Asymptotic inference for waiting times and patiences in queues with abandonment

Gorst-Rasmussen, Anders; Hansen, Martin Bøgsted

Publication date:
2007

Document Version
Publisher's PDF, also known as Version of record

AALBORG UNIVERSITY

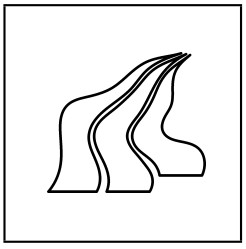
Asymptotic inference for waiting times and patiences in queues with abandonment

by

Anders Gorst-Rasmussen and Martin B. Hansen

R-2007-14 July 2007

Department of Mathematical Sciences
Aalborg University
Fredrik Bajers Vej 7 G ▪ DK - 9220 Aalborg Øst ▪ Denmark
Phone: +45 96 35 80 80 ▪ Telefax: +45 98 15 81 29
URL: www.math.auc.dk/research/reports/reports.htm



ISSN 1399–2503 ▪ On-line version ISSN 1601–7811

Asymptotic Inference for Waiting Times and Patiences in Queues with Abandonment

Anders Gorst-Rasmussen
Department of Mathematical Sciences, Aalborg University, Frederik Bajers Vej 7G, 9220 Aalborg East, Denmark
email: gorst@math.aau.dk

Martin B. Hansen
Department of Mathematical Sciences, Aalborg University, Frederik Bajers Vej 7G, 9220 Aalborg East, Denmark
email: mbh@math.aau.dk

Abstract

Motivated by applications in call center management, we propose a framework based on empirical process techniques for inference about the waiting time and patience distribution in multiserver queues with abandonment. The framework rigorises heuristics based on survival analysis of independent and identically distributed observations by allowing correlated successive waiting times. Assuming a regenerative structure of the sequence of offered waiting times, we establish asymptotic properties of estimators of limiting distribution functions and derived functionals. We discuss construction of bootstrap confidence intervals and statistical tests, including a simple bootstrap two-sample test for comparing patience distributions. The methods are exemplified in a small simulation study, and a real data example is given involving comparison of patience distributions for two customer classes in a call center.

Key words: Queues with abandonment; regenerative sequence; empirical process; dependent survival data; tele-queues.

1. Introduction. In a queuing system with abandonment, customers may abandon the waiting line before being serviced. This leads to right-censored waiting times where offered waiting times in the queue without abandonment are censored by random customer patiences. Models for queues with abandonment are of practical interest when designing and analyzing call centers where abandonment may considerably affect performance (Garnett et al. [15]). There has recently been a surge of interest in empirical applications of queuing models with abandonment to running call centers for which detailed call-by-call data are available. Statistical analyses of such data can provide both quantitative measures of performance and quality of service, as well as offer valuable insight into the qualitative nature of customer abandonment. This was demonstrated by Brown et al. [8], who applied methods from classical survival analysis to estimate cumulative distribution functions (CDFs) of waiting times and patiences, hazard rates, and related functionals. However, positive correlation of successive waiting times generally invalidates the asymptotic theory classically used to derive interval estimates and statistical tests. As pointed out by Gans et al. [14], there is a need to develop survival analytic methods which are capable of providing confidence intervals and statistical tests for call-by-call data from queues with abandonment.

Nonparametric survival techniques for dependent observations have previously been studied in the literature under mixing assumptions, and include Kaplan-Meier estimation (Cai [11]), quantile estimation (Cai and Kim [9]), and hazard rate estimation (Cai [10]). The techniques rely on mixing assumptions for the observation sequence, and computation of confidence intervals and statistical testing is often difficult and case-specific. In the present paper, we assume that the sequence of offered waiting times is regenerative. Informally, this means that the waiting time sequence splits into IID random blocks of random lengths. The assumption of regenerative offered waiting times is satisfied by the widely used $GI/G/m$ queuing model under weak assumptions (Asmussen [1, Theorem XII.2.2]), with blocks defined by system-wide busy periods. Regenerativity of the offered waiting times extends to independently right-censored waiting times:

$$\widetilde{W}_n := \min\{W_n, P_n\}, \quad n \in \mathbb{N}, \tag{1}$$

with $\{W_n\}$ the individual customer offered waiting times and $\{P_n\}$ the individual IID customer patiences, which we assume independent of $\{W_n\}$. Regenerativity of the offered waiting times is not a special property of the $GI/G/m$ queuing model. It remains a valid model whenever the arrival and service time sequences are stationary, and the waiting time sequence splits into independent blocks. The latter happens, for example, if the queuing system restarts at fixed time points, as is often the case in call centers.

In the present paper, we show how the assumption of regenerativity, when combined with techniques from the theory of empirical processes, can be used to rigorise methods for analyzing waiting times and patiences in queues. From a practical perspective, regenerativity justifies the use of various resampling methods to obtain confidence intervals and statistical tests for parameters. Emphasis will be placed on a simple blockwise bootstrap resampling

technique. Besides from contributing tools for practical inference, the paper contributes to the limited literature on nonparametric inference for queuing systems using empirical processes; see for example Bingham and Pitts [6]; Bingham and Pitts [7] – or Hansen and Pitts [19] for statistical inference involving empirical processes of regenerative observations. We remark that while this paper deals specifically with inference about waiting times and patiences, the empirical process techniques discussed here apply also to estimators for other types of regenerative sequences.

The paper is organized as follows. In Section 2, we review basic empirical process techniques for regenerative observations and state a new result concerning the validity of a functional blockwise bootstrap. Section 3 describes estimation of CDFs, nonparametric two-sample testing for the patience CDF, and estimation of various functionals of the waiting time and patience CDF of interest in call center managing. Section 4 presents a discussion of the practical use of the framework together with a simulation study. Finally, Section 5 illustrates a selection of the procedures applied to real-world data.

2. Asymptotic inference for regenerative sequences. Consider a sequence $\{C_n : n \in \mathbb{N}_0\}$ of random cycles taking values in $\bigcup_{m\geq 0}\mathbb{R}^m$, with $C_1, C_2, \ldots$ independent and identically distributed (IID) and independent of C_0. Thus each C_i is a block of random variables of random length. Defining X_n to be the nth real-valued observation in $\{C_n : n \in \mathbb{N}_0\}$, the sequence of random variables $X = \{X_n : n \in \mathbb{N}\}$ is called a regenerative sequence. The first cycle C_0 is known as the delay of the regenerative sequence. We denote by ℓ_n the length of C_n, define the renewal sequence $T_{n+1} := \ell_n + T_n$ (letting $T_0 := 0$), and let $\tau_n := \inf\{m \geq 1 : T_m > n\} - 1$ be the number of complete, observed cycles at time n. We assume ℓ_1 to be nonlattice with finite expectation. Then X admits a limiting distribution P (Asmussen [1, Corollary VI.1.5]), in the sense that $X_n \to \mathrm{P}$ in total variation where

$$\mathrm{P}(\,\cdot\,) = \mathrm{E}\sum_{i=T_1+1}^{T_2} 1_{X_i\in\,\cdot\,}/\mathrm{E}\ell_1. \tag{2}$$

Nonparametric statistical methods for regenerative sequences use regenerative analogues of the Law of Large Numbers (LLN) and the Central Limit Theorem (CLT) to establish consistency and asymptotic distributional properties of estimators. Adequately general forms of these limit results come from the theory of empirical processes which concerns the asymptotic behavior of functional estimators of the form

$$\mathrm{P}_n(f) = n^{-1}\sum_{i=1}^{n}(f(X_i) - \mathrm{P}f), \quad f \in \mathcal{F}, \tag{3}$$

uniformly over a set of measurable real-valued functions $\mathcal{F}$. The sequence $\{\mathrm{P}_n(f) : f \in \mathcal{F}\}$ of stochastic processes is called an empirical measure. A detailed review of limit results for empirical processes of IID observations can be found in van der Vaart and Wellner [28]. Limit results for empirical processes of regenerative observations have received limited attention in the literature; see Leventhal [21] and Tsai [27]. In this paper, we restrict ourselves to discussing the use of empirical process theory for estimating the limiting CDF of a regenerative sequence, $F(\,\cdot\,) := \mathrm{P}(-\infty, \cdot\,]$. This is not contrived: as we shall explain, a 'good' estimator of F can be used to define 'good' estimators of a range of functionals of the form $\phi(F)$.

From observations $X_1, \ldots, X_n$ of a regenerative sequence, we may estimate F using the empirical CDF defined for $x \in \mathbb{R}$ by $F_n(x) := n^{-1}\sum_{i=1}^{n} 1_{X_i\leq x}$. The sequence $\{F_n\}$ is the empirical measure of $\mathcal{F} = \{1_{(-\infty,t]} : t \in \mathbb{R}\}$ and defines a sequence in the space $D(\mathbb{R})$ of real cadlag functions equipped with the supremum norm $\|\,\cdot\,\|_\infty$. A Vapnik-Cervonenkis argument (Pollard [24, p. 16]) and the limit theorems of Leventhal [21] immediately lead to regenerative analogues of the classical Glivenko-Cantelli (uniform LLN) and Donsker theorems (uniform CLT).

Theorem 2.1 (Regenerative Glivenko-Cantelli/Donsker) *Let X be a regenerative process satisfying $\mathrm{E}\ell_1 < \infty$, and denote by F the CDF of the limiting distribution of X. Then*

$$\|F_n - F\|_\infty \to 0, \quad \textit{in probability.}$$

If moreover $\mathrm{E}\ell_1^2 < \infty$ then there exists a centered tight Gaussian process H_F on $\mathbb{R}$ such that

$$n^{1/2}(F_n - F) \xrightarrow{d} \mathrm{H}_F,$$

where $\xrightarrow{d}$ denotes weak convergence in $D(\mathbb{R})$.

The precise meaning of weak convergence in $D(\mathbb{R})$ is that $\mathrm{E}^*\varphi(F_n) \to \mathrm{E}\varphi(\mathrm{H}_F)$ for bounded, continuous, real-valued functions φ where E^* denotes outer expectation. This general form of weak convergence is required since F_n is generally nonmeasurable when $D(\mathbb{R})$ is equipped with the supremum norm and the Borel σ-field.

Theorem 2.1 in theory allows for approximating the sampling distribution of functionals of $F_n - F$ from the limiting Gaussian process H_F. However, this result is of little practical use since the covariance function of H_F depends on X in a nontrivial manner, precluding construction of distribution-free statistics in general. Instead, resampling methods can be used, i.e. methods which utilise (random) subsets of data to approximate sampling distributions. The strong mixing property of regenerative sequences (Thorrison [26, Theorem 3.3]) in principle enables application of the method of functional subsampling (Wolf et al. [29]) and, under additional mixing assumptions, the moving blocks bootstrap (Naik-Nimbalkar and Rajarshi [20]). However, the performance of either method relies on complex preliminary calibrations which again depend on the statistic under investigation. We suggest a simpler alternative which utilises the intrinsic structure of regenerative sequences. Here resampling is performed by sampling with replacement among regenerative cycles rather than individual observations, extending the naive bootstrap idea of sampling with replacement from IID observations (Efron [13]) to regenerative sequences. This regenerative block bootstrap (RBB) has previously been studied for the case of inference for the mean (Athreya and Fuh [2]; Datta and McCormick [12]; Bertail and Clémençon [3]) and is described algorithmically below.

Algorithm 2.1 (Regenerative blockwise bootstrap)
Given observations $\{X_i : i \leq n\}$ *of* X*, let* $\theta_n := \theta_n(X_1, \ldots, X_n)$ *denote a statistic.*

(*i*) *Divide* $\{X_i : i \leq n\}$ *into regenerative cycles* $C_1, \ldots, C_{\tau_n}$.

(*ii*) *Conditionally on* $\{X_i : i \leq n\}$ *and* τ_n*, sample* $C_1^*, \ldots, C_{\tau_n}^*$ *with replacement from* $\{C_1, \ldots, C_{\tau_n}\}$.

(*iii*) *Define the bootstrapped sample* $\{X_i^* : i = 1, 2, \ldots, n_*\}$ *where* X_i^* *is the* i*th real-valued observation of* $\{C_1^*, \ldots, C_{\tau_n}^*\}$, $T_{i+1}^* := T_i^* + l_i^*$ *(taking* $T_1^* := 0$ *and* l_i^* *to be the length of* C_i^**), and* $n_* := T_{\tau_n+1}^*$.

(*iv*) *Compute* $\theta_n^* := \theta_n(X_1^*, \ldots, X_{n_*}^*)$.

Approximate the law of θ_n *by the conditional law of* θ_n^* *given* $\{X_i : i \leq n\}$.

In the present paper, we need validity of an empirical process version of the RBB where $\theta_n := F_n$ is the empirical CDF and $\theta_n^* =: F_n^*$ its bootstrapped counterpart. Validity of the RBB in this setting may be defined in terms of a distance d metrising weak convergence on $D(\mathbb{R})$ by requiring

$$d\big(n_*^{1/2}(F_n^* - F_n), \mathrm{H}_F\big) \to 0, \quad \text{in probability;} \tag{4}$$

where the 'in probability' statement is relative to the law governing the observations. This in turn implies that the RBB estimator F_n^* $n_*^{1/2}(F_n^* - F_n)$ is a consistent estimator of $n^{1/2}(F_n - F)$ in the sense that their d-distance tends to zero in probability as $n \to \infty$. Typically, d will be the dual bounded Lipschitz distance on $D(\mathbb{R})$ (van der Vaart and Wellner [28, p. 73]). Validity of the empirical process RBB has been investigated by Radulović [25] for a class of empirical processes with observations from a discrete atomic Markov chain. In the appendix, we give a short proof of validity in the sense of (4) of the RBB for general empirical processes under the assumptions of the uniform CLT for regenerative observations of Tsai [27]. For the case of the RBB for the empirical CDF, the validity result reads as follows.

Theorem 2.2 (Bootstrap validity) *Let* X *be a regenerative sequence with* $\mathrm{E}\ell_1^2 < \infty$*. Denote by* F *the CDF of the limiting distribution of* X *and let* F_n^* *be the CDF obtained from the RBB. Then* (4) *holds.*

Estimation of the sampling distribution of F_n alone is of limited interest in applications, and it is desirable to extend the asymptotic results above to general functionals of F_n (plugin estimators). The continuous mapping theorem ensures that the RBB works for continuous real-valued functions of F_n. Another versatile tool not restricted to real-valued statistics is a functional analogue of the finite-dimensional delta-method. With the notation of Algorithm 2.1, let θ_n be a statistic of regenerative observations $X_1, \ldots, X_n$, taking values in a normed space V, and denote by θ_n^* the bootstrapped statistic obtained using the RBB. Suppose that $\phi\colon V \to W$ for some normed space W is a mapping for which there is a bounded linear operator $\mathrm{d}\phi_\theta\colon V \to W$ satisfying $\sup_{h\in K} \|t^{-1}(\phi(\theta + th) - \phi(\theta)) - \mathrm{d}\phi_\theta(h)\| \to 0$ when $t \to 0$ for every compact set $K \subseteq V$. Then ϕ is called Hadamard differentiable at θ. The next result follows from Theorem 3.9.4 and Theorem 3.9.11 of van der Vaart and Wellner [28].

Theorem 2.3 (Functional delta-method.) *Assume that there exists* $\theta \in V$ *and* $r_n \uparrow \infty$ *such that* $r_n(\theta_n - \theta) \xrightarrow{d} T$ *for a tight random element* T*, and that the RBB estimator* $r_n(\theta_n^* - \theta)$ *is a consistent estimator of* T*. If* ϕ *is Hadamard differentiable at* θ *with derivative* $\mathrm{d}\phi_\theta$ *then* $r_n(\phi(\theta_n) - \phi(\theta)) \xrightarrow{d} \mathrm{d}\phi_\theta(T)$*, and the RBB estimator* $r_n(\phi(\theta_n^*) - \phi(\theta_n))$ *is a consistent estimator of* $r_n(\phi(\theta_n) - \phi(\theta))$.

If T is tight Gaussian, linearity of $\mathrm{d}\phi_\theta$ implies that $\mathrm{d}\phi_\theta(T)$ is also tight Gaussian. One reason why the functional delta-method is so useful is the chain rule of Hadamard differentiation (van der Vaart and Wellner [28, Lemma 3.9.3]). This allows one to establish the asymptotics of a complicated statistic by representing it as a composition of simpler Hadamard differentiable maps applied to the empirical CDF.

RBB-based confidence intervals can be constructed using Efron's percentile method (Efron [13]). Namely if θ_n is an estimator of a real-valued parameter θ, and θ_n^* is obtained from the RBB using Algorithm 2.1, an approximate $(1-\alpha-\beta)\times 100\%$ confidence interval for θ is given by $[\theta_n-\xi^*_{n,\beta},\theta_n-\xi^*_{n,1-\alpha}]$ where $\xi^*_{n,\gamma}$ is the upper γth percentile of the bootstrap distribution of $\theta_n^*-\theta_n$, that is, the largest value x satisfying $\mathrm{P}^*(\theta_n^*-\theta_n\geq x)\geq 1-\gamma$. The RBB confidence interval asymptotically has level $1-\alpha-\beta$, whenever the statistic θ_n is a continuous or Hadamard differentiable function of the empirical CDF.

3. Asymptotic inference for waiting times and patiences. Let $\widetilde{W}_1,\ldots,\widetilde{W}_n$ be right-censored waiting times from a queuing system, defined as in (1) so that the underlying offered waiting times are assumed to form a regenerative sequence and the patiences are assumed to be IID random variables. Observations take the form

$$(\widetilde{W}_1,\delta_1),\ldots,(\widetilde{W}_n,\delta_n), \tag{5}$$

where δ_i is the noncensoring indicator of $\widetilde{W}_i$. If we seek features of the waiting time distribution, censoring occurs when the customer abandons the queue and vice versa for the patience distribution. Inferential procedures for such observations can be investigated with the empirical process methods of the previous section. This leads to a qualitative description of estimator asymptotics which, when combined with resampling techniques, can be used quantitatively to construct confidence intervals and statistical tests. We shall consider resampling using the RBB, but other resampling methods (see the discussion preceding Algorithm 2.1) may also be used to infer sampling distributions of the estimators of this section.

Denote by F the limiting CDF of uncensored observations from (5). A basic problem is how to estimate F from the censored observations. We suggest to use the product-limit (or Kaplan-Meier) estimator,

$$F_n(t)=1-\prod_{i:\widetilde{W}_{(i)}\leq t}\Big(1-\frac{n-i}{n-i+1}\Big)^{\delta_{(i)}},$$

where $\widetilde{W}_{(i)}$ is the ith order statistic of $\widetilde{W}_1,\ldots,\widetilde{W}_n$ and $\delta_{(i)}$ the corresponding indicator of noncensoring. The asymptotic properties of F_n can be established using Theorem 2.1 and 2.3. Denote by $H^{\mathrm{uc}}(t):=\mathrm{P}(\widetilde{W}\leq t,\delta=1)$ the limiting subdistribution function of the uncensored observations and by $\overline{H}(t):=\mathrm{P}(\widetilde{W}\geq t)$ the limiting tail function of observations. A classical result from survival analysis (Gill and Johansen [16]) states that F can be obtained from $(\overline{H},H^{\mathrm{uc}})$ via the mappings

$$(\overline{H},H^{\mathrm{uc}})\overset{\alpha}{\longmapsto}\int_{[0,\,\cdot\,]}\overline{H}(s)^{-1}\mathrm{d}H^{\mathrm{uc}}(s)=:\Lambda\overset{\beta}{\longmapsto}\prod_{s\in(0,\,\cdot\,]}(1-\mathrm{d}\Lambda(s))=1-F.$$

Here Λ is the cumulative hazard rate, and $\prod_{s\in(0,t]}$ denotes the product integral over $(0,t]$. Then F_n is in fact the plugin estimator $\beta(\alpha(\overline{H}_n,H_n^{\mathrm{uc}}))$ where

$$H_n^{\mathrm{uc}}(t)=n^{-1}\sum_{i=1}^n\delta_i 1_{\widetilde{W}_i\leq t},\quad \overline{H}_n(t)=n^{-1}\sum_{i=1}^n 1_{\widetilde{W}_i\geq t}.$$

It can be shown (Gill and Johansen [16]) that each of α, β, then $\beta\circ\alpha$ are Hadamard differentiable at $(H^{\mathrm{uc}},\overline{H})$ when the latter is viewed as an element of $D[0,\tau]\times D[0,\tau]$ for some τ with $\overline{H}(\tau)>0$. Combining this with Theorem 2.1-2.3, we conclude that the product-limit estimator based on regenerative observations is consistent, asymptotically Gaussian and can be bootstrapped. So we can use the RBB to construct both pointwise confidence bands for F (by estimating the distribution of $F(t)$ for each t) and uniform confidence bands (by estimating the distribution of $\sup_{t\in[0,\tau]}|F(t)|$). Examples will follow in the next section. By similar arguments, one obtains consistency, asymptotic Gaussianity, and bootstrap validity for the Nelson-Aalen-type estimator $\Lambda_n:=\alpha(\overline{H}_n,H_n^{\mathrm{uc}})$ of the cumulative hazard rate. Estimates of functions relating to the (cumulative) hazard rate have previously been used to explore abandonment behavior of customers in a call center (Brown et al. [8]). Note that empirical process theory, although a powerful framework, essentially deals with inference using step functions (empirical measures) and does not lend itself towards methods for smooth estimation of, for example, densities or hazard rates. Smooth estimation procedures for censored sequences under mixing assumptions are discussed by Cai [10].

One may ask whether estimators of expectations or quantiles of F based on plugging in the product-limit estimator F_n in the formulas $\mathrm{E}\xi(X) = \int_0^\infty \xi(x)F(\mathrm{d}x)$ and $F^{-1}(p) := \inf\{x : F(x) \geq p\}$ inherit the nice asymptotic properties. Such statistics may arise as key performance indicators in call center managing, where one seeks summary statistics such as expected waiting times and patiences; or median waiting times and patiences (Nederlof and Anton [22]). If the largest observation is censored, the product-limit estimator is not a CDF and plugging it in the definition of the expectation will produce infinite values. Instead, one can estimate the truncated expectation from $\int_0^\tau \xi(x)F_n(\mathrm{d}x)$ where τ satisfies $\mathrm{P}(\tilde{W} \leq \tau) < 1$. Consistency, asymptotic Gaussianity, and bootstrap validity of this estimator follows from Lemma 3.9.17 of van der Vaart and Wellner [28] and Theorem 2.3. Note that this truncated expectation is a negatively biased estimator of $\mathrm{E}\xi(X)$ and should be interpreted with care. Similarly for quantiles of F, Lemma 3.9.20 of van der Vaart and Wellner [28] implies Hadamard differentiability of the mapping taking F to its pth percentile, whenever F has a strictly positive derivative at $F^{-1}(p)$. Theorem 2.3 again implies consistency, asymptotic Gaussianity, and bootstrap validity for the estimator of the pth percentile based on F_n.

We next consider the issue of how to formally test equality of two limiting patience CDFs from right-censored regenerative patiences. This problem has to the best of our knowledge not been considered previously, but is of relevance when comparing abandonment behavior of two customer classes in a call center. Assume that we have available two independent samples of the form (5) (with censoring when the customer is serviced) of sizes n and m, such that the limiting CDFs of uncensored observations are F and G, respectively, and the limiting CDFs of the censored observations are H and I. Denote by F_n and G_n the product-limit estimators of the CDFs, and let τ be such that $H(\tau) < 1$ and $I(\tau) < 1$. We seek to test the null hypothesis

$$H_0 \colon F(t) = G(t), \quad \forall\, t \in [0, \tau] \tag{6}$$

against the two-sided alternative $F \neq G$. Denote by W the common tight Gaussian limit of $n^{1/2}(F_n - F)$ and $m^{1/2}(G_m - G)$ under the null hypothesis. Define the test statistic

$$D_{n,m} := \sqrt{(nm)/(n+m)}\|(F_n - F) - (G_m - G)\|_\infty,$$

where $\|\cdot\|_\infty$ denotes supremum over the interval $[0, \tau]$, and assume that $nm/(n+m) \to \lambda \in (0, 1)$. Then, under the null hypothesis, the continuous mapping theorem implies $D_{n,m} \xrightarrow{d} \|W\|_\infty$. The distribution of the supremum $\|W\|_\infty$ is intractable and must be approximated by resampling techniques. To this end, define the bootstrapped counterpart of $D_{n,m}$ by

$$D^*_{n,m} = \sqrt{(n_* m_*)/(n_* + m_*)}\|(F^*_n - F_n) - (G^*_m - G_m)\|_\infty.$$

Here n_*, F^*_n and m_*, G^*_m are obtained by applying the RBB to each censored sample separately. The map $(A, B) \mapsto A - B$ is Hadamard differentiable on $(D[0, \tau])^2$. Theorem 2.3, Slutsky's lemma for the bootstrap (Radulović [25, Lemma 3.1]), and Theorem 2.1-2.2 together with the continuous mapping theorem implies consistency of $D^*_{n,m}$ as an estimator of $D_{n,m}$ as $n, m \to \infty$. So the conditional distribution of the bootstrapped test statistic $D^*_{n,m}$ may be used to define critical levels for the null hypothesis (6): if $\xi^*_{n,m,\alpha}$ is the upper α percentile of the RBB distribution $\mathrm{P}^*(D^*_{n,m} \leq \cdot\,)$, then H_0 is rejected at approximate level α if $\sqrt{mn/(m+n)}\|F_n - G_m\|_\infty > \xi^*_{n,m,\alpha}$. This essentially corresponds to constructing an $(1-\alpha) \times 100\%$ uniform confidence band for $F - G$ and rejecting H_0 at level α if the band does not contain the zero function. Analogous procedures with potentially better power are easily defined for other smooth 'discrepancy functionals' $(F, G) \mapsto \phi(F, G)$ than the difference: for example the odds ratio or the cumulative hazard ratio of two limiting CDFs – or weighted versions hereof.

The above approach to hypothesis testing (constructing confidence intervals by resampling and checking whether zero is contained in the interval) applies generally to simple hypotheses $H_0 : \theta_1 = \theta_2$ whenever consistent estimators $\hat{\theta}_{1n}$ and $\hat{\theta}_{2n}$ of θ_1 and θ_2 exist which are asymptotically Gaussian and can be bootstrapped. This in turn yields a method for rigorous empirical comparison of for example medians, probabilities, and expectations. Note that, in the case of inference for expectations with respect to the limiting distribution, more efficient RBB-methods based on the percentile t-method (Hall [18]) exist (Bertail and Clémençon [4]).

4. Practical Considerations and Simulation Examples. In the previous section, we discussed methods for qualitatively and quantitatively investigating properties of estimators from right-censored waiting times. The key was the asserted regenerative structure of the offered waiting times which enabled regenerative empirical process techniques to be applied. The assumption of regenerativity is often a reasonable and parsimonious model. It holds in the general $GI/G/m$-queuing model with regeneration occurring when all servers are idle (Asmussen [1, Theorem XII.2.2]), allowing regenerative cycles to be constructed whenever all such regeneration points have been identified in an observation sequence. In call centers, with many servers and high load, there may be few or no

system wide idle periods during a typical day of operation. On the other hand, if a regenerative model is adopted, forced regeneration occurs at the end of every day when the call center closes. This suggests that (a subset of) the waiting time sequence for each separate day of operation can be used to define regenerative cycles. This idea is not restricted to $GI/G/m$-type queuing systems, but applies to any queuing system for which independent and identically distributed cycles of waiting times can be defined. Stationarity of the cycle sequence can be checked empirically by investigating stationarity of a sequence of real-valued statistics calculated from the cycles (averages, variances etc.), for example using time series plots. A sufficient condition for cycle stationarity is stationarity of the underlying observation and cycle length sequence.

We performed a small simulation study to illustrate coverages of RBB confidence intervals for selected statistics of waiting times and patiences, as well as level and power of the two-sample RBB test for patience CDFs. In all experiments, we considered an $M/M/15$ queuing system with an arrival rate of 13.5 customers per minute and a service rate of 1 customer per minute, corresponding to a system load of 90%. Waiting times were right-censored with IID patiences from various distributions. Each regenerative block used in the RBB was simulated independently and comprised 15 minutes of observations following a 15 minute start-up period, corresponding to blocks of approximately 200 successive observations in the stationary regime. A start-up period was used solely for computational reasons: the inferential methods also apply in the transient regime, but are not easily compared with theoretical results.

A typical sequence of right-censored waiting times from an $M/M/15$ queuing system with exponential patiences is shown in Figure 1 (left). Observe the positive correlation between successive observations which precludes the use of standard statistical methods for IID data. In Figure 1 (right), an example of the estimated patience CDF (superimposed on the true patience CDF) is shown.

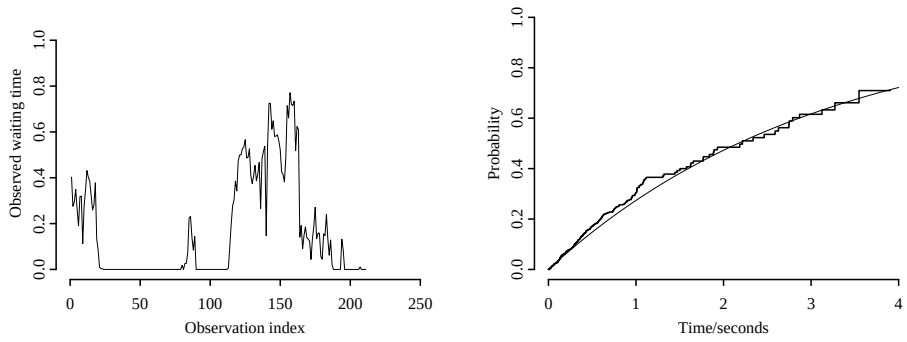


Figure 1: Left: Example of a waiting time sequence in an $M/M/15$ queue with an arrival rate of 13.5 customers per minute, a service rate of 1 customer per minute, and exponential patiences. Right: An estimate of the patience CDF (thick line) in the same queuing system system, superimposed on the true patience CDF (thin line).

Table 1 shows estimated coverages of RBB-confidence intervals for a selection of statistics of the right-censored waiting times. Observe that coverages are subject to sampling variation which can be quantified using standard methods for binomial proportions. All confidence intervals have been calculated using the percentile method. The estimated coverages in Table 1 are generally close to their nominal values, although the confidence intervals appear slightly anticonservative. We found that decreasing the rate of abandonment did not markedly impact coverage for estimates from the patience distribution, although quantile estimation becomes difficult when the rate of abandonment is small. This is due to the product-limit estimator having an atom at infinity if the largest observation is censored, frequently leading to infinite quantile estimates in the case of heavy censoring. The uniform confidence intervals and the corresponding coverages are calculated for the respective CDFs over the fixed interval $[0, 1.5]$ for all simulations. The estimated coverages for the uniform confidence intervals were sensitive to the choice of interval – too large intervals lead to poor coverages. In applications, one would typically use the interval ranging from zero to the largest uncensored observation of the sample.

The estimated level and power of the RBB two-sample test for two different types of patience distributions (exponential and lognormal with fixed logarithmic variance 1) are shown in Table 2. Test statistics were calculated over the fixed interval $[0, 1.5]$ for all simulations. The parameter of each patience distribution was adjusted to provide rates of abandonment of 20%, 10%, and 5%, respectively. The level of the test was estimated for each

rate of abandonment. We also estimated the power to detect a supremum distance deviation of 0.05, 0.1, and 0.2 from these reference patience distributions, letting each comparison distribution be stochastically larger than its reference counterpart. The test exhibits reasonable power properties, considering the small rate of abandonment: more detailed power assessments are difficult due to the lack of reference methods. The estimated levels suggest that the test is slightly conservative. As was the case for uniform RBB confidence intervals for CDFs, the test was sensitive to the choice of interval over which the test statistic was calculated.

		Coverage					
		Waiting times			Patiences		
Abandonment	$1-\alpha$	$F(1)$	$F^{-1}(0.5)$	$\|F\|_\infty$	$F(1)$	$F^{-1}(0.2)$	$\|F\|_\infty$
20%	0.90	0.84	0.97	0.85	0.87	0.90	0.93
	0.95	0.91	0.92	0.91	0.93	0.94	0.96
10%	0.90	0.87	0.88	0.85	0.88	0.86	0.90
	0.95	0.91	0.94	0.92	0.94	0.90	0.96
5%	0.90	0.84	0.86	0.82	0.85	0.33	0.90
	0.95	0.89	0.93	0.90	0.91	0.26	0.95

Table 1: Observed coverage of RBB confidence intervals for functionals of the patience CDF F in an $M/M/15$ queue with an arrival rate of 13.5 customers per minute, a service rate of 1 customer per minute, and exponential patiences. The parameter of each patience distribution was adjusted to provide the given rate of abandonment. Each figure is based on 500 independent simulations of a sequence of 25 IID blocks of average length 200. For each simulation, 4000 bootstrap replications were used. The statistic $\|F\|_\infty$ was calculated over the fixed interval $[0, 1.5]$.

			Exponential patience				Lognormal patience			
Abandonment	$1-\alpha$		Level	Power to detect Δ			Level	Power to detect Δ		
		Δ		0.05	0.10	0.20		0.05	0.10	0.20
20%	0.90		0.93	0.53	0.91	1.00	0.94	0.55	0.89	1.00
	0.95		0.98	0.31	0.79	0.99	0.99	0.31	0.80	0.99
10%	0.90		0.92	0.37	0.81	0.98	0.95	0.47	0.89	1.00
	0.95		0.96	0.22	0.65	0.95	0.98	0.38	0.79	0.98
5%	0.90		0.93	0.38	0.62	0.95	0.93	0.59	0.93	1.00
	0.95		0.97	0.11	0.46	0.87	0.97	0.41	0.85	0.99

Table 2: Observed level and power of the RBB two-sample test for detecting a difference of Δ between patience CDFs in an $M/M/15$ queue with an arrival rate of 13.5 customers per minute, a service rate of 1 customer per minute, and exponential/lognormal patience distributions. Parameters of the three reference patience distributions were adjusted to provide the given rates of abandonment (logarithmic variance of lognormal distribution fixed to 1). Comparison distributions were chosen stochastically larger than their reference distributions. Each figure is based on 500 independent simulations of a sequence of 25 IID blocks of expected length 200. For each simulation, 4000 bootstrap replications were used. The two-sample test statistic was calculated over the fixed interval $[0, 1.5]$.

5. Application to real data As an application of the methods of this paper, we considered inference from real data given by call logs from a call center of a small Israeli bank. See Brown et al. [8] for a detailed description and statistical analysis of the data. We extracted right-censored waiting times for all customers of the call center arriving during the period 2 p.m.–3 p.m. on ordinary Israeli weekdays (Sunday-Thursday) in November and December. This is representative of customer waiting experience during peak hours and would be of particular interest to a call center manager. We obtained 36 observation sequences of average length 139. Observation sequences of separate days were assumed independent. The assumption of stationarity of blocks was assessed by checking the sufficient condition of stationarity of the waiting time sequence, using time series plots and visual inspection of estimates of waiting time and patience distribution CDFs for different weekdays. We did not find evidence against the stationarity assumption.

In the following, estimates are presented as estimate (95% confidence interval). All interval estimates were constructed using the percentile method, using 4000 replications using the RBB on the 36 blocks. The product-limit estimates with uniform 95% confidence bands for the waiting time and patience CDFs are shown in Figure 2.

The median waiting time was 37 seconds (21-53), while the probability of waiting more than 3 minutes was 0.15 (0.11-0.20). The tail of the waiting time distribution is reasonably well estimated (Figure 2, left), so in this case it is meaningful to estimate the expected waiting time using the tail formula (truncating the product-limit estimate at the largest observation). The value was 81 seconds (63-98). The 20th upper percentile of the patience distribution was 52 (47-86), while the probability of having a patience greater than 3 minutes was 0.64 (0.61-0.68).

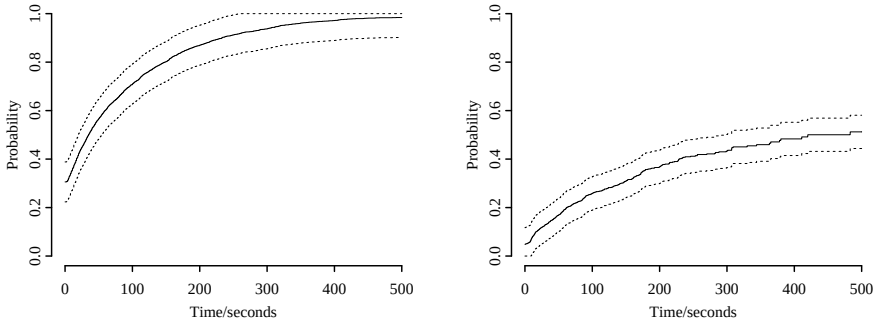


Figure 2: Left: Estimated waiting time CDF (solid line) with RBB-based 95% uniform confidence bands (dotted lines). Right: Estimated patience CDF (solid line) with RBB-based 95% uniform confidence bands (dotted lines). Observations used are for customers arriving between 2 p.m. and 3 p.m. on ordinary weekdays (Sunday-Thursday).

To illustrate the application of the RBB two-sample test, we considered comparison of patience CDFs of two different priority groups of stock market customers. We used censored waiting times collected on ordinary weekdays (Sunday-Thursday) in the period 8 a.m.–8 p.m. The large time interval was used to obtain a reasonable number of observed patiences, although waiting times are unlikely to be stationary over such an interval. For the framework of this paper, however, nonstationarity is not a theoretical issue: we only require blocks to be stationary (and independent), corresponding to the heuristic assumption that the different days of operation are 'stochastically similar'. We obtained 36 blocks of average length 170. Product-limit estimates of the CDFs are shown in Figure 3, left. Using 4000 replications in the RBB, we accepted the hypothesis of equal patience distributions, with a p-value of 0.07. To further explore the nature of the (nonsignificant) difference between the patience distribution, their absolute difference was plotted alongside a uniform 95% confidence band (Figure 3, right). There appears to be a borderline significant discrepancy around 500 seconds, indicating that patience distributions for the two customer classes may differ in the tails.

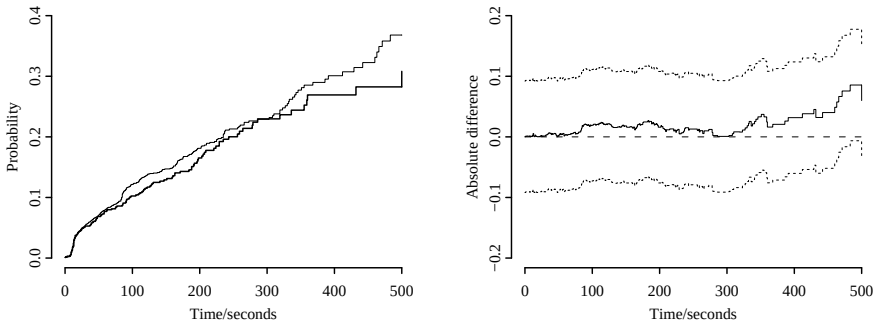


Figure 3: Left: Estimated patience CDF for regular stock markets customers (thick line) and priority stock market customers (thin line) arriving between 8 am and 8 pm on ordinary weekdays (Sunday-Thursday). Right: Estimated absolute distance between the two priority groups' CDFs (solid line) with uniform 95% confidence band (dotted lines).

Appendix A. Validity of the RBB for empirical processes. For definiteness, we assume the regenerative sequence X to be defined canonically in terms of the cycles $\{C_n : n \in \mathbb{N}_0\}$ which are given by the coordinate sequence on an infinite product space $(\Omega, \mathcal{B}, \mathrm{Q}) := (\tilde{\Omega}, \tilde{\mathcal{G}}, \mathrm{Q}') \otimes \prod_{n\geq 1}(\tilde{\Omega}, \tilde{\mathcal{G}}, \mathrm{Q}^*)$ where $\tilde{\Omega} = \bigcup_{m\geq 0}\mathbb{R}^m$ and $\tilde{\mathcal{G}}$ is the natural σ-algebra generated by $\bigcup_{n\geq 1}\mathcal{B}^n$ for the Borel σ-algebra $\mathcal{B}$ on $\mathbb{R}$.

The empirical process corresponding to the empirical measure (3) for a class of real-valued measurable functions $\mathcal{F}$ on $\mathbb{R}$ is the $\mathcal{F}$-indexed stochastic process $\{\mathrm{G}_n(f) : f \in \mathcal{F}\}$ where $\mathrm{G}_n(f) := n^{1/2}\mathrm{P}_n(f)$. The corresponding bootstrapped empirical process $\{\mathrm{G}_n^*(f) : f \in \mathcal{F}\}$ is given by

$$\mathrm{G}_n^*(f) := n_*^{1/2}\Big(n_*^{-1}\sum_{i=1}^{n_*} f(X_i^*) - \underline{n}^{-1}\sum_{i=1}^{\underline{n}} f(X_i)\Big), \qquad n \in \mathbb{N},\ f \in \mathcal{F};$$

with n_* and $\{X_i^* : i = 1, \ldots, n_*\}$ obtained from Algorithm 2.1, and $\underline{n} := T_{\tau_n+1}$. Each of G_n and G_n^* are viewed as functions with values in the metric space $\ell^\infty(\mathcal{F})$ of uniformly bounded real-valued functions on $\mathcal{F}$ equipped with the uniform norm $\|K\|_{\mathcal{F}} = \sup_{f\in\mathcal{F}}|K(f)|$.

The theorem below is the bootstrap variant of the uniform CLT by Leventhal [21] and Tsai [27]. It bears some similarities to the bootstrap uniform CLT by Radulović [25] for a class of empirical processes with observations from a discrete atomic Markov chain. However, our method of proof is distinct from his in that we avoid assuming mixing properties for the regenerative sequence and imposing bracketing conditions on the function class $\mathcal{F}$. Additionally, our approach uses Poissonization, implying that we can use the strategy of Giné and Zinn [17] to give a concise proof based on multiplier inequalities.

For a measure γ on $(\mathbb{R}, \mathcal{B})$, the $\mathcal{L}^p(\gamma)$ ε-covering number $N_p(\mathcal{F}, \varepsilon, \gamma)$ of $\mathcal{F}$ for some $\varepsilon > 0$ is the smallest number of $\mathcal{L}^p(\gamma)$ ε-balls needed to cover $\mathcal{F}$. The following combinatorial entropy is due to Pollard [23]

$$N_p(\varepsilon, \mathcal{F}) := \sup_\gamma N_p(\mathcal{F}, \varepsilon, \gamma),$$

where the supremum runs over finitely supported measures γ on $(\mathbb{R}, \mathcal{B})$. Recall that an envelope function F for $\mathcal{F}$ is any (measurable) real-valued function on Λ satisfying $f(\lambda) \leq F(\lambda)$ for all λ and f. To simplify our derivation, we assume in the following that $\mathcal{F}$ is sufficiently regular to ensure measurability of suprema of processes. See Leventhal [21] for details on the specific measurability assumptions required in the theorem.

THEOREM A.1 *Suppose that* $\mathrm{E}\ell_1^2 < \infty$. *Let* $\mathcal{F}$ *be a class of measurable real-valued functions on* $\mathbb{R}$ *with envelope function* F *such that*

$$\int_0^\infty \sqrt{\log N_2(\varepsilon, \mathcal{F})}\mathrm{d}\varepsilon < \infty, \qquad \mathrm{E}\Big(\sum_{i=T_1+1}^{T_2} F(X_i)\Big)^2 < \infty.$$

Under further measurability assumptions on $\mathcal{F}$, *there exists a tight, centered Gaussian process* $\mathrm{H_P}$ *on* $\ell^\infty(\mathcal{F})$ *such that* $\mathrm{G}_n \xrightarrow{d} \mathrm{H_P}$ *where* $\xrightarrow{d}$ *denotes weak convergence in* $\ell^\infty(\mathcal{F})$, *and the RBB is valid for the empirical process of* G_n *in the sense that*

$$d(\mathrm{G}_n^*, \mathrm{H_P}) \to 0, \quad \textit{in probability } (\mathrm{Q}) \tag{7}$$

where d *is dual bounded Lipschitz distance on* $\ell^\infty(\mathcal{F})$ *(van der Vaart and Wellner [28, p. 73]).*

PROOF. By Theorem 4.3 of Tsai [27], the hypotheses imply that G_n converges weakly in $\ell^\infty(\mathcal{F})$ to a tight, centered Gaussian process $\mathrm{H_P}$. Following Giné and Zinn [17], bootstrap validity holds if we can show the analogue of (7) for the finite-dimensional distributions of G_n^* and stochastic asymptotic equicontinuity in probability with respect to a totally bounded semimetric ρ on $\mathcal{F}$. The latter means that

$$\lim_{\delta\downarrow 0}\lim_n \|\mathrm{G}_n^*\|_{\mathcal{F}_\delta} = 0, \quad \text{in probability } (\mathrm{Q}),$$

where $\|K\|_{\mathcal{F}_\delta} := \sup\{|K(f) - K(g)| : \rho(f, g) < \delta\}$ for $K \in \ell^\infty(\mathcal{F})$. Additionally, it must hold that ρ makes $\mathrm{H_P}$ uniformly equicontinuous. As shown in Tsai [27], our assumptions imply that $\mathcal{F}$ is totally bounded in $\mathcal{L}^2(\mathrm{P})$ and that G_n is asymptotically $\mathcal{L}^2(\mathrm{P})$-equicontinuous. By elementary properties of $\mathcal{L}^p$-seminorms, both properties also hold for $\mathcal{L}^1(\mathrm{P})$-seminorm. Theorem 1.5.7 of van der Vaart and Wellner [28] then implies that $\mathrm{H_P}$ is uniformly $\mathcal{L}^1(\mathrm{P})$-equicontinuous. So we can use $\mathcal{L}^1(\mathrm{P})$-seminorm in the definition of $\|\cdot\|_{\mathcal{F}_\delta}$.

The result (7) for finite-dimensional distributions follows from the Cramér-Wold device (Billingsley [5, Theorem 29.4]) and Theorem 2.1 in Radulović [25]. The latter concerns convergence of finite-dimensional

distributions for observations from a discrete Markov chain; using basic asymptotics of renewal/regenerative processes (Asmussen [1, Section V.6 and VI.3]), the proof also works for regenerative sequences.

We proceed to show stochastic $\mathcal{L}^1(\mathrm{P})$-equicontinuity of G_n^*. Define for $j = 1, \ldots, \tau_n$ stochastic processes

$$Z_j(f) := \sum_{i=T_j+1}^{T_{j+1}} f(X_i), \quad Z_j^*(f) := \sum_{i=T_j^*+1}^{T_{j+1}^*} f(X_i^*), \quad f \in \mathcal{F}.$$

Denote by γ the distribution of the bootstrapped observations obtained from Algorithm 2.1 and by E_γ expectation with respect to γ and take $\mu := \mathrm{E}\ell_1$. Define $a_n := [n/\mu]$. Then

$$\begin{aligned}
\|(n_*/a_n)^{1/2}\mathrm{G}_n^*\|_{\mathcal{F}_\delta} \quad &\leq \quad \Big\|a_n^{-1/2}\sum_{i=1}^{\tau_n}(Z_i^* - Z_i)\Big\|_{\mathcal{F}_\delta} + (\tau_n/a_n)^{1/2}\Big\|Y_n a_n^{-1}\sum_{i=1}^{\tau_n}(Z_i - \mu\mathrm{P})\Big\|_{\mathcal{F}_\delta} \\
&\quad + (\tau_n/a_n)^{3/2}\|Y_n\mu\mathrm{P}\|_{\mathcal{F}_\delta} + (n_*/\underline{n})\underline{a}_n^{-1/2}\sum_{i=T_0+1}^{T_1}|F(X_i)| \\
&=: \quad A(n,\delta) + (\tau_n/a_n)^{1/2}B(n,\delta) + (\tau_n/a_n)^{3/2}C(n,\delta) + (n_*/\underline{n})D(n),
\end{aligned}$$

where $Y_n := (a_n/\underline{n}) \times \tau_n^{-1/2}(\underline{n} - n_*)$. By Slutsky's lemma for the bootstrap (Radulović [25, Lemma 3.1]), it is enough to show convergence in probability as $n \to \infty$, $\delta \downarrow 0$ of $A(n,\delta)$, $B(n,\delta)$, $C(n,\delta)$, and $D(n)$ separately.

It is immediate that $D(n) \to 0$ almost surely. Concerning $C(n,\delta)$, define $\bar{\ell}_{\tau_n} = \tau_n^{-1}\sum_{i=1}^{\tau_n}\ell_i$. Then $n_* - \underline{n} = \sum_{i=1}^{\tau_n}(\ell_i^* - \bar{\ell}_{\tau_n})$ is of order $O_{\mathrm{Q}}(\sqrt{n})$ as $n \to \infty$. This follows since the ℓ_i^*s are conditionally IID, so that by Markov's inequality

$$\gamma\Big(\sum_{i=1}^{\tau_n}\ell_i^* > \sqrt{n}M\Big) \leq \tau_n\gamma(\ell_1^* > \sqrt{n}M) \leq M^{-2}n^{-1}\sum_{i=1}^{\tau_n}\ell_i^2 \to 0, \qquad n, M \to \infty$$

almost surely, by the Law of Large Numbers. Slutsky's lemma for bootstrapped processes (Radulović [25, Lemma 3.1]) then implies $Y_n = O_{\mathrm{Q}}(1)$. Recalling our choice of semimetric in the definition of $\|\cdot\|_{\mathcal{F}_\delta}$, we obtain $C(n,\delta) \leq |Y_n|\mu\delta$ which converges to zero in probability as $n \to \infty$, $\delta \downarrow 0$.

Convergence of $B(n,\delta)$ to zero in probability follows from Slutsky's lemma and arguments as in the proof of Lemma 4.6 of Tsai [27]. Since $Y_n = O_{\mathrm{Q}}(1)$, we have $\lim_{\delta\downarrow 0}\lim_n \|B(n,\delta)\|_{\mathcal{F}_\delta} = 0$ in probability.

Finally, regarding $A(n,\delta)$, fix $\varepsilon > 0$, $\delta > 0$. By Markov's inequality

$$\begin{aligned}
\gamma\Big(\Big\|a_n^{-1/2}\sum_{i=1}^{\tau_n}(Z_i^* - Z_i)\Big\|_{\mathcal{F}_\delta} > \varepsilon\Big) &\leq \gamma\Big(\Big\{\Big\|a_n^{-1/2}\sum_{i=1}^{\tau_n}(Z_i^* - Z_i)\Big\|_{\mathcal{F}_\delta} > \varepsilon\Big\} \cap \{|\tau_n - a_n| \leq a_n\}\Big) + 1_{\{|\tau_n - a_n| > a_n\}} \\
&\leq \varepsilon^{-1}\mathrm{E}_\gamma\Big(\Big\|a_n^{-1/2}\sum_{i=1}^{\tau_n}(Z_i^* - Z_i)\Big\|_{\mathcal{F}_\delta} 1_{\{|\tau_n - a_n| \leq a_n\}}\Big) + o_{\mathrm{Q}}(1).
\end{aligned}$$

To bound the last expectation, we use Poissonization. Let $\{N_n\}$ be a sequence of IID symmetrised Poisson random variables with parameter $1/2$ independent of X, T, defined on the same probability space. To simplify notation, we implicitly assume all of the calculations in the following to be conditionally on $|\tau_n - a_n| \leq a_n$. By Lemma 3.6.6 of van der Vaart and Wellner [28],

$$\mathrm{E}_\gamma\Big\|a_n^{-1/2}\sum_{i=1}^{\tau_n}(Z_i^* - Z_i)\Big\|_{\mathcal{F}_\delta} \leq 4\mathrm{E}_N\Big\|a_n^{-1/2}\sum_{i=1}^{\tau_n}N_i Z_i\Big\|_{\mathcal{F}_\delta}.$$

Since $\mathrm{E}\|W_1\|_{\mathcal{F}_\delta} \leq \mathrm{E}\|W_1 + W_2\|_{\mathcal{F}_\delta}$ for centered, independent processes W_1, W_2 by Jensen's inequality,

$$\mathrm{E}_N\Big\|a_n^{-1/2}\sum_{i=1}^{\tau_n}N_i Z_i\Big\|_{\mathcal{F}_\delta} \leq \mathrm{E}_N\Big\|a_n^{-1/2}\sum_{i=1}^{a_n}N_i Z_i\Big\|_{\mathcal{F}_\delta} + \mathrm{E}_N\Big\|a_n^{-1/2}\sum_{i=a_n+1}^{\tau_n}N_i Z_i\Big\|_{\mathcal{F}_\delta} \leq 2\mathrm{E}_N\Big\|a_n^{-1/2}\sum_{i=1}^{a_n}N_i Z_i\Big\|_{\mathcal{F}_\delta}.$$

Taking expectations E_X with respect to X, T everywhere, conclude that for some universal constant C

$$\mathrm{E}_X\gamma\Big(\Big\|a_n^{-1/2}\sum_{i=1}^{\tau_n}(Z_i^* - Z_i)\Big\|_{\mathcal{F}_\delta} > \varepsilon\Big) \leq C\varepsilon^{-1}\mathrm{E}\Big\|a_n^{-1/2}\sum_{i=1}^{a_n}N_i Z_i\Big\|_{\mathcal{F}_\delta}.$$

The multiplier inequality argument in the proof of Theorem 3.6.3 of van der Vaart and Wellner [28] implies convergence to zero of the right hand side of the display as $n \to \infty, \delta \downarrow 0$. This proves stochastic equicontinuity in probability of $A(n,\delta)$ and so G_n^* is stochastically equicontinuous in probability (Q). Combining this with convergence of finite-dimensional distributions, we obtain the desired result. □

Acknowledgment. The research is supported by the Danish Natural Science Research Council, grant 272-06-0442, "Point Process Modelling and Statistical Inference". We would like to thank Professor Avishai Mandelbaum, Industrial Engineering and Management, Technion, Israel for providing access to the call center data.

References

[1] S. Asmussen, *Applied probability and queues*, 2 ed., Springer, 2003.

[2] K. B. Athreya and C. D. Fuh, *Bootstrapping Markov chain: countable case*, J. Statist. Plann. Inference **33** (1992), 311–331.

[3] P. Bertail and S. Clémençon, *Regenerative block-bootstrap for Markov chains*, Bernoulli **12** (2006), 689–712.

[4] P. Bertail and S. Clémençon, *Second order properties of regeneration-based bootstrap for Markov chains*, Test **16** (2007), 109–122.

[5] P. Billingsley, *Probability and measure*, 3 ed., Wiley, 1995.

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

[7] _____, *Nonparametric inference from the $M/G/1$ busy periods*, Comm. Statist. Stochastic Models **15** (1999), 247–272.

[8] L. Brown, N. Gans, A. Mandelbaum, A. Sakov, H. Shen, S. Zeltyn, and L. Zhao, *Statistical analysis of a telephone call center: A queuing science perspective*, J. Amer. Statist. Assoc. **100** (2005), 36–50.

[9] J. Cai and J. Kim, *Nonparametric quantile estimation with correlated failure time data*, Lifetime Data Anal. **9** (2003), 357–371.

[10] Z. Cai, *Kernel density estimation and hazard rate estimation for censored dependent data*, J. Multivariate Anal. **67** (1998), 23–34.

[11] _____, *Estimating a distribution function for censored time series data*, J. Multivariate Anal. **78** (2001), 299–318.

[12] S. Datta and W. P. McCormick, *Regeneration-based bootstrap for Markov chains*, Canad. J. Statist. **21** (1993), 181–193.

[13] B. Efron, *Bootstrap methods: Another look at the jackknife*, Ann. Statist. **7** (1979), 1–26.

[14] N. Gans, G. Koole, and A. Mandelbaum, *Commissioned paper: Telephone call centers: Tutorial, review and research prospects*, Manuf. Serv. Oper. Manag. **5** (2003), 79–141.

[15] O. Garnett, E. Mandelbaum, and M. Reiman, *Designing a call center with impatient customers*, Manuf. Serv. Oper. Manag. **4** (2002), 208–228.

[16] R. Gill and S. Johansen, *A survey of product-integration with a view towards application in survival analysis*, Ann. Statist. **18** (1990), 1501–1555.

[17] E. Giné and J. Zinn, *Bootstrapping general empirical measures*, Ann. Probab. **18** (1990), 851–869.

[18] P. Hall, *The bootstrap and Edgeworth expansion*, Springer, 1992.

[19] M. B. Hansen and S. M. Pitts, *Nonparametric inference from the $M/G/1$ workload*, Bernoulli **12** (2006), 737–759.

[20] U. V. Naik-Nimbalkar and M. B. Rajarshi, *Validity of blockwise bootstrap for empirical processes with stationary observations*, Ann. Statist. **22** (1994), 980–994.

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

[22] A. Nederlof and J. Anton, *Customer obsession: Your roadmap to profitable CRM*, Anton Press, 2002.

[23] D. Pollard, *A central limit theorem for empirical processes*, J. Austr. Math. Soc. Ser. A **33** (1982), 235–248.

[24] _____, *Convergence of stochastic processes*, Springer, 1984.

[25] D. Radulović, *Renewal type bootstrap for Markov chains*, Test **13** (2004), 147–192.

[26] H. Thorrison, *Coupling, stationarity and regeneration*, Springer, 2000.

[27] T. H. Tsai, *The uniform CLT and LIL for Markov chains*, Ph.D. thesis, University of Wisconsin, 1998.

[28] A. W. van der Vaart and J. A. Wellner, *Weak convergence and empirical processes*, Springer, 1996.

[29] M. Wolf, D. N. Politis, and J. P. Romano, *Weak convergence of dependent empirical measures with application to subsampling in function spaces*, J. Statist. Plann. Inference **79** (1999), 179–191.