The Annals of Statistics
1994, Vol. 22, No. 3, 1428–1446

NONPARAMETRIC ESTIMATION OF THE STATIONARY WAITING TIME DISTRIBUTION FUNCTION FOR THE $GI/G/1$ QUEUE

BY SUSAN M. PITTS

University College London

The $GI/G/1$ queueing model is regarded as a functional that maps the service and interarrival time distribution functions onto the stationary waiting time distribution function. By considering the output of the functional when it is applied to nonparametric estimators of the input distribution functions, we obtain a nonparametric estimator of the stationary waiting time distribution function. Using appropriate continuity and differentiability properties of the functional, we show that statistical properties of the input estimators carry over to corresponding properties for the stationary waiting time distribution function estimator.

1. Introduction. This paper is concerned with a classical stochastic model, that of the $GI/G/1$ queue. Customers arrive in a renewal process to a single server; are served in order of arrival, with successive service times being independent, identically distributed random variables, independent of the interarrival times; and then leave the system. Write μ_S and μ_T for the probability distributions [on the Borel sets of $\mathbf{R}$, concentrated on $(0, \infty)$] of the service and interarrival times, respectively. In the above description, μ_S and μ_T are unspecified; fixing these fixes the stochastic evolution of the queue and, in particular, fixes various quantities of interest. One of these is the stationary waiting time distribution, which is defined when the traffic intensity

$$\rho = \frac{\int x\mu_S(dx)}{\int x\mu_T(dx)} < 1,$$

in which case successive customer waiting times converge in distribution to a proper random variable, called the stationary waiting time, with distribution μ_W [see, e.g., Asmussen (1987), Chapter 8].

Here we consider the statistical problem of estimating the stationary waiting time distribution function given random samples from each of the service and interarrival time distributions. A functional approach is taken as in, for example, Grübel and Pitts (1992), in that the $GI/G/1$ queueing model is regarded as a functional that maps the pair consisting of the service and interarrival time distributions onto the stationary waiting time distribution. The nonparametric estimator proposed in this paper is defined by evaluating the stationary waiting time functional at the pair given by nonparametric estimators (the empirical distribution functions) of the input distributions. Statistical properties of the

Received January 1992; revised August 1993.
AMS 1991 *subject classifications*. Primary 60K25; secondary 62G05.
Key words and phrases. $GI/G/1$ queue, nonparametric estimation.

resulting output estimator are obtained by combining corresponding properties of the input estimators with local properties of the functional. This approach is the same as that taken in Grübel and Pitts (1993), where a similarly defined nonparametric estimator of the renewal function is studied.

This procedure may be compared to that taken in a parametric context, where the input distributions are supposed to belong to certain parametric families (e.g., they may both be assumed to be Erlangian). The parameters of the input distributions are estimated from the data and any output quantity is then "estimated" by calculating the corresponding quantity for a queue with the estimated input distributions.

For this parametric approach, attention in the literature is focused on estimation of the parameters of the input distributions, under various parametric assumptions and observation strategies, as in, for example, Basawa and Prabhu (1988). A nonparametric approach is taken in Gaver and Jacobs (1988), where a nonparametric estimator of the probability of a long customer delay in an $M/G/1$ queue is studied in the case where the interarrival distribution is known and a random sample from the unknown service time distribution is available. This estimator is obtained by replacing an exponential approximation for the tail of the waiting time distribution by its empirical counterpart. A non-functional approach to the problem of estimation of output quantities for queues concerns inference from direct observation of the output processes themselves [see, e.g., Heyde (1988) and the references contained there].

In Section 2 we introduce notation and give a precise description of the stationary waiting time functional in terms of harmonic renewal measures. We also define the spaces that are to form the setting for the functional. Section 3 deals with continuity of the functional and strong consistency of our estimator. In Section 4 we establish a suitable differentiability property of the functional, which, together with an asymptotic normality result for the input estimators, leads via the delta method [as described in Gill (1989)] to a corresponding asymptotic normality result for the estimator in terms of convergence in distribution to a Gaussian process. The differentiability also leads to asymptotic validity of bootstrap confidence bands in an appropriate function space for the unknown stationary waiting time distribution function. These are simultaneous bands, having a different interpretation from that of pointwise bands and allowing for consideration of global properties of the unknown F_W, for example, whether F_W stochastically dominates some given stationary waiting time distribution function. In Section 5 results from van der Vaart (1991) are used to obtain asymptotic efficiency of the estimator, again relying on the differentiability of the functional. Section 6 contains an example and discussion, and the proofs of the continuity and differentiability results are in the Appendix.

2. Definitions. Using the connection between the stationary waiting time distribution and that of the maximum of an associated random walk, we give a decomposition of the functional into several simpler maps. This connection is described in, for example, Asmussen (1987).

Let $\{X_n\}_{n \in \mathbf{N}}$ be independent, identically distributed random variables with distribution μ the same as that of $S-T$, where S and T are independent random variables with distributions μ_S and μ_T, respectively. Throughout we assume that the queue is stable, that is, $\rho < 1$, and that the distribution of $S - T$ is not concentrated on $(-\infty, 0]$ since this case is not of interest for the stationary waiting time. Let $\{Z_n\}_{n \in \mathbf{N}_0}$ be defined by $Z_0 = 0$ and $Z_n = \sum_{i=1}^{n} X_i$, so that $\{Z_n\}$ is a random walk with step distribution μ. Then the stationary waiting time distribution μ_W is the same as that of the maximum of the random walk, and we approach this distribution via harmonic renewal measures. Define the harmonic renewal measure $\nu(\mu)$ associated with a (possibly defective) probability measure μ by

$$\nu(\mu) = \sum_{k=1}^{\infty} \frac{1}{k} \mu^{*k},$$

where μ^{*k} denotes the k-fold convolution of μ. When μ is proper, $\nu(\mu)$ is an infinite measure. Write μ_+ for the (defective) distribution of the first (strict) ascending ladder height of the random walk. The following two relations together form the crux of our decomposition:

(1) $$\nu(\mu_+)(A) = \nu(\mu)\big(A \cap (0, \infty)\big) \quad \text{for all Borel sets } A,$$

and, writing $\widehat{\mu}(\theta)$ for the Fourier transform $\int \exp(i\theta x)\mu(dx)$ of μ,

(2) $$\widehat{\mu}_W(\theta) = \exp\{\nu(\mu_+)^\wedge(\theta) - \nu(\mu_+)(\mathbf{R})\} \quad \text{for all } \theta \in \mathbf{R},$$

where (1) is obtained from the Spitzer–Baxter equations [Feller (1971), 18.3] and is discussed in Grübel (1989), and (2) is Spitzer's identity for the maximum of a random walk with drift to $-\infty$ [Feller (1971), 18.5], written in terms of harmonic renewal measures. Thus the stationary waiting time functional can be decomposed as

(3) $$(\mu_S, \mu_T) \mapsto \mu \mapsto \nu(\mu) \mapsto \nu(\mu_+) \mapsto \mu_W;$$

see also Grübel and Pitts (1992).

We need to translate (3) to a map taking the pair (F_S, F_T) of service and interarrival time distribution fuctions to the stationary waiting time distribution function F_W, and this requires specification of appropriate spaces for the domain and codomain for the functional. Let D_∞ be the space of right-continuous real-valued functions on $[-\infty, \infty]$ with left-hand limits and left-continuous at ∞; this space provides a natural setting for distribution functions. However, we shall see below that we need to consider weighted spaces of right-continuous functions with left-hand limits. In order to motivate the introduction of these spaces, we turn to a consideration of convolution. We need a definition of convolution involving infinite measures. Let $\mathcal{H}$ be the set of real-valued functions H on $\mathbf{R}$ that are right-continuous and nondecreasing and that satisfy $\lim_{x \to -\infty} H(x) = 0$. For H in $\mathcal{H}$ and measurable $g\colon \mathbf{R} \to \mathbf{R}$ with the map $x \mapsto g(t-x)$ integrable for all t

in $\mathbf{R}$ with respect to Lebesgue-Stieltjes measure defined by H, the convolution $g \star H$ of g and H is given by

$$g \star H(t) = \int g(t-x)H(dx).$$

If H_1 and H_2 are in $\mathcal{H}$ and $H_1 \star H_2$ exists, then so does $H_2 \star H_1$, and the two are equal. Provided the relevant integrals exist, convolution powers of elements of $\mathcal{H}$ can be defined by letting $H^{\star 0}$ be the indicator function $I_{[0,\infty)}$ of $[0,\infty)$ and $H^{\star n} = H \star H^{\star(n-1)}$ for n in $\mathbf{N}$. We define the harmonic renewal function V associated with a (possibly defective) probability measure μ with distribution function F and

$$\int |x|F(dx) < \infty, \qquad m_1 = \int xF(dx) > 0 \tag{4}$$

to be

$$V(x) = \nu(\mu)\big((-\infty, x]\big) \quad \text{for } x \text{ in } \mathbf{R}.$$

From Heyde [(1964), Lemma 1], $0 \le V(x) < \infty$ for all x in $\mathbf{R}$, so V is in $\mathcal{H}$. If μ is defective, then $\nu(\mu)$ is finite and V can be extended to an element of D_∞ by defining $V(\infty) = \lim_{x\to\infty} V(x)$ and $V(-\infty) = 0$. However, when μ is nondefective, V is not in D_∞, and so we consider other spaces as in Grübel and Pitts (1993). For a real-valued function f on $[-\infty,\infty]$ and α and β in $\mathbf{R}$, let $T_{\alpha\beta}f$ be the function from $\mathbf{R}$ to $\mathbf{R}$ defined by

$$T_{\alpha\beta}f(x) = \begin{cases} (1+x)^\beta f(x), & \text{if } x \ge 0, \\ \big(1+|x|\big)^\alpha f(x), & \text{if } x < 0. \end{cases}$$

We write $D_{\alpha\beta}$ for the set of real-valued functions f on $\mathbf{R}$ with $T_{\alpha\beta}f$ extendable to an element of D_∞. For f in $D_{\alpha\beta}$, define

$$\|f\|_{\alpha\beta} = \|T_{\alpha\beta}f\|_\infty,$$

where $\|\cdot\|_\infty$ is the supremum norm. Then $D_{\alpha\beta}$ is a nonseparable Banach space. If β is negative, then $D_{\alpha\beta}$ contains functions which do not have a finite limit at $+\infty$. Now let F be a distribution function satisfying (4) with finite second moment m_2. The renewal function U associated with F is defined by

$$U(x) = \sum_{k=0}^{\infty} F^{\star k}(x) \quad \text{for } x \text{ in } \mathbf{R}.$$

Under the above conditions, from Daley (1980) we obtain

$$0 \le V(x) \le U(x) \le \frac{x \vee 1}{m_1} + \frac{m_2}{m_1^2}, \tag{5}$$

where $x \vee y$ is $\max\{x, y\}$. By the elementary renewal theorem U is in $D_{0,-1}$, and (5) yields a bound on $\|U\|_{0,-1}$. From Alsmeyer [(1991), Theorem 1.2], we have $V(x)/(1+x) = o(1)$ as $x \to \infty$, that is, V is in $D_{0,-1}$. We write $C_{\alpha\beta}$ for the set of continuous functions in $D_{\alpha\beta}$.

Let $F = F_T \star F_{-S}$, where F_{-S} is the distribution function of $-S$, and let V be the harmonic renewal function associated with F. Let V_+ be defined by $V_+(x) = \nu(\mu_+)((-\infty, x])$, where μ_+ is the right Wiener–Hopf factor for the random walk with step distribution $F_{-T} \star F_S$, so that, by (1), $V_+(x) = V(0-) - V(-x-)$ for $x > 0$, where $V(y-) = \lim_{t\uparrow y} V(t)$. Then, from (2),

$$F_W(x) = \exp\{-\lim_{x\to\infty} V_+(x)\} \sum_{k=0}^{\infty} \frac{1}{k!} V_+^{\star k}(x). \tag{6}$$

3. Continuity and consistency. Theorem 3.1 gives continuity of the stationary waiting time functional. Throughout, for n in $\mathbf{N}_0$, let $F_{S,n}$ and $F_{T,n}$ be the distribution functions of the service and interarrival times, respectively, for a stable $GI/G/1$ queue, and let $F_{W,n}$ be the associated stationary waiting time distribution function.

THEOREM 3.1. *Assume* $\int x^2 F_{T,n}(dx) < \infty$ *and* $\int x^\gamma F_{S,n}(dx) < \infty$ *for some* $\gamma > 2$, *and*

$$\|F_{S,n} - F_{S,0}\|_{0\gamma} \to 0 \quad \text{and} \quad \|F_{T,n} - F_{T,0}\|_\infty \to 0 \quad \text{as } n \to \infty.$$

Then

$$\|F_{W,n} - F_{W,0}\|_\infty \to 0 \quad \text{as } n \to \infty.$$

The proof of this theorem is given in the Appendix. Borovkov [(1976), Section 21] gives sufficient conditions for continuity of the stationary waiting time functional. His conditions are weaker than ours, requiring only the existence of first moments and

$$\int x F_{S,n}(dx) \to \int x F_{S,0}(dx) \quad \text{as } n \to \infty.$$

However, we shall use the techniques developed in our proof when we go on to consider differentiability of the functional.

Now suppose that F_S and F_T (and F_W) are unknown, and that $\{S_n\}_{n\in\mathbf{N}}$ are independent identically distributed random variables on a probability triple $(\Omega, \mathcal{A}, \mathbf{P})$ with distribution function F_S, and $\{T_n\}_{n\in\mathbf{N}}$ are independent identically distributed random variables on $(\Omega, \mathcal{A}, \mathbf{P})$ with distribution function F_T, with $\{T_n\}$ independent of $\{S_n\}$. Let $\widehat{F}_{S,n}$ and $\widehat{F}_{T,n}$, be the the empirical distribution functions based on $S_1, \ldots, S_n$ and $T_1, \ldots, T_n$, respectively, so that, for x in $\mathbf{R}$ and ω in Ω,

$$\widehat{F}_{S,n}(x,\omega) = n^{-1} \sum_{i=1}^{n} I_{(-\infty,x]}(S_i(\omega)) \quad \text{and} \quad \widehat{F}_{T,n}(x,\omega) = n^{-1} \sum_{i=1}^{n} I_{(-\infty,x]}(T_i(\omega)).$$

Define $\Phi(\widehat{F}_{S,n}, \widehat{F}_{T,n})$ to be zero if the queue with input $(\widehat{F}_{S,n}, \widehat{F}_{T,n})$ is not stable, and otherwise to be the stationary waiting time distribution function $\widehat{F}_{W,n}$ for the queue with this input; our estimator is $\Phi(\widehat{F}_{S,n}, \widehat{F}_{T,n})$. By the strong law of large numbers, with probability 1, $\Phi(\widehat{F}_{S,n}, \widehat{F}_{T,n})$ is well-defined as $\widehat{F}_{W,n}$ eventually.

For measurability purposes, we give $D_{\alpha\beta}$ with its open ball (projection) σ-field, $\mathcal{D}^P_{\alpha\beta}$ [see Pollard (1984), Chapter 4, for details of this notion and of convergence in distribution in nonseparable metric spaces]. It is straightforward to check that the estimator evaluated at x in $\mathbf{R}$ is a random variable and hence that the estimator is a random element of D_∞.

An appropriate consistency result for the input estimators is provided by the following weighted version of the Glivenko–Cantelli theorem.

LEMMA 3.2. *Let $\{X_i\}_{i\in\mathbf{N}}$ be independent, identically distributed random variables with continuous distribution function F. Let $\widehat{F}_n$ be the empirical distribution function based on $X_1, \dots, X_n$. Then, for $\gamma \geq 0$,*

$$E(|X_1 \vee 0|^\gamma) < \infty \quad \Rightarrow \quad \|\widehat{F}_n - F\|_{0\gamma} \to 0 \quad \textit{as } n \to \infty \textit{ almost surely.}$$

This is proved by rescaling the weighted Glivenko–Cantelli result for random variables uniformly distributed on (0, 1), which is given in Shorack and Wellner [(1986), 10.2]. Theorem 3.1 and Lemma 3.2 yield the following theorem.

THEOREM 3.3. *Suppose that F_S is continuous with, for some $\gamma > 2$,*

$$\int x^\gamma F_S(dx) < \infty, \qquad \int x^2 F_T(dx) < \infty.$$

Then, with probability 1,

$$\|\Phi(\widehat{F}_{S,n}, \widehat{F}_{T,n}) - F_W\|_\infty \to 0 \quad \textit{as } n \to \infty.$$

4. Differentiability, asymptotic normality and the bootstrap. We show that the stationary waiting time functional is differentiable along certain curves of pairs of distribution functions. Let $F_{-S,n}(x) = 1 - F_{S,n}(-x-)$ and $g_{-S}(x) = -g_S(-x-)$.

THEOREM 4.1. *Assume $\int x^\gamma F_{S,n}(dx) < \infty$ and $\int x^\gamma F_{T,n}(dx) < \infty$ for some $\gamma > 2$. Assume that, for some continuous g_S and g_T in $D_{0\gamma}$,*

$$\|\sqrt{n}(F_{S,n} - F_{S,0}) - g_S\|_{0\gamma} \to 0 \quad \textit{as } n \to \infty$$

and

$$\|\sqrt{n}(F_{T,n} - F_{T,0}) - g_T\|_{0\gamma} \to 0 \quad \textit{as } n \to \infty.$$

Then

$$\|\sqrt{n}(F_{W,n} - F_{W,0}) - \Phi'_{(F_S,F_T)}(g_S, g_T)\|_\infty \to 0 \quad as\ n \to \infty,$$

where

$$\begin{aligned}\Phi'_{(F_S,F_T)}(g_S,g_T) &= -h \star F_{W,0},\\ h(t) &= \begin{cases} g \star U(-t), & if\ t \geq 0,\\ 0, & if\ t < 0,\end{cases}\\ g &= g_{-S} \star F_{T,0} + g_T \star F_{-S,0}\end{aligned}$$

and U is the renewal function associated with $F_{-S,0} \star F_{T,0}$.

This is proved in the Appendix. The derivative $\Phi'_{(F_S,F_T)}$ at (F_S,F_T) is a linear bounded map from $D_{0\gamma} \times D_{0\gamma}$ to D_∞.

The appropriate asymptotic normality result for the input estimators is given next. Let $B\,[= B(t,\omega)$ for t in $[0,1]$ and ω in $\Omega]$ be a standard Brownian bridge, and let $B \circ F$ be defined by $(B \circ F)(x,\omega) = B(F(x),\omega)$. We call $B \circ F$ a Brownian bridge rescaled by F. As in Pollard [(1984), Chapter 4], we say that random elements $\{X_n\}$ in $D_{\alpha\beta}$ converge in distribution to a random element X in $D_{\alpha\beta}$ and write $X_n \to_d X$ as $n \to \infty$ in $D_{\alpha\beta}$, if $E(f(X_n)) \to E(f(X))$ as $n \to \infty$ for all real-valued bounded continuous measurable functions f on $D_{\alpha\beta}$. For random elements X and Y of $D_{\alpha\beta}$ we write $X =_d Y$ if X and Y have the same distribution.

LEMMA 4.2. *Let $\{X_i\}_{i\in\mathbf{N}}$, F and $\widehat{F}_n$ be as in Lemma 3.2. Let $\gamma > 0$ and suppose that $E(|X_1 \vee 0|^\gamma) < \infty$. Then, for every $0 \leq \beta < \gamma/2$,*

$$n^{1/2}(\widehat{F}_n - F) \to_d B \circ F \quad as\ n \to \infty\ in\ D_{0\beta}.$$

This is proved using the corresponding result for uniformly distributed random variables given in Shorack and Wellner [(1986), 3.7.1]; see also Pyke and Shorack (1968). The result is translated to the general case by rescaling.

THEOREM 4.3. *Suppose that F_S and F_T are continuous and satisfy*

$$\int x^{2\gamma} F_S(dx) < \infty \quad and \quad \int x^{2\gamma} F_T(dx) < \infty,$$

for some $\gamma > 2$. Then

$$\sqrt{n}\big(\Phi(\widehat{F}_{S,n},\widehat{F}_{T,n}) - F_W\big) \to_d Z \quad as\ n \to \infty\ in\ D_\infty,$$

where Z is a Gaussian process obtained by applying the derivative of the stationary waiting time functional to the pair $(B_1 \circ F_S, B_2 \circ F_T)$ for independent Brownian bridges B_1 and B_2.

PROOF. Apply Lemma 4.2 to obtain convergence in distribution in $D_{0\gamma'}$, $2 < \gamma' < \gamma$, of $\sqrt{n}(\widehat{F}_{S,n} - F_S)$ and of $\sqrt{n}(\widehat{F}_{T,n} - F_T)$. The limiting rescaled Brownian bridges concentrate on a separable subspace of $D_{0\gamma'}$, and so we can apply the Skorohod–Dudley–Wichura theorem [see, e.g., Shorack and Wellner (1986), 2.3.5] to conclude that there exists a probability space $(\widetilde{\Omega}, \widetilde{\mathcal{A}}, \widetilde{\mathbf{P}})$ and random elements $\widetilde{F}_{S,n} =_d \widehat{F}_{S,n}$, $\widetilde{F}_{T,n} =_d \widehat{F}_{T,n}$ and $\widetilde{B}_1$ and $\widetilde{B}_2$ independent Brownian bridges, defined on this space such that, with $\widetilde{\mathbf{P}}$-probability 1,

$$\sqrt{n}(\widetilde{F}_{S,n} - F_S) \to \widetilde{B}_1 \circ F_S \quad \text{and} \quad \sqrt{n}(\widetilde{F}_{T,n} - F_T) \to \widetilde{B}_2 \circ F_T \qquad \text{in } D_{0\gamma'}.$$

Fix $\widetilde{\omega}$ in a set of $\widetilde{\mathbf{P}}$-probability 1 such that the above relation holds and, in addition,

$$\int x\widetilde{F}_{S,n}(dx) \to \int xF_S(dx) \quad \text{and} \quad \int x\widetilde{F}_{T,n}(dx) \to \int xF_T(dx). \tag{7}$$

Write

$$g_S(x) = \widetilde{B}_1\big(F_S(x), \widetilde{\omega}\big) \quad \text{and} \quad g_T(x) = \widetilde{B}_2\big(F_T(x), \widetilde{\omega}\big).$$

Using (7), eventually the conditions of Theorem 4.1 are satisfied, and we obtain

$$\sqrt{n}\big(\Phi(\widetilde{F}_{S,n}, \widetilde{F}_{T,n}) - F_W\big) \to \Phi'_{(F_S,F_T)}(g_S, g_T) \quad \text{as } n \to \infty \text{ in } D_\infty.$$

Write $\widetilde{Z}$ for the process obtained by applying $\Phi'_{(F_S,F_T)}$ to the sample paths of $(\widetilde{B}_1 \circ F_S, \widetilde{B}_2 \circ F_T)$. Since this map is linear and bounded, $\widetilde{Z}$ is Gaussian in D_∞. Further, $\sqrt{n}(\Phi(\widetilde{F}_{S,n}, \widetilde{F}_{T,n}) - F_W) \to_d \widetilde{Z}$ in D_∞, and this implies that $\sqrt{n}(\Phi(\widehat{F}_{S,n}, \widehat{F}_{T,n}) - F_W) \to_d Z$ in D_∞. □

The process Z has zero means and its covariances can be calculated by tracking through the processes obtained at each stage of the decomposition of the functional. Defining

$$\begin{aligned} c(t,u) = &\int\int F_T\big((t-x) \wedge (u-y)\big)F_{-S}(dx)F_{-S}(dy) \\ &+ \int\int F_{-S}\big((t-x) \wedge (u-y)\big)F_T(dx)F_T(dy), \end{aligned}$$

we find that

$$\begin{aligned} &\operatorname{cov}\big(Z(t,\cdot), Z(u,\cdot)\big) \\ &\quad = \int_{[0,u]}\int_{[0,t]}\int\int c(-t-x-v, -u-y-w)U(dv)U(dw)F_W(dx)F_W(dy) \\ &\qquad - 2H \star F(t)H \star F(u), \end{aligned} \tag{8}$$

where

$$H(x) = \begin{cases} F \star U(-x), & \text{if } x \geq 0, \\ 0, & \text{if } x < 0, \end{cases}.$$

$F = F_{-S} \star F_T$ and U is the renewal function associated with F.

In order to assess the precision of the estimator, some form of confidence region is needed. Since we regard the estimate as an element of D_∞, it is natural to consider confidence regions in this function space. This leads to a consideration of simultaneous (rather than pointwise) confidence bands for the unknown distribution function. It would be straightforward to obtain asymptotically correct $100\alpha\%$ confidence bands if the distribution of $\|Z\|_\infty$ were known, because if $\mathbf{P}(\|Z\|_\infty \leq q(\alpha)) = \alpha$ and $\mathbf{P}(\|Z\|_\infty = q(\alpha)) = 0$, then Theorem 4.3 implies

$$\mathbf{P}\big(\sqrt{n}\|\Phi\big(\widehat{F}_{S,n}, \widehat{F}_{T,n}\big) - \Phi(F_S, F_T)\|_\infty \leq q(\alpha)\big) \to \alpha \quad \text{as } n \to \infty.$$

However, this distribution is unknown. To deal with this problem we use the bootstrap. Write $F^\star_{S,n}$ and $F^\star_{T,n}$ for the empirical distribution functions based on random samples of size n drawn from distributions $\widehat{F}_{S,n}$ and $\widehat{F}_{T,n}$, respectively. For large n, we estimate the distribution of $\|\sqrt{n}(\Phi(\widehat{F}_{S,n}, \widehat{F}_{T,n}) - \Phi(F_S, F_T))\|_\infty$ by that of $\|\sqrt{n}(\Phi(F^\star_{S,n}, F^\star_{T,n}) - \Phi(\widehat{F}_{S,n}, \widehat{F}_{T,n}))\|_\infty$. Monte Carlo methods are used to approximate the α quantile, $\widehat{q}_n(\alpha)$ of the distribution of

$$\big\|\sqrt{n}\big(\Phi(F^\star_{S,n}, F^\star_{T,n}) - \Phi(\widehat{F}_{S,n}, \widehat{F}_{T,n})\big)\big\|_\infty.$$

The confidence band is then calculated as

$$\Phi(\widehat{F}_{S,n}, \widehat{F}_{T,n}) \pm n^{-1/2} \times \text{approximation to } \widehat{q}_n(\alpha).$$

The justification for this procedure is given next and depends on the differentiability of the functional. The proof is not included here as it is a simple modification of the proof of Theorem 2.3 in Grübel and Pitts (1993), which in turn follows the steps given in Gill [(1989), Theorems 4 and 5].

THEOREM 4.4. *Suppose* $0 < \alpha < 1$. *Assume that* F_S *and* F_T *are continuous with*

$$\int x^{2\gamma} F_S(dx) < \infty, \qquad \int x^{2\gamma} F_T(dx) < \infty,$$

for some $\gamma > 2$. *Then*

$$\lim_{n \to \infty} \mathbf{P}\big(\sqrt{n}\|\Phi(\widehat{F}_{S,n}, \widehat{F}_{T,n}) - \Phi(F_S, F_T)\|_\infty \leq \widehat{q}_n(\alpha)\big) = \alpha.$$

If interest is in $F_W(x)$ for fixed x, then modification of the above result gives asymptotic validity of pointwise bootstrap confidence intervals for $F_W(x)$. From (8), $\mathrm{var}(Z(x))$ is complicated, and it seems that studentization to improve the order of accuracy of pointwise bootstrap confidence intervals is not practicable in this case.

5. Asymptotic efficiency. Asymptotic efficiency is used here in the sense of van der Vaart (1991), and is related to a convolution theorem, Theorem 2.1 in van der Vaart's paper. Efficiency of the stationary waiting time distribution function estimator will be proved by first showing that $(I-\widehat{F}_{S,n}, I-\widehat{F}_{T,n})$ (where I is $I_{[0,\infty)}$) is efficient for $(I-F_S, I-F_T)$ in $D_{0\gamma} \times D_{0\gamma}$, and then using the differentiability of the map taking $(I-F_S, I-F_T)$ to F_W ($\gamma > 2$, essentially Theorem 4.1) and applying Theorem 3.1 in van der Vaart (1991); see that paper for further details.

For $\gamma > 0$, let $\mathcal{P}_1$ be the set of probability measures μ on $\mathbf{R}$, concentrated on $(0,\infty)$ with $\int x^{2\gamma}\mu(dx) < \infty$. For μ in $\mathcal{P}_1$, let

$$T_1(\mu) = \left\{g \in L_2(\mu)\colon \int g\,d\mu = 0,\ \int |x|^{2\gamma} g^2(x)\mu(dx) < \infty\right\}, \tag{9}$$

and let $\mathcal{P}_1(\mu)$ be the collection of maps $[0,1] \to \mathcal{P}_1$, $t \mapsto \mu_t$, such that, as $t \downarrow 0$,

$$\int \left[\frac{1}{t}\{(d\mu_t)^{1/2} - (d\mu)^{1/2}\} - \frac{1}{2}g(d\mu)^{1/2}\right]^2 \to 0, \tag{10}$$

for some g in $T_1(\mu)$, and

$$\int |x|^{2\gamma}\mu_t(dx) \to \int |x|^{2\gamma}\mu(dx). \tag{11}$$

For μ in $\mathcal{P}_1$, using the construction in Groeneboom and Wellner [(1992), page 7], for every g in $T_1(\mu)$ we can find a path in $\mathcal{P}_1(\mu)$ such that (10) holds. We think of μ_S and μ_T as elements of $\mathcal{P}_1$.

Let $\mathcal{P}$ be the set of probability measures μ on $(\mathcal{X}, \mathcal{B}) = (\mathbf{R}^2, \mathcal{B}(\mathbf{R}^2))$ [$\mathcal{B}(\cdot)$ denotes the Borel σ-field] of the form $\mu = \mu_S \otimes \mu_T$ ($\otimes$ denotes product measure), where μ_S and μ_T are in $\mathcal{P}_1$. For μ in $\mathcal{P}$, let $T(\mu)$ be the set of g in $L_2(\mu)$ with $g(x,y) = g_S(x) + g_T(y)$ for some g_S in $T_1(\mu_S)$ and some g_T in $T_1(\mu_T)$ [so that $g_S(x) = \int g(x,y)\mu_T(dy)$ etc.]. Let $\mathcal{P}(\mu)$ be the set of paths $[0,1] \to \mathcal{P}$, $t \mapsto \mu_t = \mu_{S,t} \otimes \mu_{T,t}$ such that $t \mapsto \mu_{S,t}$ and $t \mapsto \mu_{T,t}$ are paths in $\mathcal{P}_1(\mu_S)$ and $\mathcal{P}_1(\mu_T)$ satisfying (10) with g_S and g_T, respectively. Then (10) is satisfied by $\{\mu_t\}$ with $g = g_S + g_T$. Let $\kappa\colon \mathcal{P} \to D_{0\gamma} \times D_{0\gamma}$ take $\mu_S \otimes \mu_T$ to $(I-F_S, I-F_T)$, where $D_{0\gamma} \times D_{0\gamma}$ has norm given by $\|(f,g)\| = \max\{\|f\|_{0\gamma}, \|g\|_{0\gamma}\}$ for f and g in $D_{0\gamma}$. Using (9)–(11), we have the following lemma.

LEMMA 5.1. *For $\mu \in \mathcal{P}$, there exists a bounded linear map $\kappa'_\mu\colon T(\mu) \to D_{0\gamma} \times D_{0\gamma}$ such that $(1/t)(\kappa(\mu_t) - \kappa(\mu)) \to \kappa'_\mu(g)$ for every path in $\mathcal{P}$ satisfying* (10). *Further, $\kappa'_\mu(g) = (\int I_{(\cdot,\infty)} g_S\,d\mu_S, \int I_{(\cdot,\infty)} g_T\,d\mu_T)$.*

Following van der Vaart (1991), a regular estimator sequence T_n for κ at μ in $(B,\mathcal{A}) = (D_{0\gamma} \times D_{0\gamma}, \mathcal{D}^P_{0\gamma} \times \mathcal{D}^P_{0\gamma})$ ($\mathcal{D}^P_{0\gamma} \times \mathcal{D}^P_{0\gamma}$ is the product σ-field) satisfies

$$\mathcal{L}_{\mu_{n^{-1/2}}}\left[\sqrt{n}\big(T_n - \kappa(\mu_{n^{-1/2}})\big)\right] \to L, \tag{12}$$

for every path $t \mapsto \mu_t$ in $\mathcal{P}(\mu)$, where L is a tight probability measure on $(B, \mathcal{A})$, and the convergence is convergence in distribution in $(B, \mathcal{A})$. By Lemma 5.1 and the convolution theorem [Theorem 2.1 in van der Vaart (1991)], there exists a tight Borel measure N_μ on B with $N_\mu[\overline{\kappa'_\mu(T(\mu))}] = 1$ and $N_\mu \circ b^{\star -1} = N(0, \|\tilde{\kappa}_{\mu, b^\star}\|_\mu^2)$ for all $b^\star \in B^\star$, where $B^\star$ is the space of all continuous linear maps $b^\star \colon B \to \mathbf{R}$; $\|f\|_\mu^2 = \int f^2\, d\mu$ for f in $L_2(\mu)$; and $\tilde{\kappa}_{\mu, b^\star}$ in $T(\mu)$ is the gradient of κ in direction $b^\star$ given by $b^\star \circ \kappa'_\mu(g) = \int \tilde{\kappa}_{\mu, b^\star} g\, d\mu$ for all $g \in T(\mu)$. By the convolution theorem, L is the same as the law of the sum of two independent random elements of $(B, \mathcal{B}(B))$, one of which has distribution N_μ; an estimator sequence T_n is asymptotically efficient for κ at μ in B [relative to $T(\mu)$ and $\mathcal{P}(\mu)$] if it satisfies (12) with $L = N_\mu$.

Let $e_{1,n}$ be the map from $(\mathbf{R}^n, \mathcal{B}(\mathbf{R}^n))$ to $(D_{0\gamma}, \mathcal{D}^P_{0\gamma})$ that takes $(x_1, \dots, x_n)$ to $I - (1/n)\sum_{i=1}^n I_{[x_i, \infty)}$, and let $E_{S,n}$ be $e_{1,n}(S_1, \dots, S_n)$. Define

$$E_n = (I - \widehat{F}_{S,n}, I - \widehat{F}_{T,n}) = \big(e_{1,n}(S_1, \dots, S_n), e_{1,n}(T_1, \dots, T_n)\big).$$

LEMMA 5.2. *Let $\gamma' > \gamma > 0$ and assume that F_S and F_T are continuous, $\int x^{2\gamma'} \mu_S(dx) < \infty$ and $\int x^{2\gamma'} \mu_T(dx) < \infty$. Then E_n is efficient for κ at $\mu = \mu_S \otimes \mu_T$ in $D_{0\gamma} \times D_{0\gamma}$ [relative to $T(\mu)$ and $\mathcal{P}(\mu)$].*

PROOF. Define w_u to be the map taking f in $D_{0\gamma}$ to $w_u(f) = (T_{0\gamma} f)(u)$ in $\mathbf{R}$ and consider $D'_{0\gamma} = \{w_u \colon u \in \mathbf{R}\}$, a subset of $D^\star_{0\gamma}$. It is easily checked that $\|f\|_{0\gamma} \le \sup_{w_u \in D'_{0\gamma}} |w_u(f)|$ for all f in $D_{0\gamma}$. Using (2.10) in van der Vaart (1991) and noting that, for $u \ge 0$, the gradient of the map taking μ_S to $I - F_S$ in direction w_u is $(1+u)^\gamma (I_{(u,\infty)}(\cdot) - (1 - F_S(u))$, we have that $w_u \circ E_{S,n}$ is efficient for $w_u \circ (I - F_S)$. Furthermore, using the conditions on F_S and Lemma 4.2, we have that $\{\mathcal{L}_{P_S}[\sqrt{n}(E_{S,n} - (I - F_S))]\}$ is tight in the sense given in van der Vaart (1991) and introduced by Dudley (1966). Hence, by Theorem 2.2 in van der Vaart (1991), $E_{S,n}$ is efficient for $I - F_S$ in $D_{0\gamma}$. The lemma follows by Theorem 4.1 of van der Vaart (1991). □

Theorem 4.1 can be modified with $1/t_n, t_n \to 0$ as $n \to \infty$, replacing $\sqrt{n}$ everywhere. Suppose now that $\gamma > 2$, and let U be the subset of $D_{0\gamma} \times D_{0\gamma}$ consisting of pairs $(I - F_S, I - F_T)$, where F_S and F_T are proper distribution functions with $F_S(0) = F_T(0) = 0$, $\int x^\gamma F_S(dx) < \infty$, $\int x^\gamma F_T(dx) < \infty$ and $\int x F_S(dx) < \int x F_T(dx)$. Define $\Psi \colon U \to D_\infty$ to be the map taking $(I - F_S, I - F_T)$ to F_W. With probability 1, E_n is in U eventually. When E_n is not in U, set $\Psi(E_n) = 0$, so that $\Psi(E_n)$ is $\Phi(\widehat{F}_{S,n}, \widehat{F}_{T,n})$. Then we see that the modified version of Theorem 4.1 means that Ψ is Hadamard differentiable at $(I - F_S, I - F_T)$ in U tangentially to $C_{0\gamma} \times C_{0\gamma}$ [as in van der Vaart (1991); see Gill (1989)]. Applying Theorem 3.1 of van der Vaart (1991) we obtain the following theorem.

THEOREM 5.3. *Let $\gamma' > \gamma > 2$ and suppose that F_S and F_T are continuous, $\int x F_S(dx) < \int x F_T(dx)$, $\int x^{2\gamma'} F_S(dx) < \infty$ and $\int x^{2\gamma'} F_T(dx) < \infty$. Then $\Psi(E_n)$ is efficient for F_W in D_∞ relative to $T(\mu_S \otimes \mu_T)$ and $\mathcal{P}(\mu_S \otimes \mu_T)$.*

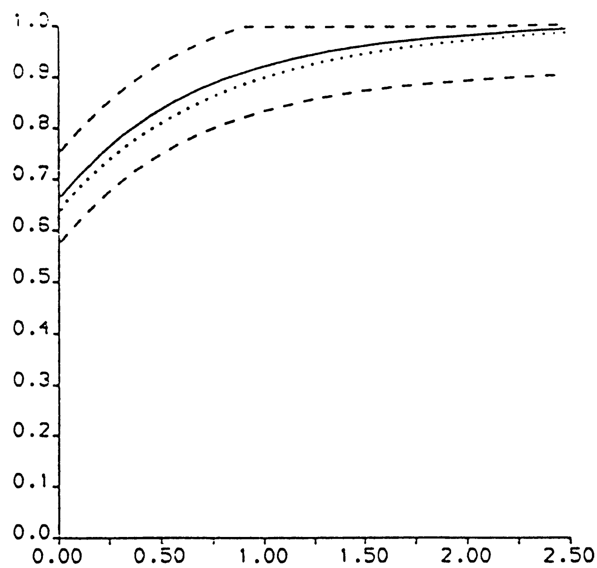

FIG. 1. *The distribution function F_W (dotted), the estimate (solid) and the bootstrap confidence band (broken lines) for $\rho = 0.5$, $n = 300$.*

6. Example. We illustrate the estimator for a $GI/M/1$ queue where the interarrival times are uniformly distributed on $(0, 2)$ and the service times are exponentially distributed with mean 0.5 (so ρ is 0.5). We simulate random samples from the input distributions and calculate (numerically) an approximation to the resulting estimate using the program GG1WAIT in Grübel (1991), which uses the fast Fourier transform algorithm. Figure 1 shows F_W, the estimate and an approximate 90% bootstrap confidence band when $n = 300$ using 500 bootstrap repetitions.

In other cases studied, in order to maintain comparable discrepancies, n needed to be larger for larger ρ-values. The most obvious feature is that the bootstrap confidence band is of constant width. It might well be judged more appropriate to have nonconstant-width simultaneous confidence bands for distribution function estimators, and this is an area for further investigation. One way to achieve this would entail analysis of the functional as a map to $D_{0\gamma}$, $\gamma > 0$.

In summary, we have established continuity and differentiability properties of the stationary waiting time functional, and these have been used to obtain strong consistency, asymptotic normality, asymptotic validity of the bootstrap and asymptotic efficiency of the stationary waiting time distribution function estimator. Similar reasoning would yield asymptotic efficiency of the empirical renewal function in Grübel and Pitts (1993). The proofs in the Appendix allow for easy modification and extension to corresponding results for a variety of stochastic models where quantities of interest are related to the Wiener–Hopf factors of a random walk. For other stochastic models, whether or not such an approach is fruitful depends on the analytic tractability of the functional in question.

APPENDIX

We need to consider functions $H: \mathbf{R} \to \mathbf{R}$ with $\sup_x |(T_{0,-1}H)(x)| < \infty$, but with $T_{0,-1}H$ not necessarily extendable to an element of D_∞. For notational convenience, we write $\|H\|_{0,-1}$ for $\sup_x |(T_{0,-1}H)(x)|$ in this case.

LEMMA A.1. *Let $\varepsilon > 0$. There exists a constant $c_1(\varepsilon)$ such that*

$$\|f \star H\|_{0,-1} \le c_1(\varepsilon)\|f\|_{2+\varepsilon,0}\|H\|_{0,-1}.$$

PROOF. The proof is similar to that of Lemma 3.2(ii) of Grübel and Pitts (1993). For $x < 0$,

$$\begin{aligned}|f \star H(x)| &\le \int_{(-\infty,0]} |f(x-y)|H(dy) + \sum_{k=1}^{\infty} \int_{(k-1,k]} |f(x-y)|H(dy) \\ &\le \|f\|_{2+\varepsilon,0}\,\|H\|_{0,-1} + \sum_{k=1}^{\infty} H(k) \sup_{y \le -k+1} |f(y)|,\end{aligned}$$

and $H(k)\sup_{y\le -k+1}|f(y)| \le (1+k)k^{-2-\varepsilon}\|f\|_{2+\varepsilon,0}\|H\|_{0,-1}$.

For $x \ge 0$ use similar arguments on noting that

$$\begin{aligned}(1+x)^{-1}|f \star H(x)| &\le (1+x)^{-1} \textstyle\int_{(-\infty,x]} |f(x-y)|H(dy) \\ &\quad + (1+x)^{-1} \textstyle\sum_{k=1}^{\infty} \int_{(x+k-1,x+k]} |f(x-y)|H(dy).\end{aligned}$$ □

PROOF OF THEOREM 3.1. The proof of the lemma below is straightforward.

LEMMA A.2. *Let $\{F_n\}_{n\in \mathbf{N}_0}$ and $\{G_n\}_{n\in\mathbf{N}_0}$ be distribution functions with, for n in $\mathbf{N}_0$, $F_n(0) = 0$ and $G_n(0) = 1$, and with G_n in $D_{\gamma 0}$ for some $\gamma \ge 0$. Assume that*

$$\|F_n - F_0\|_\infty \to 0 \quad \text{and} \quad \|G_n - G_0\|_{\gamma 0} \to 0 \quad \text{as } n \to \infty.$$

Then

$$\|F_n \star G_n - F_0 \star G_0\|_{\gamma 0} \to 0 \quad \text{as } n \to \infty.$$

The next lemma refers to stochastic ordering of distribution functions. For distribution functions F and G, F is stochastically smaller than or equal to G, $F \preceq G$, if $1 - F(x) \le 1 - G(x)$ for all x in $\mathbf{R}$.

LEMMA A.3. *Let $\gamma > 0$ and $\rho > 0$. Let $\{F_n\}_{n \in \mathbf{N}_0}$ be distribution functions with, for all n in $\mathbf{N}_0$,*

$$\int_{(-\infty,\,0]} |x|^{\gamma} F_n(dx) < \infty \quad \text{and} \quad \int_{(0,\,\infty)} x^{\rho} F_n(dx) < \infty,$$

and

$$\|F_n - F_0\|_{\gamma 0} \to 0 \quad \text{as } n \to \infty.$$

Then there exists a sequence $\{G_n\}_{n \in \mathbf{N}}$ of distribution functions, with $I_{[0,\,\infty)} - G_n$ in $D_{\gamma\rho}$, $G_n \preceq F_m$, for $m \geq n$ and $m = 0$, and $\|G_n - F_0\|_{\gamma\rho} \to 0$ as $n \to \infty$.

PROOF. Let $G_n = F_0 \vee \sup_{j \geq n} F_j$, so that $G_n \preceq F_j$ for $j \geq n$ and $j = 0$. For $t < 0$,

$$(1 + |t|)^{\gamma} |G_n(t) - F_0(t)| \leq \sup_{j \geq n} \|F_j - F_0\|_{\gamma 0}. \tag{13}$$

For any $t_0 > 0$ and $0 \leq t \leq t_0$,

$$(1 + t)^{\rho} |G_n(t) - F_0(t)| \leq (1 + t_0)^{\rho} \sup_{j \geq n} \|F_j - F_0\|_{\gamma 0}. \tag{14}$$

For $t > t_0$,

$$\begin{aligned} (1 + t)^{\rho} |G_n(t) - F_0(t)| &\leq \sup_{j \geq n} |F_j(t) - F_0(t)| I_{\{F_j(t) > F_0(t)\}}(t) \\ &\leq (1 + t)^{\rho} \big(1 - F_0(t)\big) \\ &\leq \int_{[t_0,\,\infty)} (1 + x)^{\rho} F_0(dx). \end{aligned} \tag{15}$$

From (13), (14) and (15), $\|G_n - F_0\|_{\gamma\rho} \to 0$. The remaining properties of G_n are easily checked. □

LEMMA A.4. *Let $\{F_n\}_{n \in \mathbf{N}_0}$ be distribution functions with, for all n and for some $\gamma > 2$,*

$$\int x F_n(dx) > 0, \qquad \int_{(0,\,\infty)} x^2 F_n(dx) < \infty, \qquad \int_{(-\infty,\,0]} |x|^{\gamma} F_n(dx) < \infty$$

and

$$\|F_n - F_0\|_{\gamma 0} \to 0 \quad \text{as } n \to \infty.$$

Let $-M_n$ be the infimum of the random walk with step distribution F_n, and let $F_{M,n}$ be the distribution function of M_n. Then

$$\|F_{M,n} - F_{M,0}\|_{\infty} \to 0 \quad \text{as } n \to \infty.$$

PROOF. By Lemma A.3 there exists a distribution function G, with positive mean and finite second moment, such that $G \preceq F_0$ and $G \preceq F_n$ for all n large enough. For n and k in $\mathbf{N}$, define

$$H_{n,k} = \begin{cases} I_{[0,\infty)}, & \text{if } k = 1, \\ \sum_{i=0}^{k-1} F_n^{\star(k-1-i)} \star F_0^{\star i}, & \text{if } k \geq 2, \end{cases}$$

and

$$H_n = \sum_{k=1}^{\infty} \frac{H_{n,k}}{k},$$

so that, if V_n is the harmonic renewal function associated with F_n, we have $V_n - V_0 = (F_n - F_0) \star H_n$. Since $\|H_n\|_{0,-1} \leq \|U_G\|_{0,-1} < \infty$, where U_G is the renewal function associated with G, we have $\|V_n - V_0\|_{0,-1} \to 0$ as $n \to \infty$ by Lemma A.1.

Let V_{n+} be the harmonic renewal function associated with the first strict ascending ladder height for the random walk with step distribution $\lim_{x\uparrow \cdot}(1 - F(-x))$. For $x > 0$ we have $V_{n+}(x) = V_n(0-) - V_n(-x-)$ and, using (6),

$$\begin{aligned} &|F_{M,n}(x) - F_{M,0}(x)| \\ &\quad \leq \exp\{-V_n(0-)\}|(V_{n+} - V_{0+}) \star J_n(x)| \\ &\quad\quad + |\exp\{-V_n(0-)\} - \exp\{-V_0(0-)\}| \sum_{k=0}^{\infty} \frac{1}{k!} V_{0+}^{\star k}(x), \end{aligned}$$

where

$$J_{n,k} = \begin{cases} I_{[0,\infty)}, & \text{if } k = 1, \\ \sum_{i=0}^{k-1} V_{n+}^{\star(k-1-i)} \star V_{0+}^{\star i}, & \text{if } k > 1, \end{cases}$$

and

$$J_n = \sum_{k=1}^{\infty} \frac{1}{k!} J_{n,k}.$$

For n large enough,

$$|(V_{n+} - V_{0+}) \star J_n(x)| \leq \|V_{n+} - V_{0+}\|_{\infty} \exp\{V_0(0-) + 1\}.$$

Since $\|V_n - V_0\|_{0,-1} \to 0$ and $\sum_{k=0}^{\infty}(1/k!)V_{0+}^{\star k}(x) \leq \exp\{V_0(0-)\} < \infty$, the lemma is proved. □

Theorem 3.1 now follows from Lemmas A.2 and A.4. □

Proof of Theorem 4.1.

LEMMA A.1. *Let* $\{F_n\}_{n\in\mathbf{N}_0}$ *and* $\{G_n\}_{n\in\mathbf{N}_0}$ *be distribution functions with, for* n *in* $\mathbf{N}_0$ *and some* $\gamma \ge 0$, G_n *in* $D_{\gamma 0}$, $G_n(0) = 1$ *and* $F_n(0) = 0$. *Suppose*

$$\sqrt{n}(F_n - F_0) \to f \quad \text{in } D_\infty \quad \text{and} \quad \sqrt{n}(G_n - G_0) \to g \quad \text{in } D_{\gamma 0},$$

where f *in* D_∞ *and* g *in* $D_{\gamma 0}$ *are continuous. Then*

$$\sqrt{n}(F_n \star G_n - F_0 \star G_0) \to g \star F_0 + f \star G_0 \quad \text{as } n \to \infty \text{ in } D_{\gamma 0}.$$

PROOF. For $x < 0$ and n large enough,

$$\begin{aligned}(1+|x|)^\gamma \big|\{\sqrt{n}(F_n \star G_n - F_0 \star G_0) - f \star G_0 - g \star F_0\}(x)\big| \\ \le \|\sqrt{n}(G_n - G_0) - g\|_{\gamma 0} + \|\sqrt{n}(F_n - F_0) - f\|_\infty \{\|G_0\|_{\gamma 0} + 1\} \\ + (1+|x|)^\gamma |(f \star G_n - f \star G_0)(x)|.\end{aligned}$$

For the third term, use the continuity of f to approximate it by a polynomial over $[0, t_0]$, $t_0 > 0$, and use $\lim_{x\to\infty} f(x) = 0$. Similar arguments work for $x \ge 0$. □

The next lemma is the same as Grübel and Pitts [(1993), Proposition 3.11].

LEMMA A.2. *Let* $\{F_n\}_{n\in\mathbf{N}_0}$ *be distribution functions and* $\gamma > 2$. *Suppose that, for all* n,

$$\int x F_n(dx) > 0, \qquad \int |x|^\gamma F_n(dx) < \infty \quad \text{and} \quad \|F_n - F_0\|_{\gamma\gamma} \to 0 \quad \text{as } n \to \infty.$$

Let U_n *be the renewal function associated with* F_n. *Then*

$$\|U_n - U_0\|_{0,-1} \to 0 \quad \text{as } n \to \infty.$$

LEMMA A.3. *Let* α *and* β *be such that* $0 \le \beta < \alpha < \infty$, *and let* g *be in* $D_{\alpha\alpha}$. *Then, for each* $\varepsilon > 0$, *there exists a function* g_ε *which is a linear combination of indicator functions of intervals of the form* $[a, b)$, $-\infty < a < b < \infty$, *such that*

$$\|g - g_\varepsilon\|_{\beta\beta} < \varepsilon.$$

This is similar to Grübel and Pitts [(1993), Lemma 3.12].

LEMMA A.4. *Let* $\{F_n\}_{n\in\mathbf{N}_0}$ *be distribution functions with, for all* n *and for some* $\gamma > 2$, $\int x F_n(dx) > 0$ *and* $\int |x|^\gamma F_n(dx) < \infty$. *Assume that*

$$\|\sqrt{n}(F_n - F_0) - g\|_{\gamma 0} \to 0 \quad \text{as } n \to \infty,$$

where g *is continuous and in* $D_{\gamma\gamma}$. *Let* $F_{M,n}$ *be as in Lemma* A.4. *Then*

$$\|\sqrt{n}(F_{M,n} - F_{M,0}) + h \star F_{M,0}\|_\infty \to 0 \quad \text{as } n \to \infty,$$

where

$$h(t) = \begin{cases} g \star U(-t), & \text{if } t \geq 0, \\ 0, & \text{if } t < 0, \end{cases}$$

and U is the renewal function associated with F_0.

PROOF. Let G be as in Lemma A.4. Using notation from that lemma, we have

$$\begin{aligned} &\|\sqrt{n}(V_n - V_0) - g \star U\|_{0,-1} \\ (16) \qquad &\leq \left\|\left(\sqrt{n}(F_n - F_0) - g\right) \star H_n\right\|_{0,-1} + \|g \star H_n - g \star U\|_{0,-1}. \end{aligned}$$

The first term tends to zero by Lemma A.1. For the second term, by Lemma A.7 we can find $\{g_l\}_{l \in \mathbf{N}}$ such that g_l is a linear combination of indicator functions of intervals of the form $[a, b)$, $-\infty < a < b < \infty$, with $\|g_l - g\|_{\gamma'\gamma'} \to 0$ as $l \to \infty$ for $2 < \gamma' < \gamma$, so that

$$\begin{aligned} (17) \qquad \|g \star H_n - g \star U\|_{0,-1} \leq \|(g_l - g) \star U\|_{0,-1} + \|(g_l - g) \star H_n\|_{0,-1} \\ + \|g_l \star H_n - g_l \star U\|_{0,-1}. \end{aligned}$$

It is easy to show that the first two terms on the right-hand side of (17) tend to zero, using Lemma A.1 and $\|H_n\|_{0,-1} \leq \|U_G\|_{0,-1}$ for all n large enough. For the last term in (17) we use the following result.

If G_1 and G_2 are distribution functions with positive means and finite second moments such that $G_1 \leq F_n \leq G_2$, then, writing U_i for the renewal function associated with G_i, $i = 1, 2$, we have

$$(18) \qquad \|g_l \star H_n - g_l \star U\|_{0,-1} \leq c_2(g_l)\|U_2 - U_1\|_{0,-1},$$

where $c_2(g_l)$ is a constant depending on g_l but not on G_1 and G_2. This follows from

$$\|I_{[a,b)} \star H_n - I_{[a,b)} \star U\|_{0,-1} \leq \left(2 + |a| + |b|\right)\|U_2 - U_1\|_{0,-1}$$

and the triangle inequality. We apply Lemma A.3 to $\{F_n\}$ and to the distribution functions $1 - F_n(-x-)$ to obtain G_1 and G_2, each having positive mean and finite second moment, with $G_1 \leq F_n \leq G_2$ for $n = 0$ and all n large enough, and such that $\|U_1 - U_2\|_{0,-1}$ is arbitrarily small. This is possible by Lemma A.6 since we can choose G_1 and G_2 arbitrarily close to F_0 in $\|\cdot\|_{\gamma\gamma}$. Applying (18) we see that $\|g \star H_n - g \star U\|_{0,-1} \to 0$ as $n \to \infty$, and so, from (16),

$$(19) \qquad \|\sqrt{n}(V_n - V_0) - g \star U\|_{0,-1} \to 0 \quad \text{as } n \to \infty.$$

Hence $\|\sqrt{n}(V_{n+} - V_{0+}) - h_1\|_\infty \to 0$ as $n \to \infty$, where, noting that $g \star U$ is continuous,

$$h_1(t) = \begin{cases} g \star U(0) - g \star U(-t), & \text{if } t > 0, \\ 0, & \text{otherwise.} \end{cases}$$

Writing $K_n = \sum_{k=0}^{\infty}(1/k!)V_{n+}^{\star k}$, we have

$$\begin{aligned}
&\|\sqrt{n}(F_{M,n} - F_{M,0}) + h \star F_{M,0}\|_\infty \\
&\quad \le \left|\exp\{-\|V_{n+}\|_\infty\} - \exp\{-\|V_{0+}\|_\infty\}\right| \|h_1 \star K_0\|_\infty \\
&\qquad + \|\sqrt{n}(K_n - K_0) - h_1 \star K_0\|_\infty \\
&\qquad + \Big|\sqrt{n}\Big(\exp\{-\|V_{n+}\|_\infty\} - \exp\{-\|V_{0+}\|_\infty\}\Big) \\
&\qquad\qquad + \exp\{-\|V_{0+}\|_\infty\}g \star U(0)\Big| \|K_0\|_\infty.
\end{aligned}$$

The first and the third terms tend to zero by (19). For the remaining term,

$$\begin{aligned}
&\|\sqrt{n}(K_n - K_0) - h_1 \star K_0\|_\infty \\
&\quad \le \left\|\left(\sqrt{n}(V_{n+} - V_{0+}) - h_1\right) \star J_n\right\|_\infty + \|h_1 \star J_n - h_1 \star K_0\|_\infty \\
&\quad \le \left\|\left(\sqrt{n}(V_{n+} - V_{0+}) - h_1\right) \star J_n\right\|_\infty + |g \star U(0)|\, \|J_n - K_0\|_\infty \\
&\qquad + \|h \star J_n - h \star K_0\|_\infty.
\end{aligned}$$

By arguments similar to those of Lemma A.4, the first two terms can be made arbitrarily small by taking n large enough. Note that h satisfies $h(t) = 0$ for $t < 0$ and $h(t) \to 0$ as $t \to \infty$. Approximating h by a polynomial over $[0, t_0]$ for t_0 large enough, we can show that $\|h * J_n - h * K_0\|_\infty \to 0$ as $n \to \infty$, and the lemma is proved. □

Theorem 4.1 now follows from Lemmas A.5 and A.8. □

Acknowledgments. I would like to thank Rudolf Grübel for many helpful discussions in the course of this work. The proof of Lemma A.3 was suggested by a referee. I would like to thank both the referee and an Associate Editor for comments leading to substantial improvements in the paper.

REFERENCES

ALSMEYER, G. (1991). Some relations between harmonic renewal measures and certain first passage times. *Statist. Probab. Lett.* **12** 19–27.

ASMUSSEN, S. (1987). *Applied Probability and Queues.* Wiley, New York.

BASAWA, I. V. and PRABHU, N. U. (1988). Large sample inference from single server queues. *Queueing Systems* **3** 289–304.

BOROVKOV, A. A. (1976). *Stochastic Processes in Queueing Theory.* Springer, New York.

DALEY, D. J. (1980). Tight bounds for the renewal function of a random walk. *Ann. Probab.* **8** 615–621.

DUDLEY, R. M. (1966). Weak convergence of probabilities on nonseparable metric spaces and empirical measures on Euclidean spaces. *Illinois J. Math.* **10** 109–126.

FELLER, W. (1971). *An Introduction to Probability Theory and Its Applications*, **2** 2nd ed. Wiley, New York.

GAVER, D. P. and JACOBS, P. A. (1988). Nonparametric estimation of the probability of a long delay in the $M/G/1$ queue. *J. Roy. Statist. Soc. Ser. B* **50** 392–402.

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

GROENEBOOM, P. and WELLNER, J. A. (1992). *Information Bounds and Nonparametric Maximum Likelihood Estimation.* Birkhäuser, Boston.

GRÜBEL, R. (1989). The fast Fourier transform in applied probability theory. *Nieuw Arch. Wisk.* **7** 289–300.

GRÜBEL, R. (1991). Algorithm AS 265: $G/G/1$ via fast Fourier transform. *J. Roy. Statist. Soc. Ser. C* **40** 355–365.

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). Nonparametric estimation in renewal theory I: The empirical renewal function. *Ann. Statist.* **21** 1431–1451.

HEYDE, C. C. (1964). Two probability theorems and their application to some first passage problems. *J. Austral. Math. Soc.* **4** 214–222.

HEYDE, C. C. (1988). Asymptotic efficiency results for the method of moments with application to estimation for queueing processes. In *Queueing Theory and its Applications: Liber Amicorum for J. W. Cohen* (O. J. Boxma and R. Syski, eds.) 405–411. North-Holland, Amsterdam.

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

PYKE, R. and SHORACK, G. R. (1968). Weak convergence of a two-sample empirical process and a new approach to Chernoff–Savage theorems. *Ann. Math. Statist.* **39** 755–771.

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

VAN DER VAART, A. W. (1991). Efficiency and Hadamard differentiability. *Scand. J. Statist.* **18** 63–75.

DEPARTMENT OF STATISTICAL SCIENCE
UNIVERSITY COLLEGE LONDON
GOWER STREET
LONDON WC1E 6BT
ENGLAND