1996).

7. Indicator data: proofs

7.1 *Preliminaries*

We now define weighted versions of D_∞, as in Grübel and Pitts (1993). For $\beta \geq 0$ and $f : [0,\infty) \to \mathbb{R}$, let $T_\beta f$ be given by

$$(T_\beta f)(x) = (1+x)^\beta f(x), \qquad x \geq 0.$$

Let D_β be the set of all $f : [0,\infty) \to \mathbb{R}$ such that $T_\beta f$ is (extendable to) an element of D_∞. For f in D_β, write $\|f\|_\beta$ for $\|T_\beta f\|_\infty$. In the following proofs, we write $\|f\|_\beta$ for $\sup_t |(T_\beta f)(t)|$ for a function f with $T_\beta f$ bounded, but not necessarily in D_β.

The next lemma relates $\|\cdot\|_\beta$-norms and convolution. Let $\mathcal{V}$ be the set of all real-valued functions V on $[0,\infty)$, that are right-continuous and nondecreasing.

LEMMA 7.1. *Assume $V \in \mathcal{V}$, $f : [0,\infty) \to \mathbb{R}$. Let $\beta \geq 0$. Then*

$$\|f \star V\|_\beta \leq 2^\beta \|f\|_\beta \{\| \|V\|_\infty I_{[0,\infty)} - V\|_\beta + \|V\|_\infty\}.$$

PROOF. See Pitts (1994*b*), Lemma 2.3. □

We also need the space L^1 of functions $f : [0,\infty) \to \mathbb{C}$ with $\|f\| = \int |f(t)|dt < \infty$. For f and g in L^1, define $f*g$ by $f*g(t) = \int_0^\infty f(t-x)g(x)dx$. Then $(L^1, \|\cdot\|, *)$ is a commutative Banach algebra without a unit. Let $L = \{(f,\alpha) : f \in L^1, \alpha \in \mathbb{C}\}$ be the space that results when we append a unit element to L^1. We write δ_0 for the unit element $(0,1)$, and $f + \alpha\delta_0$ for (f,α).

We need the following two results. If $\beta > 1$, then

$$\|f_n - f\|_\beta \to 0 \Rightarrow \|f_n - f\| \to 0. \tag{7.1}$$

If h is a bounded function and g is in L then

$$\|h_n - h\|_\infty \to 0 \quad \text{and} \quad \|g_n - g\| \to 0 \Rightarrow \|h_n * g_n - h * g\|_\infty \to 0. \tag{7.2}$$

7.2 *Proof of Theorem 6.1*

The main part of this proof is to show that the functionals Φ and Ψ are continuous. Suppose that λ and $\{\lambda_n\}_{n=1}^{\infty}$ are positive numbers, and that C and $\{C_n\}_{n=1}^{\infty}$ are distribution functions concentrated on $(0, \infty)$. We show that, if $C_n - C \to 0$ in an appropriate D-space and $\lambda_n \to \lambda$, then $\Phi(\lambda_n, C_n) \to \Phi(\lambda, C)$ in D_∞, and similarly for Ψ.

The key step is a continuity result for the functional taking a probability density function onto the corresponding renewal density, given in Proposition 7.4 below. First we need the following definition. For $f : [0, \infty) \to \mathbb{R}$ with $\int |f(x)|dx < \infty$, define $\Sigma f : [0, \infty) \to \mathbb{R}$ by $(\Sigma f)(t) = \int_t^\infty f(x)dx$. Let $\Sigma(f + \alpha\delta_0) = \Sigma f$. Then $\|\Sigma f\|_\infty \leq \|f\|$, and if $\int_0^\infty x|f(x)|dx < \infty$ then Σf is in L^1. The next lemma collects some results about convergence of f_n, Σf_n and $\Sigma\Sigma f_n$ $(= \Sigma(\Sigma f_n))$.

LEMMA 7.2. *Let f and $\{f_n\}_{n=1}^{\infty}$ be probability density functions on $[0, \infty)$ with $\|f_n - f\|_\infty \to 0$ as $n \to \infty$. Then*

$$\text{(i)} \quad \|f_n - f\| \to 0 \qquad \textit{and} \qquad \text{(ii)} \quad \|\Sigma f_n - \Sigma f\|_\infty \to 0.$$

In addition, let $m_{1,n} = \int x f_n(x)dx < \infty$ for all n, $m_1 = \int x f(x)dx < \infty$, and suppose $m_{1,n} \to m_1$ as $n \to \infty$. Then

$$\text{(iii)} \quad \|\Sigma f_n - \Sigma f\| \to 0 \qquad \textit{and} \qquad \text{(iv)} \quad \|\Sigma\Sigma f_n - \Sigma\Sigma f\|_\infty \to 0.$$

PROOF. (i) is Scheffé's Theorem, see Billingsley (1968), p. 224, and (ii) follows from (i). (iii) follows from (ii) and Theorem 1 in Pratt (1960), and (iii) implies (iv). □

The next lemma gives a representation of the normalised renewal measure, involving inverses of elements of L. An element x in L has an inverse if there exists an element $x^{*(-1)}$ in L such that $x * x^{*(-1)} = \delta_0$.

LEMMA 7.3. *Let f be a probability density function on $[0, \infty)$ with $\int x^2 f(x)dx < \infty$, and with associated renewal density u. Then, as elements of L,*

$$u + \delta_0 - \frac{1}{m_1} = \frac{1}{m_1}(\Sigma\Sigma f - \Sigma f + m_1\delta_0) * (\Sigma f - f + \delta_0)^{*(-1)}. \tag{7.3}$$

PROOF. For proving equality of elements of L, see Grübel and Pitts (1992), Section 3. If $f + \alpha\delta_0$ is in L, let $(f + \alpha\delta_0)\tilde{}(\theta) = \int_0^\infty e^{i\theta x} f(x)dx + \alpha$. From Grübel (1989), Section 3, we have that $u + \delta_0 - (1/m_1)$ is in L with

$$\left(u + \delta_0 - \frac{1}{m_1}\right)^{\sim}(\theta) = \frac{m_1 - (\Sigma f)\tilde{}(\theta)}{m_1(1 - \tilde{f}(\theta))},$$

for $\theta \neq 0$. From the proof of Theorem 1 in Grübel (1986), we have

$$(\Sigma f - f + \delta_0) * (\delta_0 - e_1) = \delta_0 - f,$$

and similarly we obtain

$$(\Sigma\Sigma f - \Sigma f + m_1\delta_0) * (\delta_0 - e_1) = m_1\delta_0 - \Sigma f.$$

Hence for $\theta \neq 0$,

$$\left(u + \delta_0 - \frac{1}{m_1}\right)\tilde{}(\theta) = \frac{(\Sigma\Sigma f - \Sigma f + m_1\delta_0)\tilde{}(\theta)}{m_1(\Sigma f - f + \delta_0)\tilde{}(\theta)}.$$

This also holds for $\theta = 0$ by continuity. Finally the δ_0-parts of both sides of (7.3) are equal to δ_0. □

This representation is used in the next result, which gives uniform convergence of renewal densities.

PROPOSITION 7.4. *Let f and $\{f_n\}_{n=1}^{\infty}$ be probability density functions on $[0, \infty)$ satisfying $\int x^2 f_n(x)dx < \infty$ for all n and $\int x^2 f(x)dx < \infty$. Assume that $m_{1,n} = \int x f_n(x)dx \to m_1 = \int x f(x)dx$ and $\|f_n - f\|_\infty \to 0$ as $n \to \infty$. Then there exist versions u_n and u of renewal densities associated with f_n and f such that*

$$\|u_n - u\|_\infty \to 0 \quad \text{as} \quad n \to \infty.$$

PROOF. Let

$$g = (\Sigma f - f + \delta_0)^{*(-1)} - \delta_0 \quad \text{in} \quad L. \tag{7.4}$$

Then g is in L^1 and there is a version g such that

$$g(t) = -(\Sigma f - f)(t) - (\Sigma f - f) * g(t) \quad \text{for all} \quad t. \tag{7.5}$$

Similarly let g_n be a version of $(\Sigma f_n - f_n + \delta_0)^{*(-1)} - \delta_0$ such that $g_n(t) = -(\Sigma f_n - f_n)(t) - (\Sigma f_n - f_n) * g_n(t)$ for all t. Since the map $x \mapsto x^{*(-1)}$ is $\|\cdot\|$-$\|\cdot\|$ continuous at invertible $x \in L$, we know that $\|g_n - g\| \to 0$, by Lemma 7.2. Using Lemma 7.3, let u be given by

$$u - m_1^{-1} = m_1^{-1}(\Sigma\Sigma f - \Sigma f) + g + m_1^{-1} g * (\Sigma\Sigma f - \Sigma f),$$

and similarly for $u_n - m_{1,n}^{-1}$. Then, for all t,

$$\begin{aligned}
&\left|\left(u_n(t) - \frac{1}{m_{1,n}}\right) - \left(u(t) - \frac{1}{m_1}\right)\right| \\
&\quad \leq \left\|\frac{1}{m_{1,n}}(\Sigma\Sigma f_n - \Sigma f_n) - \frac{1}{m_1}(\Sigma\Sigma f - \Sigma f)\right\|_\infty + \|g_n - g\|_\infty \\
&\quad + \left\|g_n * \frac{1}{m_{1,n}}(\Sigma\Sigma f_n - \Sigma f_n) - g * \frac{1}{m_1}(\Sigma\Sigma f - \Sigma f)\right\|_\infty .
\end{aligned}$$

The first term on the right-hand-side tends to zero by Lemma 7.2. The third term tends to zero using $\|g_n - g\| \to 0$ and (7.2). Similar methods work for the second term, on using (7.5). The proposition now follows since $m_{1,n}^{-1} \to m_1^{-1}$. □

Note that, if f is bounded, then g and u above are bounded.

Consider now continuity of our functionals Φ and Ψ. As a first step, it is easy to check that, for $\beta \geq 0$,

$$\lambda_n \to \lambda \quad \text{as} \quad n \to \infty \Rightarrow \|E_{\lambda_n} - E_\lambda\|_\beta \quad \text{as} \quad n \to \infty. \tag{7.6}$$

Let $F_n = E_{\lambda_n} \star C_n$ with density $f_n = e_{\lambda_n} \star C_n$. We need the following results about convergence of F_n and f_n.

LEMMA 7.5. *Suppose that* $\beta \geq 0$, $\|1 - C\|_\beta < \infty$, $\|C_n - C\|_\beta \to 0$ *and* $\lambda_n \to \lambda$ *as* $n \to \infty$. *Then*

(i) $\|F_n - F\|_\beta \to 0$;

(ii) *if* $\beta > 1$ *then* $\int x^{\beta'} dF_n(x) \to \int x^{\beta'} dF(x)$ *for* $1 \leq \beta' < \beta$;

(iii) $\|f_n - f\|_\infty \to 0$.

PROOF. Lemma 7.1 and (7.6) give (i), and (ii) follows easily from (i). Using

$$f = \lambda(C - F), \tag{7.7}$$

(iii) follows from (i). □

Let $m_2 = \int x^2 dC(x)$, $m_{2,n} = \int x^2 dC_n(x)$, $H_n = \Phi(\lambda_n, C_n)$ and $\tilde{G}_n = \Psi(\lambda_n, C_n)$.

PROPOSITION 7.6. *Suppose* $\beta > 1$, $\|1 - C\|_\beta < \infty$, $\|C_n - C\|_\beta \to 0$, $\lambda_n \to \lambda$, $m_{2,n} < \infty$, *and* $m_2 < \infty$. *Then, as* $n \to \infty$,

$$\text{(i)} \quad \|H_n - H\|_\infty \to 0 \quad \textit{and} \quad \text{(ii)} \quad \|\tilde{G}_n - \tilde{G}\|_\infty \to 0.$$

PROOF. We first show that

$$\|e_{\lambda_n} \star U_n - e_\lambda \star U\|_\infty \to 0. \tag{7.8}$$

The conditions of Proposition 7.4 are satisfied, using Lemma 7.5(ii) and (iii). With appropriate versions of u_n and u,

$$\|e_{\lambda_n} \star U_n - e_\lambda \star U\|_\infty \leq \|e_{\lambda_n} - e_\lambda\|_\infty + \|e_\lambda * u_n - e_\lambda * u\|_\infty.$$

The first term tends to zero because of (7.6) and $\lambda_n \to \lambda$. The other term converges to zero by (7.2), using $\|e_{\lambda_n} - e_\lambda\| \to 0$ and Proposition 7.4, and (7.8) is proved.

To prove (i), we have

$$|H_n(t) - H(t)| \le \frac{1}{\lambda_n}\left|\log_e\left(\frac{e_{\lambda_n} \star U_n(t)}{\lambda_n}\right) - \log_e\left(\frac{e_\lambda \star U(t)}{\lambda}\right)\right| + \left|\frac{1}{\lambda_n} - \frac{1}{\lambda}\right|\left|\log_e\left(\frac{e_\lambda \star U(t)}{\lambda}\right)\right|. \tag{7.9}$$

The function $t \mapsto e_\lambda \star U(t)$ is continuous and positive, with value λ when $t = 0$ and finite positive limit as $t \to \infty$. Thus this function is bounded away from zero on $[0, \infty)$. Using the mean value theorem, (7.8), and $\lambda_n \to \lambda$, we have that the first term on the right-hand-side of (7.9) tends to zero uniformly in t. Since $|\log_e((e_\lambda \star U)/\lambda)|$ is bounded, the second term also tends to zero uniformly in t. The proof of (ii) is similar. □

We now consider strong consistency of our input estimators $\hat{\lambda}_n$ and $\hat{C}_n$. Using the Strong Law of Large Numbers, we obtain that $\hat{\lambda}_n \to \lambda$ with probability one. From Lai (1974) (see Shorack and Wellner (1986), Section 10.2) it easily follows that, since $\int x^2 dC(x) < \infty$, we have $\|\hat{C}_n - C\|_\gamma \to 0$ as $n \to \infty$ almost surely for $0 \le \gamma \le 2$. Combining Proposition 7.6 with these strong consistency results for the input estimators gives Theorem 6.1, using the methods of Grübel and Pitts (1993).

7.3 *Proof of Theorem 6.2*

The finite-dimensional delta method gives asymptotic normality of $\hat{\lambda}_n$ as follows. By the Central Limit Theorem, if $\mu = 1/\lambda$ and $\hat{\mu}_n = \frac{1}{n}\sum_{i=1}^n Z_i$, then $\sqrt{n}(\hat{\mu}_n - \mu) \to_d \mathrm{N}(0, \lambda^{-2})$. If $\phi : \mathbb{R} \to \mathbb{R}$ is differentiable at μ, then $\sqrt{n}(\phi(\hat{\mu}_n) - \phi(\mu)) \to_d \mathrm{N}(0, (\phi'(\mu))^2\lambda^{-2})$. Applying this with $\phi(t) = t^{-1}$, we obtain

$$\sqrt{n}(\hat{\lambda}_n - \lambda) \to_d \mathrm{N}(0, \lambda^2). \tag{7.10}$$

Asymptotic normality of the other input estimator $\hat{C}_n$ follows from a classical result on weak convergence of weighted empirical processes (O'Reilly (1974), Csörgő *et al.* (1986), see also Shorack and Wellner (1986), Section 3.7). Write B for a standard Brownian bridge. Then under the conditions of Theorem 6.2 for $0 \le \beta < \gamma$ we have $B(F) \in D_\beta$ with probability one and

$$\sqrt{n}(\hat{C}_n - C) \to_d B(F) \quad \text{as} \quad n \to \infty \quad \text{in} \quad D_\beta. \tag{7.11}$$

In order to apply the infinite-dimensional delta method, we must establish differentiability of Φ and Ψ, see Gill (1989). This is done in Proposition 7.9. We first prove differentiability for the renewal density functional.

PROPOSITION 7.7. *Let f and $\{f_n\}_{n=1}^\infty$ be probability density functions on $[0, \infty)$ with $\int x^2 f(x)dx < \infty$, $\int x^2 f_n(x)dx < \infty$ for all n, and f bounded. Suppose that*

$$\|\sqrt{n}(f_n - f) - g_f\|_\beta \to 0,$$

for some $\beta > 2$, where $\|g_f\|_\beta < \infty$. Then there exist versions u_n, u of renewal densities associated with f_n and f, such that

$$\|\sqrt{n}(u_n - u) - v\|_\infty \to 0,$$

where $v = T(g_f)$ for some linear bounded map T.

PROOF. Since $\beta > 2$, the assumptions and (7.1) imply that

$$(7.12) \quad \|\sqrt{n}(f_n - f) - g_f\|_\infty \to 0 \quad \text{and} \quad \|\sqrt{n}(f_n - f) - g_f\| \to 0,$$

and

$$(7.13) \quad \|\sqrt{n}(\Sigma f_n - \Sigma f) - \Sigma g_f\|_\infty \to 0 \quad \text{and} \quad \|\sqrt{n}(\Sigma f_n - \Sigma f) - \Sigma g_f\| \to 0.$$

From the representation in Lemma 7.3, the required limit involves the limit of $\sqrt{n}(g_n - g)$. The map taking a in an appropriate subset of L onto $(\delta_0 + a)^{*(-1)}$ in L, is Fréchet differentiable there, with derivative given by $x \mapsto -(\delta_0 + a)^{*(-2)} * x$ (see Rudin (1974), 10.36). Thus, since (7.12) and (7.13) hold,

$$\sqrt{n}(g_n - g) \to -(\delta_0 + \Sigma f - f)^{*(-2)} * (\Sigma g_f - g_f) \quad \text{in} \quad L.$$

Let v_1 be the version of the right-hand-side given by

$$v_1 = -(\Sigma g_f - g_f) - 2g * (\Sigma g_f - g_f) - g^{*2} * (\Sigma g_f - g_f),$$

where g is as in (7.5). Then v_1 satisfies

$$v_1 = -(\Sigma g_f - g_f) - (\Sigma f - f) * v_1 - (\Sigma g_f - g_f) * g.$$

Using this v_1 and (7.5), we have

$$\begin{aligned}
&|(\sqrt{n}(g_n - g) - v_1)(t)| \\
&\quad \le |(\sqrt{n}\{(\Sigma f_n - f_n) - (\Sigma f - f)\} - (\Sigma g_f - g_f))(t)| \\
&\quad\quad + |\sqrt{n}(g_n - g) * (\Sigma f_n - f_n)(t) - v_1 * (\Sigma f - f)(t)| \\
&\quad\quad + |\{\sqrt{n}((\Sigma f_n - f_n) - (\Sigma f - f)) - (\Sigma g_f - g_f)\} * g(t)|.
\end{aligned}$$

Each of these terms tends to zero uniformly in t by (7.12), (7.13) and (7.2), and so

$$(7.14) \qquad \|\sqrt{n}(g_n - g) - v_1\|_\infty \to 0.$$

Let $\bar{g}_f = \int x g_f(x)dx$, and

$$\begin{aligned}
v = &\frac{1}{m_1}(\Sigma\Sigma g_f - \Sigma g_f) - \frac{\bar{g}_f}{m_1^2}(\Sigma\Sigma f - \Sigma f) + v_1 + \frac{1}{m_1}(\Sigma\Sigma f - \Sigma f) * v_1 \\
&+ \frac{1}{m_1} g * (\Sigma\Sigma g_f - \Sigma g_f) - \frac{\bar{g}_f}{m_1^2} g * (\Sigma\Sigma f - \Sigma f) - \frac{\bar{g}_f}{m_1^2}.
\end{aligned}$$

Then, with the versions of u_n and u as in Proposition 7.4,

$$
\begin{aligned}
(7.15) \quad & \|\sqrt{n}(u_n - u) - v\|_\infty \\
&\le \left\| \sqrt{n}\left\{ \frac{1}{m_{1,n}}(\Sigma\Sigma f_n - \Sigma f_n) - \frac{1}{m_1}(\Sigma\Sigma f - \Sigma f) \right\} \right. \\
&\qquad \left. - \frac{1}{m_1}(\Sigma\Sigma g_f - \Sigma g_f) + \frac{\bar{g}_f}{m_1^2}(\Sigma\Sigma f - \Sigma f) \right\|_\infty \\
&\quad + \|\sqrt{n}(g_n - g) - v_1\|_\infty \\
&\quad + \left| \left(\sqrt{n}\left\{ \frac{1}{m_{1,n}} g_n * (\Sigma\Sigma f_n - \Sigma f_n) - \frac{1}{m_1} g * (\Sigma\Sigma f - \Sigma f) \right\} \right. \right. \\
&\qquad - \frac{1}{m_1}(\Sigma\Sigma f - \Sigma f) * v_1 - \frac{1}{m_1} g * (\Sigma\Sigma g_f - \Sigma g_f) \\
&\qquad \left. \left. + \frac{\bar{g}_f}{m_1^2} g * (\Sigma\Sigma f - \Sigma f) \right)(t) \right| \\
&\quad + \left| \sqrt{n}\left(\frac{1}{m_{1,n}} - \frac{1}{m_1} \right) + \frac{\bar{g}_f}{m_1^2} \right|.
\end{aligned}
$$

From (7.13) we obtain

$$(7.16) \qquad \sqrt{n}((1/m_{1,n}) - (1/m_1)) \to -\bar{g}_f/m_1^2.$$

This, together with (7.12) and (7.13), shows that the first and last terms on the right-hand-side of (7.15) tend to zero, as does the second term by (7.14). The third term is dealt with similarly, using (7.2). □

We remark that, in L,

$$
\begin{aligned}
v + \frac{\bar{g}_f}{m_1^2} = &- \frac{\bar{g}_f}{m_1^2}(\Sigma\Sigma f - \Sigma f) * (\Sigma f - f + \delta_0)^{*(-1)} \\
&+ \frac{1}{m_1}(\Sigma\Sigma g_f - \Sigma g_f) * (\Sigma f - f + \delta_0)^{*(-1)} \\
&- \frac{1}{m_1}(\Sigma\Sigma f - \Sigma f + m_1\delta_0) * (\Sigma f - f + \delta_0)^{*(-2)} * (\Sigma g_f - g_f).
\end{aligned}
$$

This is the Fréchet derivative at f of the map from an appropriate subset of L into L, taking $a = f + \alpha\delta_0$ to $(\int xf(x)dx)^{-1}(\Sigma\Sigma a - \Sigma a + (\int xf(x)dx)\delta_0) * (\Sigma a - a + \delta_0)^{*(-1)}$, where the derivative is evaluated at g_f.

Grübel (1989) obtains a differentiability result for the renewal density functional when f is an exponential density and f_n is a mixture of f and another fixed density f_1 with the same mean m_1 as f. In this case, our derivative agrees with the one given there.

We use Proposition 7.7 in our proof of the differentiability of Φ and Ψ. Assume that $\{\lambda_n\}$ and λ are positive numbers such that $\sqrt{n}(\lambda_n - \lambda) \to \nu$, where $\nu \in \mathbb{R}$. For $\lambda > 0$ let $f_\lambda(t) = te^{-\lambda t}1_{[0,\infty)}(t)$. Then, for any $\beta \ge 0$, we have

$$(7.17) \qquad \|\sqrt{n}(E_{\lambda_n} - E_\lambda) - \nu f_\lambda\|_\beta \to 0.$$

We next obtain differentiability results for F_n and f_n.

LEMMA 7.8. *For $\beta \geq 0$, assume $\|1 - C\|_\beta < \infty$,*

$$\sqrt{n}(\lambda_n - \lambda) \to \nu, \qquad \|\sqrt{n}(C_n - C) - g_C\|_\beta \to 0,$$

where $\|g_C\|_\beta < \infty$. Then

$$\|\sqrt{n}(f_n - f) - g_f\|_\beta \to 0,$$

where $g_f = (\nu/\lambda)f + \lambda g_C - \lambda\nu f_\lambda \star C - \lambda g_C \star E_\lambda$.

PROOF. Let $g_F = \nu f_\lambda \star C + g_C \star E_\lambda$. We have

$$\begin{aligned}\|\sqrt{n}(F_n - F) - g_F\|_\beta \leq &\|\{\sqrt{n}(E_{\lambda_n} - E_\lambda) - \nu f_\lambda\} \star C_n\|_\beta \\ &+ \|\{\sqrt{n}(C_n - C) - g_C\} \star E_\lambda\|_\beta \\ &+ |\nu| \|f_\lambda \star C_n - f_\lambda \star C\|_\beta.\end{aligned}$$

The first two terms on the right-hand-side tend to zero by Lemma 7.1 and (7.17). Integrating by parts in the third term and using $\|C_n - C\|_\beta \to 0$, we obtain convergence to zero for this term, which yields

$$\|\sqrt{n}(F_n - F) - g_F\|_\beta \to 0. \tag{7.18}$$

The lemma follows by applying (7.7). □

It can be shown that $g_F = -\Sigma g_f$, so that (7.18) implies $\|\sqrt{n}(\Sigma f_n - \Sigma f) - \Sigma g_f\|_\beta$ tends to zero.

PROPOSITION 7.9. *Assume $\beta > 2$, $\|1 - C\|_\beta < \infty$,*

$$\sqrt{n}(\lambda_n - \lambda) \to \nu \quad \text{and} \quad \|\sqrt{n}(C_n - C) - g_C\|_\beta \to 0,$$

where $\|g_C\|_\beta < \infty$. Then

$$\text{(i)} \quad \|\sqrt{n}(H_n - H) - g_H\|_\infty \to 0 \quad \text{and} \quad \text{(ii)} \quad \|\sqrt{n}(\tilde{G}_n - \tilde{G}) - g_G\|_\infty \to 0,$$

where

$$g_H = -\frac{\nu}{\lambda}H + \frac{\nu}{\lambda^2} - \frac{g_1}{\lambda^2 e_\lambda \star U}, \qquad g_1 = \nu f'_\lambda \star U + v \star E_\lambda,$$

and

$$g_G = \frac{v}{e_\lambda \star U} - \frac{u g_1}{(e_\lambda \star U)^2},$$

and v and u are as in Proposition 7.7.

PROOF. We first show that

$$\|\sqrt{n}(e_{\lambda_n} \star U_n - e_\lambda \star U) - g_1\|_\infty \to 0. \tag{7.19}$$

We have

$$\begin{aligned}\|\sqrt{n}(e_{\lambda_n} \star U_n - e_\lambda \star U) - g_1\|_\infty \le\ & \|\sqrt{n}(e_{\lambda_n} - e_\lambda) - \nu f'_\lambda\|_\infty \\ & + \|\sqrt{n}(u_n - u) * e_{\lambda_n} - v * e_\lambda\|_\infty \\ & + \|\{\sqrt{n}(e_{\lambda_n} - e_\lambda) - \nu f'_\lambda\} * u\|_\infty.\end{aligned} \tag{7.20}$$

Applying (7.17) we have, for all $\gamma > 0$,

$$\|\sqrt{n}(e_{\lambda_n} - e_\lambda) - \nu f'_\lambda\|_\gamma \to 0 \quad \text{as} \quad n \to \infty,$$

and so the first term on the right-hand-side of (7.20) tends to zero. For the third term, taking $\gamma > 1$, we have $\|\sqrt{n}(e_{\lambda_n} - e_\lambda) - \nu f'_\lambda\|$ converges to zero by (7.1). In addition, this version of u is bounded, and so the third term tends to zero. For the second term we apply Proposition 7.7. Note that f is bounded. By Lemma 7.8, we have $\sqrt{n}(f_n - f)$ converges to g_f in $\|\cdot\|_\beta$ ($\beta > 2$), and further $\|g_f\|_\beta < \infty$. Hence the conditions of Proposition 7.7 are satisfied, and (7.2) yields that the second term on the right-hand-side of (7.20) converges to zero, and (7.19) is proved.

For the functional Φ, for $t \ge 0$

$$\begin{aligned}&|(\sqrt{n}(H_n - H) - g_H)(t)| \\ &\quad \le \frac{1}{\lambda}\left|\left(\sqrt{n}\left\{\log_e\left(\frac{e_{\lambda_n} \star U_n}{\lambda_n}\right) - \log_e\left(\frac{e_\lambda \star U}{\lambda}\right)\right\} + \frac{\nu}{\lambda} - \frac{g_1}{\lambda e_\lambda \star U}\right)(t)\right| \\ &\quad + |\lambda H(t)|\left|\sqrt{n}\left(\frac{1}{\lambda_n} - \frac{1}{\lambda}\right) + \frac{\nu}{\lambda^2}\right| \\ &\quad + \left|\frac{\nu}{\lambda} - \frac{g_1(t)}{e_\lambda \star U(t)}\right|\left|\frac{1}{\lambda_n} - \frac{1}{\lambda}\right|.\end{aligned}$$

Since H is bounded and since $\sqrt{n}(\lambda_n - \lambda) \to \nu$ implies $\sqrt{n}(\lambda_n^{-1} - \lambda^{-1}) \to -\nu\lambda^{-2}$, the second term tends to zero. For the third term, $\lambda_n^{-1} \to \lambda^{-1}$ and $g_1/(e_\lambda \star U)$ is bounded. The first term tends to zero uniformly in t using the mean value theorem and (7.19). The second part of the proposition follows similarly, using Proposition 7.7 and (7.19). □

Theorem 6.2 now follows from Proposition 7.9, (7.11) and (7.10), on applying the delta method, as in Grübel and Pitts (1993), Section 3.7.

7.4 *Proof of Theorem 6.3*

The key to this proof is the following lemma, which implies that $\hat{R}_n$ and R_n are eventually close, and similarly for $\hat{S}_n$ and S.

LEMMA 7.10. *Assume $\int x^{2\gamma} dC(x) < \infty$ for some $\gamma > 2$. Then $\hat{R}_n \to_d R$ and $\hat{S}_n \to_d S$ in D_∞.*

PROOF. This is a straightforward adaptation of the proof of Proposition 3.15 in Grübel and Pitts (1993). □

We obtain Theorem 6.3 on using the above lemma and Lemma 3.16 in Grübel and Pitts (1993).

8. Simulation studies

In order to obtain the full benefit from the theory developed here, we would need extensive laboratory data on real particles diffusing according to the Ornstein-Uhlenbeck dynamics of Section 3, or of biological particles such as spermatazoa and leukocytes. We hope to discuss this further elsewhere.

So far as simulation studies rather than analysis of laboratory data are concerned, one needs a thorough simulation of $M/G/\infty$ queues with a range of choices of G, taken from suitable parametric families, for example. Again, we defer further consideration here.

We begin with a simulation study of the $M/M/\infty$ case, for two reasons:

(i) the $M/M/\infty$ queue is particularly easy to simulate;

(ii) the motivation for the paper was a desire to study the adequacy of the exponential approximation, that is, of the $M/M/\infty$ queue as a model for the Ornstein-Uhlenbeck dynamics. As the work of Sections 6 and 7 makes clear, this reduces to study of local properties: that is, functional derivatives near exponentiality. Our $M/M/\infty$ study is of the functional derivative *at* exponentiality, which we offer as the most basic case.

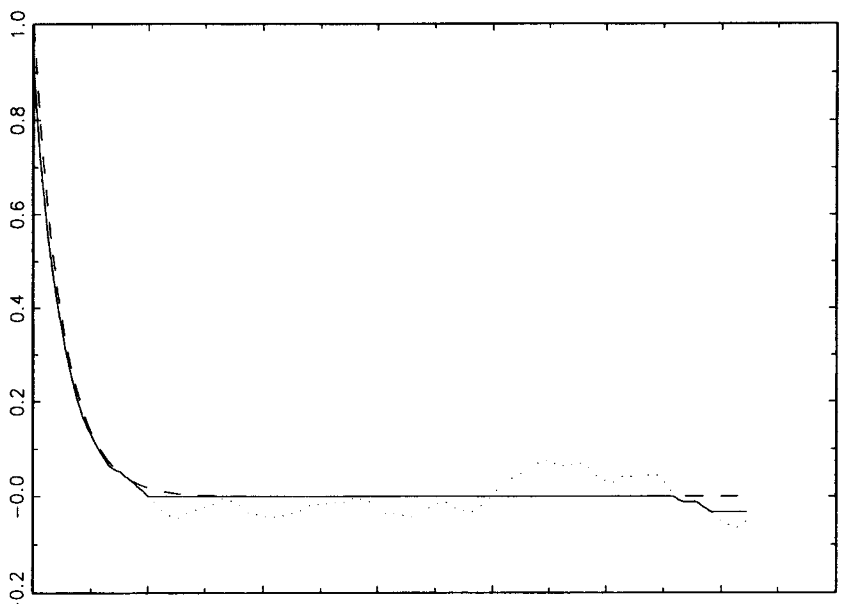

Fig. 1.

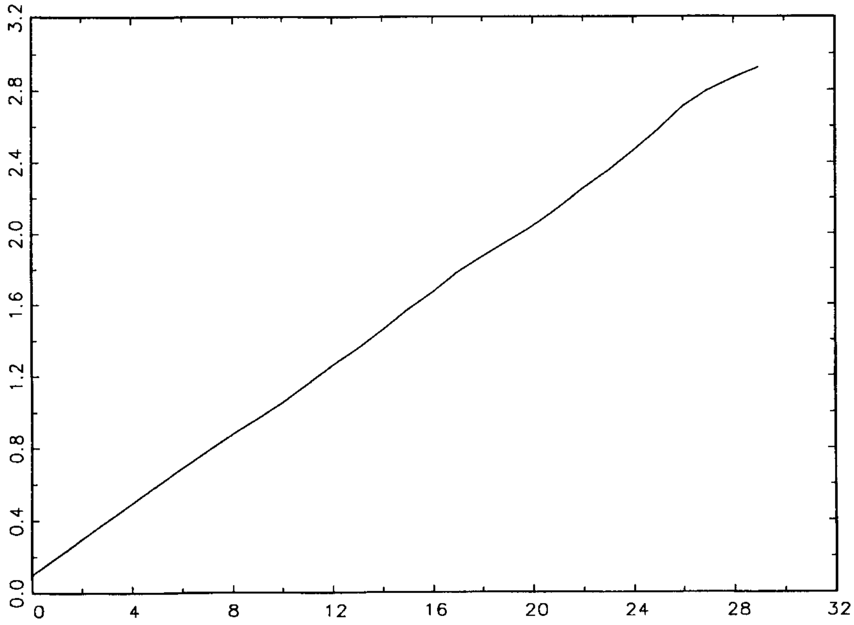
0.0
0.4
0.8
1.2
1.6
2.0
2.4
2.8
3.2
0
4
8
12
16
20
24
28
32

Fig. 2.

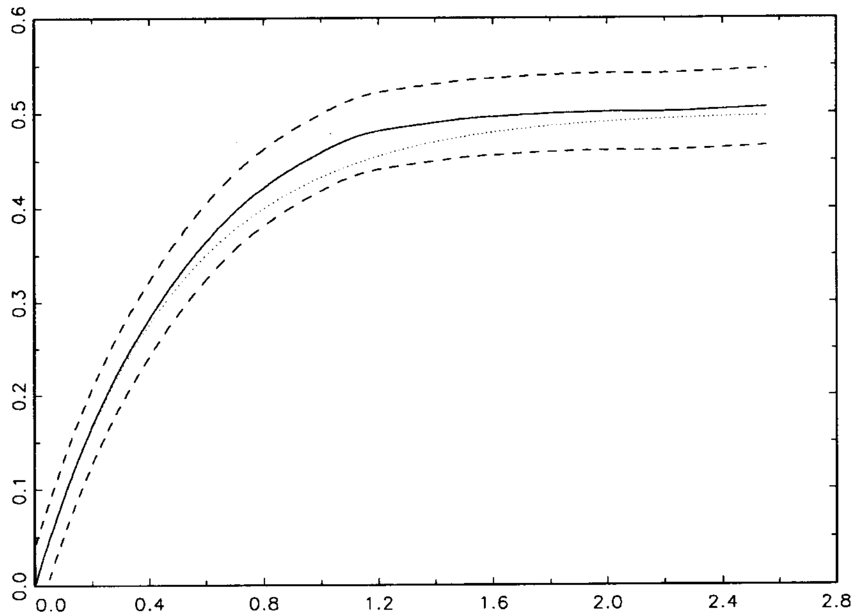
0.0
0.1
0.2
0.3
0.4
0.5
0.6
0.0
0.4
0.8
1.2
1.6
2.0
2.4
2.8

Fig. 3.

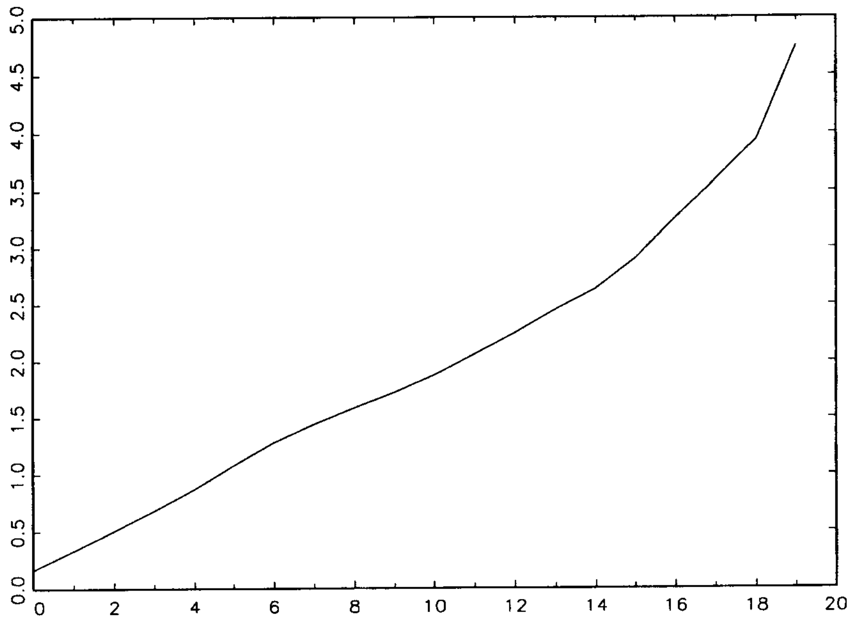

Fig. 4a.

Our two simulations, for count and indicator data, were conducted separately for convenience. For the first, we used a sample of 1,000 events (arrivals and departures) for $M(1)/M(2)/\infty$, an $M/M/\infty$ queue with traffic intensity $\frac{1}{2}$, with $h = 0.05$, after a burn-in of 10,000 events. In Fig. 1 we show the true ρ (dashed curve), exactly exponential, the estimated curve r after monotone regression as in Section 5 (solid curve), and the estimated r without monotone regression (dotted curve—a much worse fit). In Fig. 2 we plot $-\log r_j$ against jh, again as in Section 5, to illustrate near-linearity. For the indicator data, we used 300 busy and 300 idle periods. Figure 3 shows the true (non-normalized) integrated tail H (dotted curve), the estimated H (solid curve), and the nominal 90% bootstrap confidence curves (with 300 bootstrap repetitions).

Note 1. For the indicator-data case of Section 6, where our asymptotic results depend on having a large number of both idle and busy periods, we point out that if $\mu = \lambda\alpha$ is at all large (say, μ of the order of 5 or 6 even) most of the time-axis is occupied by busy periods, and so an inconveniently long time will be needed to accumulate enough idle periods. Thus keeping μ low is important for simulation purposes.

Note 2. Reference to $(*)$ shows that use of idle and busy periods is directly informative, not about G itself—our primary object of interest—but about the normalised integrated tail $G^*(x) := \alpha^{-1}\int_0^x (1 - G(u))du$. To pass from G^* to G involves differentiation, and there is an unavoidable source of difficulty both theoretically and numerically. It may well happen, for example, that G^* is closer

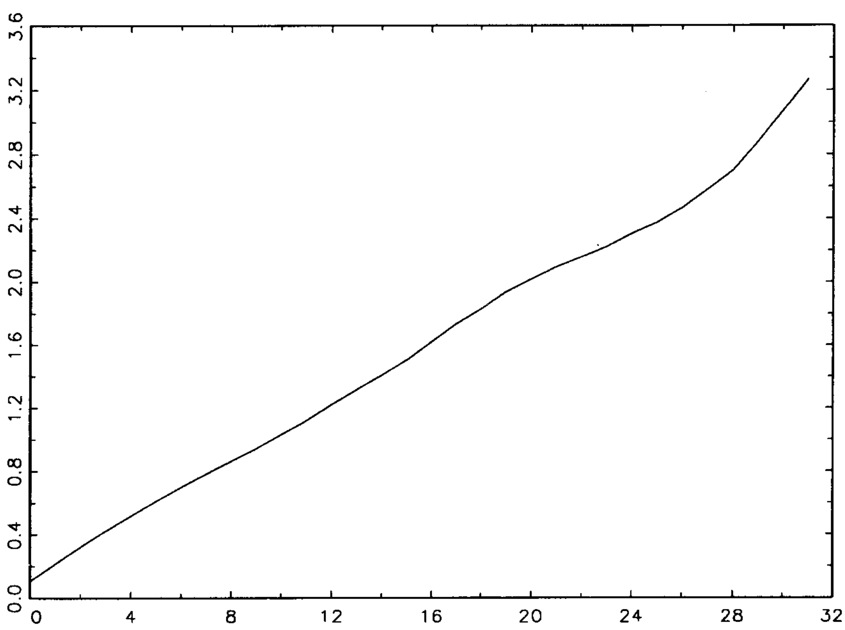

Fig. 4b.

to exponential than G itself. So although G^* is exactly exponential if and only if G is, nevertheless this dependence on G^* may blunt our ability to detect departures from exponentiality in G in practice.

Note 3. We have not conducted an exhaustive simulation study of the behaviour of these estimators in situations with non-exponential service times. However we have considered the $M/E_2/\infty$ case (with mean service time 1/2) and the $M/H_2/\infty$ case (with service time distribution given by an equal mixture of an exponential distribution with mean 1/4 and an exponential distribution with mean 3/4). For comparison, we include in Fig. 4a and Fig. 4b respectively, the resulting plots corresponding to Fig. 2.

Acknowledgements

We are much indebted to both referees for their very helpful comments.

References

Andersen, P. K., Borgan, O., Gill, R. D. and Keiding, N. (1993). *Statistical Models Based on Counting Processes*, Springer, New York.

Baccelli, P. and Brémaud, P. (1994). *Elements of Queueing Theory. Palm-martingale Calculus and Stochastic Recurrences*, Springer, New York.

Barlow, R. E., Bartholomew, D. J., Bremner, J. M. and Brunk, H. D. (1972). *Statistical Inference under Order Restrictions*, Wiley, New York.

Bartlett, M. S. (1955/78). *An Introduction to Stochastic Processes*, with Special Reference to Methods and Applications, Cambridge University Press, Cambridge (1st ed. 1955, 2nd ed. 1966, 3rd ed. 1978).

Beneš, V. E. (1957). Fluctuations of telephone traffic, *Bell System Technical Journal*, **36**, 965–973.

Billingsley, P. (1961). *Statistical Inference for Markov Processes*, University of Chicago Press, Chicago.

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

Billingsley, P. (1979). *Probability and Measure*, Wiley (2nd ed. 1986), New York.

Bingham, N. H. and Dunham, B. (1997). Estimating diffusion coefficients for count data: Einstein-Smoluchowski theory revisited. *Ann. Inst. Statist. Math.*, **49**, 667–679.

Brown, M. (1970). An $M/G/\infty$ estimation problem, *Ann. Math. Statist.*, **41**, 651–654.

Csörgő, M. and Zitikis, R. (1996). Mean residual life processes, *Ann. Statist.*, **24**, 1717–1739.

Csörgő, M., Csörgő, S., Horváth, L. and Mason, D. M. (1986). Weighted empirical and quantile processes, *Ann. Probab.*, **14**, 31–85.

Doering, C. R., Hagan, P. S. and Levermore, C. D. (1989*a*). Mean exit times for particles driven by weakly coloured noise, *SIAM J. Appl. Math.*, **49**, 1480–1513.

Doering, C. R., Hagan, P. S. and Levermore, C. D. (1989*b*). The distribution of exit times for weakly coloured noise, *J. Statist. Phys.*, **54**, 1321–1351.

Doob, J. L. (1953). *Stochastic Processes*, Wiley, New York.

Genstat 5 Committee (1993). *Genstat 5 Release 3 Reference Manual*, Oxford University Press, Oxford.

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

Gill, R. D. and van der Vaart, A. W. (1993). Non- and semi-parametric maximum likelihood estimators and the von Mises method, II, *Scand. J. Statist.*, **20**, 271–288.

Glynn, P. W. and Whitt, W. (1986). A central-limit version of $L = \lambda W$, *Queueing Systems Theory Appl.*, **1**, 191–215.

Glynn, P. W. and Whitt, W. (1988). Ordinary central limit theorem and weak law of large numbers versions of $L = \lambda W$, *Math. Oper. Res.*, **13**, 674–692.

Glynn, P. W. and Whitt, W. (1989). Indirect estimation via $L = \lambda W$, *Oper. Res.*, **37**, 82–103.

Grübel, R. (1986). On harmonic renewal measures, *Probab. Theory Related Fields*, **71**, 393–404.

Grübel, R. (1989). Stochastic models as functionals: some remarks on the renewal case, *J. Appl. Probab.*, **26**, 296–303.

Grübel, R. and Pitts, S. M. (1992). A functional approach to the stationary waiting time and idle period distributions of the $GI/G/1$ queue, *Ann. Probab.*, **20**, 1754–1778.

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. (1988). *An Introduction to the Theory of Coverage Processes*, Wiley, New York.

Hall, P. and Heyde, C. C. (1980). *Martingale Limit Theory and Its Applications*, Academic Press, Boston.

Hannan, E. J. (1970). *Multiple Time Series*, Wiley, New York.

Hannan, E. J. (1976). The asymptotic distribution of serial covariances, *Ann. Statist.*, **4**, 396–399.

Hannan, E. J. and Heyde, C. C. (1972). On limit theorems for quadratic functions of discrete time series, *Ann. Math. Statist.*, **43**, 2058–2066.

Hannan, E. J. and Kavalieris, L. (1983). The convergence of autocorrelations and autoregressions, *Austral. J. Statist.*, **25**, 287–297.

Hesse, C. H. (1991). The one-sided barrier problem for an integrated Ornstein-Uhlenbeck process, *Comm. Statist. Stochastic Models*, **7**, 447–480.

Krengel, U. (1985). *Ergodic Theorems*, Walter de Gruyter, Berlin.

Lai, T. L. (1974). Convergence rates in the strong law of large numbers for random variables taking values in Banach spaces, *Bull. Inst. Math. Acad. Sinica*, **2**, 67–85.

Lindley, D. V. (1956). The estimation of velocity distributions from counts, *Proceedings of the International Congress of Mathematicians* (Amsterdam, 1954) III, 427–444, North-Holland, Amsterdam.

Milne, R. K. (1970). Identifiability for random translates of Poisson processes, *Zeit. Wahrscheinlichkeitsth.*, **15**, 195–201.

Nozari, A. and Whitt, W. (1988). Estimating average production intervals using inventory measurements: Little's law for partially observable processes, *Oper. Res.*, **36**, 308–323.

O'Reilly, N. E. (1974). On the weak convergence of empirical processes in sup-norm metrics, *Ann. Probab.*, **2**, 642–651.

Pickands, J. and Stine, R. A. (1997). Estimation for an $M/G/\infty$ queue with incomplete information, *Biometrika*, **84**, 295–308.

Pitts, S. M. (1994*a*). Nonparametric estimation of the stationary waiting-time distribution for the $GI/G/1$ queue, *Ann. Statist.*, **22**, 1428–1446.

Pitts, S. M. (1994*b*). Nonparametric estimation of compound distributions with applications in insurance, *Ann. Inst. Statist. Math.*, **46**, 537–555.

Pollard, D. (1984). *Convergence of Stochastic Processes*, Springer, New York.

Pratt, J. W. (1960). On interchanging limits and integrals, *Ann. Math. Statist.*, **31**, 74–77.

Rao, C. R. (1973). *Linear Statistical Inference and Its Applications*, 2nd ed., Wiley, New York.

Reynolds, J. F. (1968). On the autocorrelation and spectral functions of queues, *J. Appl. Probab.*, **5**, 467–475.

Reynolds, J. F. (1972). On linearly regressive processes, *J. Appl. Probab.*, **9**, 208–213.

Reynolds, J. F. (1975). The covariance structure of queues and related processes: a survey of recent work, *Advances in Applied Probability*, **7**, 383–415.

Riordan, J. (1951). Telephone traffic time averages, *Bell System Technical Journal*, **30**, 1129–1144.

Rudin, W. (1974). *Functional Analysis*, Tata McGraw-Hill, New Delhi.

Schassberger, R. (1978). Insensitivity of steady-state distributions of generalized semi-Markov processes with speeds, *Advances in Applied Probability*, **10**, 836–851.

Shorack, G. R. and Wellner, J. A. (1986). *Empirical Processes with Applications to Statistics*, Wiley, New York.

Takács, L. (1969). On Erlang's formula, *Ann. Math. Statist.*, **40**, 71–78.

van der Vaart, A. and Wellner, J. A. (1996). *Weak Convergence and Empirical Processes*, Springer, New York.

Vere-Jones, D. (1968). Some applications of probability generating functionals to the study of input-output streams, *J. Roy. Statist. Soc. Ser. B*, **30**, 321–333.

Whitt, W. (1991). A review of $L = \lambda W$ and extensions, *Queueing Systems Theory Appl.*, **9**, 235–268 (Correction: *ibid.* (1992). **12**, 431–432).