Statistica Sinica **25** (2015), 1313-1335
doi:http://dx.doi.org/10.5705/ss.2014.092

NONPARAMETRIC ESTIMATION OF COMPONENT RELIABILITY BASED ON LIFETIME DATA FROM SYSTEMS OF VARYING DESIGN

Peter Hall[1,2], Yin Jin[2] and Francisco J. Samaniego[2]

[1] *University of Melbourne and* [2] *University of California, Davis*

Abstract: Failure time data collected from fielded systems provide indirect information about the performance of the system's components. Since it is often difficult to create or simulate field conditions in laboratory settings, the process of drawing inferences about component characteristics from data on system performance is of practical importance. However, there is very little literature on this problem that treats such inferences from a nonparametric perspective, and less literature still that allows the systems of interest to be of arbitrary design. The present paper focuses on nonparametric estimation of a common component reliability function using independent samples from coherent systems of varying design whose components have independent, identically distributed lifetimes. Two estimation approaches are studied. The first is conventional, and is based on treating each of the estimation problems separately; it is shown that these mixture estimators are consistent, and their asymptotic behaviours are characterized. The second estimator is quite unconventional. It is obtained by solving multiple point-wise maximum likelihood estimation problems simultaneously, and combining the separate estimators, each at fixed time points, to obtain an overall estimator of the reliability function. We show that the latter approach produces a legitimate reliability function and that, asymptotically, it is uniformly superior to all the estimators of the first type. Related estimators of the lifetime density and failure rate functions are also obtained, and their theoretical and numerical properties are described.

Key words and phrases: Coherent system, component reliability, density estimation, failure rate estimation, inverse problem, local likelihood, maximum likelihood, nonparametric estimation, reliability polynomial, survival function, system design.

1. Introduction

1.1. Motivation and context

Estimating the reliability of the components in an engineered system is important in engineering practice. The data obtained on the performance of fielded systems is the most relevant source of data for such estimation, as laboratory experiments on individual components often fail to simulate the environment in which the components are actually used. The problem treated here has the essential features of an inverse problem, as the goal is to invert the functional relationship between system and component reliabilities in order to gain insight into the behaviour of the components. Solutions to such problems have appeared in the reliability literature, but the approach we take to these problems differs from previous work in important respects.

First, unlike most previous work, we take a fully nonparametric approach to the problem. Further, while earlier work tends to focus on special system designs (e.g., series or parallel systems), our approach allows for the systems from which lifetime data are obtained to be of arbitrary design, subject only to the standard assumption that the systems be coherent, that is, that a system's performance be monotone in its components' performance and that every component be relevant. In the estimation problems considered here, we make the realistic assumption that only system lifetime data are available, that is, there is no auxiliary information on component performance. Finally, we treat data from multiple systems of varying design.

Our core assumptions are as follows. It is assumed that a random sample of lifetimes is available from each of $m \geq 2$ systems having possibly different designs, and that the m samples are mutually independent. It is also assumed that the components of each system have lifetimes that are independent, with a common continuous distribution function $F(x) = P(X \leq x)$ with survival function $\bar{F}(x) = 1 - F(x)$. Under these assumptions we consider several problems: estimation of $\bar{F}$ itself and, under the assumed absolute continuity of $\bar{F}$, estimation of the corresponding density f and of the failure rate $r = f/\bar{F}$.

Among the assumptions posited in the preceding paragraph, the assumption that the lifetimes of the components of all m systems are independent and identically distributed deserves further comment. Samaniego (2007) discusses a variety of systems to which the assumption applies; examples include wafers or chips in digital computers, and batteries in flashlights. Large scale applications include the case of computer hardware that is generally referred to by the term Raid (Patterson, Gibson, and Katz, 1988). The well documented efficiency and speed of Raid computers has led to their widespread use. A Raid computer with n independent disks can be designed to perform as a k-out-of-n system, which fails upon the kth disk failure. With the aid of a randomization device, Raid computers can be used to simulate the performance of an arbitrary coherent system in n components with independent, identically distributed lifetimes.

1.2. Problem description

A special case of the problem of primary interest here was treated by Bhattacharya and Samaniego (2010). For a coherent system comprised of n components having independent and identically distributed lifetimes with survival function $\bar{F}$, it is known that the relationship between the system's reliability function $\bar{F}_T$ and the reliability function $\bar{F}$ of its components can be written as

$$\bar{F}_T\,(t) = \sum_{j=1}^{n} d_j\ \left\{\bar{F}(t)\right\}^j\ , \tag{1.1}$$

where the vector of coefficients $d = (d_1, \ldots, d_n)$ is generally referred to as the domination vector, following Satyanarayana and Prabhakar (1978). The domination vector d is a topological invariant that depends solely on a system's design and does not depend on the underlying component distribution F. For fixed $t \geq 0$ we may write $p = \bar{F}(t)$ for the probability that a given component is working at time t, and we may represent the probability $h(p)$ that the system is working at time t as

$$h(p) = \sum_{j=1}^{n} d_j\, p^j\ . \tag{1.2}$$

The function h in (1.2) is referred to as the system's reliability polynomial. It is well known that h is a continuous, strictly increasing function of $p \in [0, 1]$, with $h(0) = 0$ and $h(1) = 1$.

Assume that we have a random sample consisting of the failure times of N systems of identical design whose components have lifetimes that are independent, with a common reliability function $\bar{F}$. Now, $\bar{F}_T(t) = h\left\{\bar{F}(t)\right\}$ for all $t \geq 0$, and we have that the standard, asymptotically optimal estimator of $\bar{F}_T$, namely, the empirical survival function, which is also the nonparametric maximum likelihood estimator of $\bar{F}$ and is given by

$$\hat{\bar{F}}_{T,N}(t) = \frac{1}{N} \sum_{j=1}^{N} I_{(t,\infty)}(t_j)\,, \tag{1.3}$$

where

$$I_A(t) = \begin{cases} 1\,, t \in A, \\ 0\,, t \notin A\,. \end{cases} \tag{1.4}$$

It follows from the continuity and invertibility of the function h that the nonparametric maximum likelihood estimator of the component reliability function $\bar{F}(t)$ may be obtained as

$$\hat{\bar{F}}_N(t) = h^{-1}\ \left\{\hat{\bar{F}}_{T_N}(t)\right\}\ . \tag{1.5}$$

Bhattacharya and Samaniego (2010) identify the asymptotic behaviour of $\hat{\bar{F}}_N(t)$ as

$$N^{1/2}\left\{\hat{\bar{F}}_N(t) - \bar{F}(t)\right\} \to V\ , \tag{1.6}$$

where the convergence is in distribution and the random variable V has a normal distribution with zero mean and variance given by $(\sum id_i[h^{-1}\{\bar{F}_T(t)\}]^{i-1})^{-2}F(t)$ $\bar{F}(t)$, where d and h are specified in (1.1) and (1.2).

In the present work we address the following more general problem. Suppose that m coherent systems of arbitrary design are fielded, where each system is based on components whose lifetimes are independent and a have common distribution F, and suppose that a random sample of lifetimes is drawn from each system. Specifically, for $i = 1, \ldots, m$, let $T_{i1}, T_{i2}, \ldots, T_{iN_i}$ be independent and identically distributed random variables with distribution F_{T_i}. If the ith system has n_i components and domination vector d_i, then the reliability polynomials of the m systems are

$$h_i(p) = \sum_{j=1}^{n_i} d_{ij} p^j, \quad i = 1, \ldots, m. \tag{1.7}$$

In this circumstance, the random sample of lifetimes from each individual system provides an avenue for consistent estimation of $\bar{F}$. Specifically, letting $\hat{\bar{F}}_{T_i,N_i}(t)$ denote the empirical survival function based on lifetime data from the ith system, we have that, for $i = 1, \ldots, m$, each of the estimators

$$\hat{\bar{F}}_{N_i}(t) = h_i^{-1} \left\{ \hat{\bar{F}}_{T_i,N_i}(t) \right\} \tag{1.8}$$

is a consistent, asymptotically normal estimator of $\bar{F}(t)$. Indeed, every convex combination of the estimators in (1.8) is, similarly, a consistent, asymptotically normal estimator of $\bar{F}(t)$. While this is an interesting and viable class of estimators of $\bar{F}(t)$, the class is not expected to contain estimators that achieve the smallest possible asymptotic variance. We therefore extend our investigation by taking a likelihood approach to the estimation of $\bar{F}(t)$. In doing so, we shall use a natural estimator in the class of mixtures of the estimators in (1.8) as an initial value in Newton-Raphson iterations aimed at maximizing local likelihoods. Specifically, the estimator

$$\hat{\bar{F}}_{[1]}(t) = \frac{1}{N} \sum_{i=1}^{m} N_i \, \hat{\bar{F}}_{N_i}(t) \,, \tag{1.9}$$

where $N = \sum_i N_i$, is utilized as our initial estimator of $\bar{F}$.

For arbitrary $m \geq 2$ the problem of finding the nonparametric maximum likelihood estimator of $\bar{F}$ has, in general, proven to be intractable. The likelihood-based approach we take here is unconventional, proceeding locally rather than globally. We begin by noting that, for fixed t, the estimation of the parameter $p(t) = \bar{F}(t)$ is a regular parametric problem for which the maximum likelihood estimator of p is consistent and asymptotically normal with minimum variance. While the maximum likelihood estimator $\hat{p}(t)$ of $p(t)$ will not generally be available in closed form, it can be approximated reliably through Newton-Raphson iterations using an initial value $\hat{\bar{F}}_{[1]}(t)$. It is easily shown that combining this collection of estimated $p(t)$s yields a legitimate reliability function $\hat{\bar{F}}_{[2]}(t)$ that estimates the true component reliability function $\bar{F}(t)$ for all $t \geq 0$.

In Theorem 1, we obtain the asymptotic distribution of the estimator $\hat{\bar{F}}_{[2]}(t)$, assuming that, for $i = 1, \ldots, m$, $N_i \to \infty$ in such a way that $N_i/N \to b_i > 0$, where $b = (b_1, \ldots, b_m)$ is an m-dimensional probability vector. Further, we investigate the asymptotic behaviour of the class of general mixtures of the individual estimators in (1.8), putting us in a position to prove the asymptotic domination of the likelihood-based estimator $\hat{\bar{F}}_{[2]}(t)$ over this natural class of competitors.

Many applications of reliability theory involve the use of parametric models for the lifetimes of components of systems. In such applications, the estimation problems that arise tend to be focused on estimating the probability density function f of the components under study. This of course is usually accomplished through estimation of the finite-dimensional parameter that defines the model of interest. This work motivates us to extend our investigation to the problem of density estimation. Our results in this area will enable a reliability practitioner to obtain an estimate of the shape of the true density function before a decision is made about whatever parametric model might be under consideration. Our investigation leads to results which state explicit conditions under which our nonparametric estimator of $f = F'$ converges to f at an optimal rate. Next, we focus on the estimation of the failure rate $r = f/\bar{F}$. Our final theorem gives the rate of convergence and asymptotic distribution of our proposed estimator of r.

It is worth noting that the results obtained here apply as well when sampling lifetimes from mixed systems (i.e., stochastic mixtures of coherent systems) rather than from coherent systems. This follows from the fact that mixed systems are fully characterized by their reliability polynomials and thus admit the same treatment given here to data from coherent systems. More details on mixed systems may be found in Samaniego (2007). Further, a comment on the iid assumption made on the component lifetimes of all the systems treated here will help to clarify the intended applicability of this work. As noted by Navarro, Samaniego, and Balakrishnan (2011), this assumption ensures the identifiability of the component reliability function F, while even in the quite direct extension to the assumption of independent but not identically distributed (inid) component lifetimes, component lifetimes are not identifiable parameters of the distribution of system lifetimes.

1.3. Literature review

A number of authors have studied the estimation of a component lifetime distribution F from system failure times in the presence of additional information. An early example is Moeschberger and David (1971), who treat the estimation problem in a competing risks framework. Meilijson (1981) and Bueno (1988) consider the estimation of F based on system failure times together with autopsy statistics on the systems' components. System failure times have also been employed in the estimation of F in the sampling scenario usually referred to as "masked data". In this version of the problem, it is typically assumed that one has access to data on system failure times, possibly accompanied by partial information on the component failures that were responsible for the failure of the system. Authors who have studied the estimation of component characteristics from masked data include Miyakawa (1984), Usher and Hodgson (1988) and Guess, Usher, and Hodgson (1991).

Estimation of the component lifetime distribution from system failure time data also arises in other contexts. Boyles and Samaniego (1986) derived the nonparametric maximum likelihood estimator of the underlying component distribution based on nomination sampling, that is, sampled maxima from independent samples; such a collection may be thought of as a sample of failure times from independent parallel systems of varying sizes. See also Boyles and Samaniego (1987). Inference about the underlying distribution F based on ranked set sampling, a sampling approach equivalent to the observation of independently drawn order statistics, has been treated by Stokes and Sager (1988) and by Kvam and Samaniego (1993a,b, 1994). A ranked set sample may be viewed as a set of independent lifetimes from k-out-of-n systems with varying k and n.

Much of the work cited above makes parametric assumptions about the component reliability $\bar{F}$. The assumption of exponentiality is the most prevalent, although other models, notably the Weibull distribution, have also been studied. None of the work cited in the preceding two paragraphs treats the case of general systems. There is, of course, a substantial literature on nonparametric inference in reliability. Recent relevant papers from this literature include Balakrishnan, Ng, and Navarro (2011a), Balakrishnan, Ng, and Navarro (2011b), Chahkandi, Ahmadi, and Baratpour (2014) and Eryilmaz (2011).

The present paper proceeds as follows. Estimation of the reliability function $\bar{F}$, its density f and its failure rate r is treated in Section 2. Theoretical results on the asymptotic behaviour of our estimators are stated in Section 2, with most of the formal proofs relegated to the Appendix. In Section 3 we illustrate the performance of the proposed estimators on simulated data and discuss simulation results aimed at approximating each estimator's integrated squared error. Our estimators of $\bar{F}$, f and r are shown to perform well even for samples of moderate size.

2. Methodology and its Basic Properties

2.1. Estimation of the lifetime distribution

We use notation introduced in Section 1, writing m for the number of systems with different designs and assuming that a random sample of lifetime data has been recorded for each system. All m systems are assumed to operate solely with components whose lifetimes are independent with common distribution F.

The components' reliability function $1 - F$ is denoted by $\bar{F}$. The available data consist of a random sample of N_i system lifetimes $T_{i1}, \ldots, T_{iN_i}$ for $i = 1, \ldots, m$.

To estimate $\bar{F}$ we begin with a collection of indicator functions which track whether or not a given system survives beyond time t. Let h_i, defined in (1.2), be the reliability polynomial of system i, and let

$$X_i(t) = \sum_{j=1}^{N_i} I_{(t,\infty)}(T_{ij}), \tag{2.1}$$

where the indicator function I is as defined in (1.4). Since $X_i(t)$ has the Binomial $\text{Bin}\{N_i, h_i(p)\}$ distribution, where $p = p(t) = \bar{F}(t)$, we can write the likelihood for data from the ith system as

$$L_i\{p, X_i(t)\} = \binom{N_i}{X_i(t)} \{h_i(p)\}^{X_i(t)} \{1 - h_i(p)\}^{N_i - X_i(t)}.$$

Here and below we typically suppress the argument of $p(t)$. The likelihood L of the entire dataset is of course the product of the individual likelihoods above:

$$L(p,t) = \prod_{i=1}^{m} L_i\{p, X_i(t)\} = \prod_{i=1}^{m} \binom{N_i}{X_i(t)} \{h_i(p)\}^{X_i(t)} \{1 - h_i(p)\}^{N_i - X_i(t)}. \tag{2.2}$$

The maximum likelihood estimator of $p(t) = \bar{F}(t)$, denoted by $\hat{p}(t)$, is obtained by maximizing $L(p, X(t))$ with respect to p, typically by numerical means such as the Newton-Raphson algorithm. We initialize the Newton-Raphson algorithm using the estimator $\hat{\bar{F}}_{[1]}$ in (1.9), which in turn is based on the system-specific estimators $\hat{\bar{F}}_{N_i}$, defined in (1.8) for $1 \le i \le m$. If the estimators $\hat{p}(t)$ are combined, for all $t > 0$, the result is a step function, to be designated as $\hat{\bar{F}}_{[2]}(t)$, with at most $N = \sum_{1 \le i \le m} N_i$ jump points.

We can view $\hat{\bar{F}}_{[1]}$ as a member of a larger class of estimators having the form

$$\hat{\bar{F}}_{\mathbf{k}}(t) = \sum_{i=1}^{m} k_i \, \hat{\bar{F}}_{N_i}(t), \tag{2.3}$$

where $\mathbf{k} = (k_1, \ldots, k_m)$, which in practice would depend on $N_1, \ldots, N_m$ as in equation (1.9), is a vector of nonnegative components that have the property $\sum_i k_i = 1$, and $\hat{\bar{F}}_{N_i}$ is as in (1.8).

Theorem 1 below is proved in the Appendix, and asserts that the estimator $\hat{\bar{F}}_{[2]}$ is a proper survival function, that $\hat{\bar{F}}_{[2]}$ is strongly consistent for $\bar{F}$, and that $\hat{\bar{F}}_{[2]}$ is superior to $\hat{\bar{F}}_{[1]}$ in terms of asymptotic statistical performance. Indeed, $\hat{\bar{F}}_{[2]}$ is superior to any estimator $\hat{\bar{F}}_{\mathbf{k}}$ having the form in (2.3).

In establishing this property of $\hat{\bar{F}}_{\mathbf{k}}$ we assume that

$$\begin{array}{l}(2.4.1)\ k_1,\dots,k_m \text{ are nonnegative functions of } \{N_i\} \text{ satisfying } \\ \sum_i k_i = 1, \\ (2.4.2)\ N_1,\dots,N_m \text{ all diverge as } N\to\infty.\end{array} \tag{2.4}$$

We make the following assumption throughout our theoretical work:

$$\begin{array}{l}(2.5.1) \text{ each } h_i \text{ is a reliability polynomial of a coherent system,} \\ (2.5.2) \text{ the component lifetime distribution is nondegenerate.}\end{array} \tag{2.5}$$

Our reference to consistency, and other asymptotic properties, in Theorem 1 and in subsequent theoretical results, pertains to properties as $N=\sum_i N_i$ diverges, and in particular, we consider each N_i to be a function of N.

Theorem 1. *Assume that* (2.5) *holds. Then:* (a) $\hat{\bar{F}}_{[2]}(t)$ *is a non-increasing, right-continuous function which satisfies the conditions* $\hat{\bar{F}}_{[2]}(0)=1$ *and* $\hat{\bar{F}}_{[2]}(\infty)=0$*;* (b) *with probability 1,* $\hat{\bar{F}}_{[2]}(t)$ *converges to* $\bar{F}(t)$*, uniformly in* $t\in[0,\infty]$*; and* (c) *if* (2.4) *holds then* $\hat{\bar{F}}_{[2]}$ *is asymptotically superior to any estimator* $\hat{\bar{F}}_{\mathbf{k}}$ *having the form in* (2.3)*, and in particular to* $\hat{\bar{F}}_{[1]}$*, as an estimator of* $\bar{F}$*, in the sense that the asymptotic variance of* $\hat{\bar{F}}_{[2]}$ *does not exceed that of* $\hat{\bar{F}}_{[1]}$*. The estimators* $\hat{\bar{F}}_{[1]}$ *and* $\hat{\bar{F}}_{[2]}$ *of* $\bar{F}$ *are asymptotically equivalent at* t *if and only if each* $h_i'\{\bar{F}(t)\}\neq 0$ *and*

$$k_i = C\,\frac{N_i}{N}\,\frac{h_i'\{\bar{F}(t)\}^2}{h_i\{\bar{F}(t)\}\,[1-h_i\{\bar{F}(t)\}]} \tag{2.6}$$

for $i=1,\dots,m$*, where* $C=C(t)>0$ *is chosen so that* $\sum_i k_i=1$*.*

Since the choice of $\{k_i\}$ given by (2.6) involves the unknown function $\bar{F}$, there is no practical way of obtaining a mixture estimator with such coefficients.

We shall show in the Appendix that $\hat{\bar{F}}_{[2]}$ is asymptotically normally distributed, and we note that the precursor $\hat{\bar{F}}_{[1]}$ also has that property. In that work, and in Theorems 1 and 2 in the present section, it is assumed that m, the number of systems, is kept fixed as N diverges.

2.2. Estimation of the lifetime density and failure rate

Our estimator of the lifetime distribution $F=1-\bar{F}$ is of course $\hat{F}=1-\hat{\bar{F}}$. We can estimate the corresponding density, $f=F'$, by passing a kernel smoother through $\hat{F}$:

$$\hat{f}(x)=\frac{1}{w}\int K\left(\frac{x-y}{w}\right)d\hat{F}(y)\,, \tag{2.7}$$

where $w>0$ is a bandwidth and K is a kernel function.

An appropriate bandwidth can be determined empirically, as follows. Fit a plausible model to the data, for example a Weibull or even an exponential model, using a conventional parametric approach such as maximum likelihood or even the method of moments. Then, by numerical simulation from that model, determine the bandwidth w that would minimize integrated squared error,

$$\mathrm{ISE}(w) = \int \big\{\hat{f}(x\,|\,w) - f(x)\big\}^2\,dx\,, \tag{2.8}$$

if the model were correct. Here, to express the fact that $\hat{f}$ is a functional of w, we have denoted it by $\hat{f}(\,\cdot\,|\,w)$.

In our numerical work, which involves simulating from non-exponential Weibull models, we shall use the exponential distribution as the model for bandwidth choice, to emphasize that in practice the model used does not have to be a particularly good fit to the data. The methodology that we are employing here is a generalization, to the present problem, of a technique that is already widely used for density estimation in much simpler settings than ours. There, the fitted model is referred to as the reference distribution, and it is typically taken to be normal with its mean and variance equal to those of the data; see e.g. Silverman (1986, pp. 45ff, 86f).

The main properties of $\hat{f}$ will be given in Theorem 2, below. Further, we note that, once we have estimators $\hat{f}$ of f and $\hat{\bar{F}}_{[2]}$ of $\bar{F}$, we can construct an estimator $\hat{r} = \hat{f}/\hat{\bar{F}}_{[2]}$ of the failure rate, $r = f/\bar{F}$. We shall show in the Appendix that $\hat{\bar{F}}_{[2]}$ converges to $\bar{F}$ at rate $N^{-1/2}$, and it will follow from Theorem 2 that, even if the bandwidth is chosen optimally, $\hat{f}$ converges to f at a strictly slower rate than that at which $\hat{\bar{F}}_{[2]}$ converges to $\bar{F}$. In fact, the convergence rate for $\hat{f}$ is $w^2 + (Nw)^{-1/2}$, and since $w \to 0$ as $N \to \infty$ then $w^2 + (Nw)^{-1/2}$ converges to zero more slowly than $N^{-1/2}$. Therefore the convergence rate of $\hat{r}$ to r is really that of $\hat{f}$ to f.

Indeed, it will follow from Theorem 2 that, under the assumptions there,

$$\hat{r}(t) = r(t) + \frac{\hat{f}(t) - f(t)}{\bar{F}(t)} + o_p\big\{w^2 + (Nw)^{-1/2}\big\}\,. \tag{2.9}$$

A central limit theorem for $\hat{r}$ is readily proved from (2.9) and the central limit theorem for $\hat{f}$ given in (2.13): $\hat{r} - r$ is asymptotically normally distributed with bias of the form $\kappa_2\, w^2\, f''/(2\,\bar{F})$ and variance of the form $(Nw)^{-1}\,\tau_N^2/\bar{F}^2$, where $\kappa_2 = \int u^2\, K(u)\, du$ and

$$\tau_N(t)^2 = \kappa\, f(t) \sum_{i=1}^{m} \frac{N_i}{N}\, \frac{a_i\{F(t)\}^2}{c_N\{F(t)\}^2}\, h_i'\{F(t)\}\,. \tag{2.10}$$

As a prelude to stating Theorem 2 we note the following assumptions on f, K and w:

$$\begin{aligned}&\text{(2.11.1) } K \text{ is a bounded, symmetric, compactly supported probability density, and has a bounded derivative on the real line,}\\ &\text{(2.11.2) } w = w(N) \to 0 \text{ as } N \to \infty, \text{ at a rate that is sufficiently slow to ensure that } (Nw)^{-1} (\log N)^2 \to 0 \text{ as } N \to \infty,\\ &\text{(2.11.3) } f \text{ has first two bounded, continuous derivatives in a neighbourhood of } t, \text{ and } f(t) > 0.\end{aligned} \tag{2.11}$$

If (2.11.1) holds then the quantities $\kappa = \int K(u)^2 \, du$ and κ_2 are finite and positive. Condition (2.11.1) encompasses most kernels that are used in practice, except for the standard normal density, which can be addressed by making minor modifications to our proof. Likewise, (2.11.2) is conventional, since it covers all cases where $w \to 0$ and $Nw \to \infty$ at polynomial rates, and (2.11.3) is the standard second order assumption used in theory for density estimation. If we ask only that the first two derivatives of f be bounded, without the requirement of continuity asserted in (2.11.3), then Theorem 2 continues to hold provided we replace the relatively concise term $(1/2) \kappa_2 w^2 f''(t)$ in (2.13) by simply $O(w^2)$; then we may of course drop the $o_p(w^2)$ term.

Define

$$a_i(p) = \frac{h_i'(p)}{h_i(p) \{1 - h_i(p)\}} , \quad c_N(p) = \sum_{i=1}^{m} \frac{N_i}{N} \frac{h_i'(p)^2}{h_i(p) \{1 - h_i(p)\}} . \tag{2.12}$$

If (2.5) and (2.11) hold, then $\tau_N(t)$ in (2.10) is bounded away from zero and infinity as $N \to \infty$.

Theorem 2. *Let t be such that $0 < F(t) < 1$, and assume that* (2.5) *and* (2.11) *hold. Then,*

$$\hat{f}(t) - f(t) = \frac{1}{2} \kappa_2 w^2 f''(t) + (Nw)^{-1/2} \tau_N(t) Z_N(t) + o_p\big(w^2\big) , \tag{2.13}$$

where, for each t, the random variable $Z_N(t)$, is asymptotically distributed as normal $\mathrm{N}(0,1)$.

The $(1/2)\kappa_2 w^2 f''(t)$ term on the right-hand side of (2.13) is the conventional asymptotic bias for a kernel estimator of f. However, while the error-about-the-mean term in (2.13), i.e. $(Nw)^{-1/2} \tau_N(t) Z_N$, is of the same order of magnitude as its counterpart for a standard kernel estimator with the same sample size and bandwidth, it generally has larger variance in the case of (2.13).

An immediate corollary of (2.13) is that $\hat{f}$ has asymptotic mean squared error of order $(Nw)^{-1} + w^4$, which is minimized by taking w of size $N^{-1/5}$. That in turn implies a mean square convergence rate of $O(N^{-4/5})$. This rate

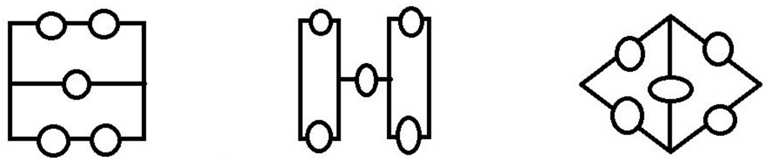

Figure 1. Three five-component systems.

of convergence is optimal for estimating densities with two derivatives. Standard methods for establishing the performance of bandwidth selectors could be adapted to prove that, in this problem, the bandwidth suggested earlier in this section asymptotically minimizes the order of magnitude of ISE(w), in (2.8).

A proof of Theorem 2 is deferred to the Appendix.

Theorem 3. *Under the asumption of Theorem* 2*,*

$$\hat{r}(t)-r(t) = \frac{1}{2}\,\kappa_2\,w^2\,\frac{f''(t)}{\bar{F}(t)}+(Nw)^{-1/2}\,\frac{\tau_N(t)\,Z_N(t)}{\bar{F}(t)}+o_p\{w^2+(Nw)^{-1/2}\}\,, \quad (2.14)$$

where the random variable $Z_N(t)$ is asymptotically distributed as normal $\mathrm{N}(0,1)$*.*

3. Discussion and Numerical Results

The purpose of the present section is to address two practical questions: (i) What can be said about the performance characteristics of the proposed estimators when the available sample sizes are small to moderate? (ii) What can be said about the global, rather than pointwise, performance of the proposed estimators as the finite sample size is moderately but steadily increased? We have examined these questions via simulation, and provide a brief commentary on our findings.

We shall be especially brief about the results on typical fits of our estimators in small to moderate samples. Figure 2 below show typical results obtained from fitting the estimator $\hat{\bar{F}}$ to $\bar{F}$, based on simulated samples of size 30 from $m = 3$ systems, all assumed to have components with iid lifetimes and a common distribution F. Schematic diagrams of these three systems from which lifetime data were drawn are shown in Figure 1.

For each of the systems above, a sample of thirty iid system lifetimes were generated from four commonly used lifetime distributions. In our simulations the parametric models used as component distributions F were: an exponential, an IFR Weibull, a DFR gamma and a lognormal distribution. Typical fits of the estimator to the true distribution are displayed in the panels of Figure 2. The specific models employed in the runs displayed here are Exp (1), Weibull $(2, 1)$, Gamma $(0.5, 2)$ and Lognormal $(0.5, 1)$.

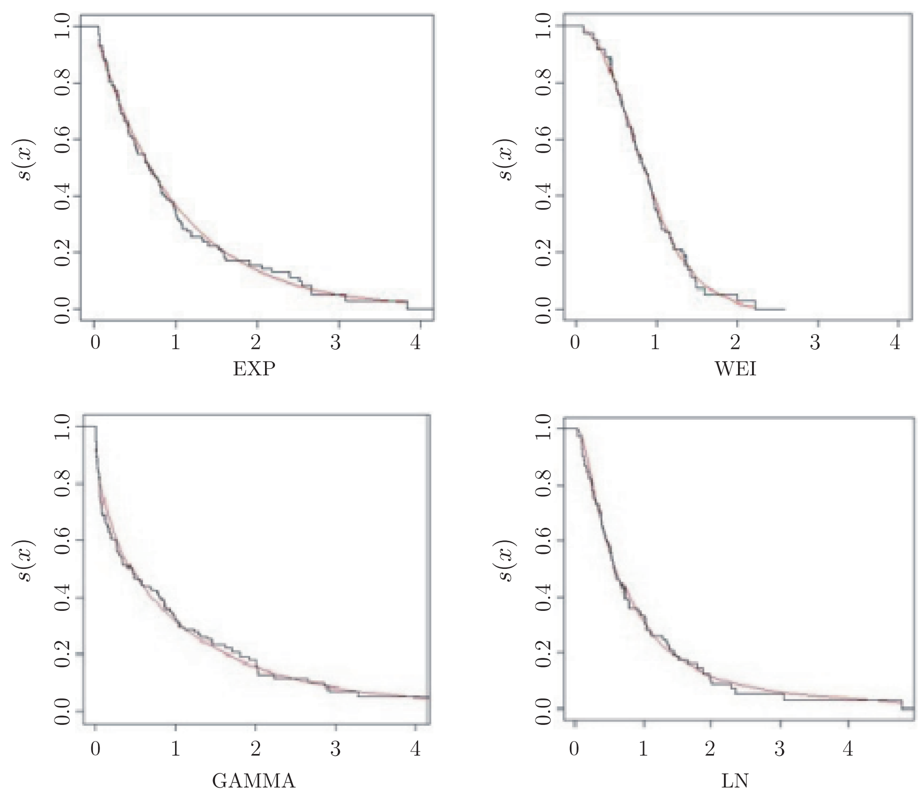

Figure 2. Plot of $\hat{\bar{F}}(t)$, represented by the unbroken line, and $\bar{F}(t)$, represented by the dotted line, from simulated system lifetimes under four models for $\bar{F}(t)$.

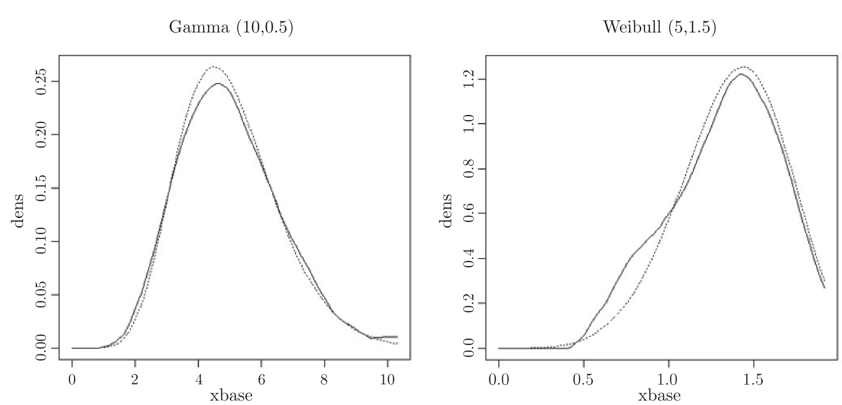

Figure 3. Plot of $\hat{f}(t)$, represented by unbroken line, and $f(t)$, represented by the dotted line, from simulated system lifetimes for $F =$ Gamma $(10, 0.5)$ or Weibull $(5, 1.5)$.

Table 1. Estimated truncated integrated squared error (TISE), based on approximately optimal bandwidth choice, for varying sample sizes.

$N_1 = N_2 = N_3$	TISE of $\hat{\hat{F}}$	TISE of $\hat{f}$	TISE of $\hat{r}$
30	0.0013	0.0267	14.3220
35	0.0010	0.0248	8.1653
40	0.0009	0.0223	6.0183
45	0.0008	0.0210	3.8107
50	0.0007	0.0210	2.8070
55	0.0006	0.0184	2.3740
60	0.0006	0.0184	1.9263
65	0.0005	0.0169	1.3720
70	0.0005	0.0164	1.2190
75	0.0004	0.0164	1.2398
80	0.0004	0.0150	1.1393
85	0.0004	0.0150	0.9908
90	0.0004	0.0141	0.9913
95	0.0004	0.0140	0.8010
100	0.0003	0.0139	0.8310

Similar results were found for our proposed density estimator based on samples of size 50 from these same three systems. The panels of Figure 3 show typical runs when 50 iid lifetimes were drawn from the Gamma $(10, 0.5)$ and Weibull $(5, 1.5)$ distributions for component lifetimes. The bandwidth w was chosen to minimize the integrated squared error using the empirical approach, based on the reference distribution discussed in Section 2.

Turning our attention to global performance measures for the proposed estimators, we calculated, from 1,000 replicated simulation runs, the median integrated squared error of each of the estimators. Results are tabulated below as the number of sampled system lifetimes, taken to be equal for each of the three systems in Figure 1, grew from 30 to 100. The integrated squared error was calculated between the 5% and 95% percentiles of an underlying Weibull distribution, since, as expected, the density estimator performs poorly in the tails of the distribution. We shall therefore refer to the approximated integral as truncated integrated squared error.

It can be seen from Table 1 that integrated squared error decreases steadily as the number of samples in each system increases from 30 to 100.

Acknowledgement

The research of F.J. Samaniego and P. Hall was supported in part by Grant W911NF11-1-0428 from the U.S. Army Research Office, and Grant DMS- 0906795 from the National Science Foundation, respectively.

Appendix

A.1. Central limit theorem for $\hat{\bar{F}}$

Theorem A.1, below, asserts that, for N sufficiently large, $\hat{\bar{F}}_{[2]}(t)$ is approximately normally distributed with mean $\bar{F}(t)$ and variance $\sigma^2_{[2]}\{\bar{F}(t)\}/N$, where $\sigma^2_{[2]}(p) = c_N(p)^{-1}$ and $c_N(p)$ is as in (2.12). The theorem also gives a uniform bound for the distance between $\hat{\bar{F}}_{[2]}$ and $\bar{F}$. As noted above Theorem 1, asymptotic properties are stated in the context of large N, where $N = \sum_i N_i$. We consider each N_i to be a function of N.

Theorem A.1. *Assume that* (2.5) *holds. Then* (a) *for each* $t \in (0, \infty)$, $N^{1/2}\{\hat{\bar{F}}_{[2]}(t) - \bar{F}(t)\}/\sigma_{[2]}\{\bar{F}(t)\}$ *is asymptotically normal* $\mathrm{N}(0,1)$, *and* (b) *if, in addition to* (2.5), $f = F'$ *exists and is bounded away from zero and infinity on an open subset* $\mathcal{O}$ *of the real line,* $\hat{\bar{F}} - \bar{F}$ *admits the following uniform bound on every compact subinterval* $\mathcal{T}$ *of* $\mathcal{O}$:

$$\sup_{t\in\mathcal{T}} \left|\hat{\bar{F}}_{[2]}(t) - \bar{F}(t)\right| = O_p\Big\{\big(N^{-1}\,\log N\big)^{1/2}\Big\}\,.$$

Similarly it can be proved that if (2.5) holds then the precursor $\hat{\bar{F}}_{[1]}(t)$ of $\hat{\bar{F}}_{[2]}(t)$ is approximately normally distributed with mean $\bar{F}(t)$ and variance $\sigma^2_{[1]}\{\bar{F}(t)\}/N$, where

$$\sigma^2_{[1]}(p) = \sum_{i=1}^{m} \frac{N_i}{N}\,\frac{h_i(p)\,\{1 - h_i(p)\}}{h_i'(p)^2}\,.$$

In the remainder of this section we give abbreviated proofs of Theorems 1, 2 and A.1.

A.2. Proof of part (b) of Theorem 1.

Since both $\hat{\bar{F}}_{[2]}$ and $\bar{F}$ are increasing, uniformly bounded functions then the uniform convergence claimed in part (b) of Theorem 1 follows if we establish pointwise convergence.

If $\hat{\bar{F}}_{[2]}(t)$ is not strongly consistent for $\bar{F}(t)$ then there exists a deterministic subsequence $\mathcal{S}$ of values of N such that, first, each N_i/N has a proper limit, b_i say, as $N \to \infty$ through $\mathcal{S}$, and secondly, $\hat{\bar{F}}_{[2]}(t)$ does not converge to $\bar{F}(t)$ almost surely as N diverges through $\mathcal{S}$. We shall confine attention to $N \in \mathcal{S}$, and argue by contradiction.

Since $\hat{\bar{F}}_{[2]}(t)$ does not converge almost surely to $\bar{F}(t)$ as $N \to \infty$ through $\mathcal{S}$ then there exists $\epsilon > 0$ such that

$$P\left\{\left|\hat{\bar{F}}_{[2]}(t) - \bar{F}(t)\right| > 2\,\epsilon \quad \text{fasl } N \in \mathcal{S}\right\} > 2\,\epsilon\,, \tag{A.1}$$

where the abbreviation "fasl" means "for all sufficiently large". Cases where $\hat{\bar{F}}_{[2]}(t)$ is arbitrarily close to 0 or 1 can be dismissed using a subsidiary argument, and so (A.1) implies that, for constants a and b satisfying $0 < a < b < 1$,

$$P\left\{\left|\hat{\bar{F}}_{[2]}(t) - \bar{F}(t)\right| > \epsilon \quad \text{and} \quad a \le \hat{\bar{F}}_{[2]}(t) \le b \quad \text{fasl } N \in \mathcal{S}\right\} > \epsilon\,. \tag{A.2}$$

We interpret p as a particular value of that quantity, for example the solution $p = \hat{\bar{F}}_{[2]}(t)$ of $D(p) = 0$ (as in Section 2), where

$$D(p) = \sum_{i=1}^{m} \left\{\frac{X_i(t)}{h_i(p)} - \frac{N_i - X_i(t)}{1 - h_i(p)}\right\} h_i'(p)\,, \tag{A.3}$$

$X_i(t)$ is as in (2.1), and h_i is the ith reliability polynomial in (1.2). As $N \to \infty$ through $\mathcal{S}$, $D(p)/N$ converges almost surely to

$$d_\infty(p) = \sum_{i=1}^{m} b_i \left\{\frac{\bar{F}_{T_i}(t)}{h_i(p)} - \frac{1 - \bar{F}_{T_i}(t)}{1 - h_i(p)}\right\} h_i'(p)\,,$$

uniformly in $p \in [a, b]$, where $\bar{F}_{T_i}$ is the reliability function for a generic lifetime T_i from the ith system. Now, $\bar{F}_{T_i}(t) = h_i\{\bar{F}(t)\}$ for $i = 1, \ldots, m$, and so

$$d_\infty(p) = \sum_{i=1}^{m} b_i \left[\frac{h_i\{\bar{F}(t)\}}{h_i(p)} - \frac{1 - h_i\{\bar{F}(t)\}}{1 - h_i(p)}\right] h_i'(p)\,, \tag{A.4}$$

which of course equals 0 if $p = \bar{F}(t)$. Therefore, if $p = \hat{\bar{F}}_{[2]}(t)$ is interpreted as the solution of $D(p) = 0$, then we can deduce from (A.2) that

$$d_\infty(p) = 0 \text{ for at least one } p \in (0,1) \text{ for which } p \ne \bar{F}(t)\,. \tag{A.5}$$

However, since h_i is an increasing function then the term within braces in (A.4) is a decreasing function of p, and therefore d_∞ also has this property. This result, and the fact that each h_i is a proper reliability polynomial, imply that the equation $d_\infty(p) = 0$ has a unique solution. Therefore it is not possible for (A.5) to be true, and so the initial assumption that $\hat{\bar{F}}_{[2]}(t)$ does not converge to $\bar{F}(t)$ almost surely must have been false.

A.3. Proofs of Theorem 1 (a) and (c) and of Theorem A.1(a)

Part (a) of theorem 1 may be deduced from the following considerations:

The fact that $\hat{p}(t)$ is a step function follows from the fact that the likelihood function $L(p, t)$ in (2.2) is constant between consecutive observed system lifetimes. To see that it is non-increasing, let $t_j < t_{j+1}$ be observed system lifetimes.

Then the ratio $L(p, t_{j+1})/L\{\hat{p}(t_j), t_{j+1}\}$ may be written as $L(p, t_j)/L\{\hat{p}(t_j), t_j\} \times g_i(p)/g_i\{\hat{p}(t_j)\}$ for some $i \in \{1, 2, \ldots, m\}$, where, for $x \in [0, 1]$, $g_i(x) = \{1 - h_i(x)\}/h_i(x)$. Since these functions $\{g_i\}$ are strictly decreasing, it follows that

$$L(p, t_{j+1}) < L\{\hat{p}(t_j), t_{j+1}\} \leq L\{\hat{p}(t_{j+1}), t_{j+1}\} \quad \text{for all} \quad p > \hat{p}(t_j),$$

which immediately implies that $\hat{\bar{F}}_{[2]}(t_{j+1}) \leq \hat{\bar{F}}_{[2]}(t_j)$. Finally, since $0 < t_j < \infty$ for $1 \leq j \leq N$, $\hat{\bar{F}}_{[2]}(0) = 1$ and $\hat{\bar{F}}_{[2]}(\infty) = 0$. Before treating part (c) of Theorem 1, we will derive part (a) of Theorem A.1. In the arguments below, if $\mathcal{N}_N$ is a random variable, and $a_N > 0$ and b_N are real numbers depending on N, we say that "$\mathcal{N}_N$ is asymptotically normally distributed with mean b_N and variance a_N^2" to signify that $(\mathcal{N}_N - b_N)/a_N$ is asymptotically normally distributed with zero mean and unit variance.

Let $p_0 = \bar{F}(t) \in (0, 1)$ be fixed, and recall that $X_i(t)$ and $c_N(p)$ are defined in (2.1) and (2.12), respectively. The derivative with respect to p of the logarithm of $L\{p, X_i(t)\}$, the latter defined in (2.2), is given by $D(p)$ in (A.3), and can be shown by Taylor expansion to equal

$$N^{1/2}\, c_N(p_0)\, V_N(p_0) - N\, c_N(p_0)\, (p - p_0) + O_p\Big\{N^{1/2}\, |p - p_0| + N\, (p - p_0)^2\Big\},$$

uniformly in p such that $|p - p_0| \leq \epsilon$ if $\epsilon > 0$ is sufficiently small, where

$$V_N(p) = \frac{1}{N^{1/2}\, c_N(p)} \sum_{i=1}^{m} a_i(p)\, \{X_i(t) - N_i\, h_i(p)\}\,. \tag{A.6}$$

Equivalently, defining $\delta = \delta(p, p_0) = N^{1/2}\,(p - p_0)$, we have that

$$\frac{D(p)}{N^{1/2}\, c_N(p_0)} = V_N(p_0) - \delta + O_p\big(N^{-1/2}\, |\delta| + N^{-1/2}\, \delta^2\big)\,, \tag{A.7}$$

uniformly in p in a neighbourhood of p_0. Using the fact that the variables $X_i(t)$ are independent sums of independent, uniformly bounded random variables, it can be proved from (A.6), using Lyapounov's central limit theorem, that $N^{1/2}\, c_N(p_0)\, V_N(p_0)$ is asymptotically normally distributed with zero mean and variance equal to the sum of the variances, i.e. to

$$\sum_{i=1}^{m} N_i\, a_i(p_0)^2\, h_i(p_0)\, \{1 - h_i(p_0)\}\,. \tag{A.8}$$

Since $p = \hat{\bar{F}}_{[2]}(t)$ is the solution of the equation $D(p) = 0$ when $p_0 = \bar{F}(t)$, it follows from (A.7) that the corresponding version of δ equals $V_N(p_0) + o_p(N^{-1/2})$, and therefore that $N^{1/2}\, \{\hat{\bar{F}}_{[2]}(t) - \bar{F}(t)\}$ is asymptotically normally distributed with zero mean and variance equal to $c_N(p_0)^{-2}$ multiplied by the quantity in (A.8). That is, part (a) of Theorem A.1 holds.

Next we establish part (c) of Theorem 1. Recall the definition of $\hat{\bar{F}}_{\mathbf{k}}$ in (2.3):

$$\hat{\bar{F}}_{\mathbf{k}}(t) = \sum_{i=1}^{m} k_i \hat{\bar{F}}_{N_i}(t) = \sum_{i=1}^{m} k_i h_i^{-1}\Big\{\hat{\bar{F}}_{T_i N_i}(t)\Big\} = \sum_{i=1}^{m} k_i h_i^{-1}\Big\{\frac{1}{N_i}\sum_{j=1}^{N_i} I_{(t,\infty)}(T_{ij})\Big\}.$$

Using this formula, Taylor expansion and Lyapounov's central limit theorem, it can be proved that a suitably standardized version of $\hat{\bar{F}}_{\mathbf{k}}(t)$ is asymptotically normal. More specifically, one may show that

$$\sqrt{N}\{\hat{\bar{F}}_{\mathbf{k}}(t) - \bar{F}(t)\} \xrightarrow{D} Y \sim \mathrm{N}\big(0, \sigma_k^2(t)\big),$$

where

$$\sigma_{\mathbf{k}}^2(t) = \sum_{i=1}^{m} \frac{k_i^2}{b_i} A_i\{\bar{F}(t)\}, \tag{A.9}$$

with

$$A_i(p) = \frac{h_i(p)\,\{1 - h_i(p)\}}{h_i'(p)^2}\,, \quad p = p(t) = \bar{F}(t)$$

and b_i equal to an m-dimentional probability vector, with $b_i > 0$ for all i, as discussed in the penultimate paragraph of Section 1.2. With the understaning that $p = p(t) = \bar{F}(t)$, one can deduce from part (a) of Theorem A.1 that

$$\sigma_{[2]}(t)^2 = \left\{\sum_{i=1}^{m} \frac{b_i}{A_i(p)}\right\}^{-1} \le \sigma_{\mathbf{k}}(t)^2\,, \tag{A.10}$$

where the inequality follows from following property, derived using the Cauchy-Schwarz inequality:

$$1 = \sum_{i=1}^{m} k_i = \sum_{i=1}^{m} k_i \left\{\frac{A_i(p)}{b_i}\,\frac{b_i}{A_i(p)}\right\}^{1/2} \le \left\{\sigma_{\mathbf{k}}(t)^2 \sum_{i=1}^{m} \frac{b_i}{A_i(p)}\right\}^{1/2}.$$

Part (c) of Theorem 1 follows directly from (A.10). Note that $\hat{\bar{F}}_{[1]}$ is a special case of $\hat{\bar{F}}_{\mathbf{k}}$, and that, if k_i is defined by (2.6), then in view of (A.9), $\sigma_{\mathbf{k}}^2(t) = \sigma_{[2]}^2(p)$, still with $p = \bar{F}(t)$. This establishes that aspect of part (c) of Theorem 1 pertaining to (2.6).

A.4. Proof of part (b) of Theorem A.1

In the argument below we take $a < b$ to be such that $\mathcal{P} = [a, b]$ is a subset of $\{\bar{F}(t) : t \in \mathcal{O}\}$, where $\mathcal{O}$ is the open set mentioned in the statement of part (b) of Theorem A.1. Furthermore, we let $\mathcal{P}'$ denote a regular grid of B points in the compact interval $\mathcal{P}$, with edge width asymptotic to a constant multiple of N^{-k} as N increases, where k will be chosen large but fixed. Thus, B is asymptotic to a constant multiple of N^k.

Let $D_1(u,p)$ denote the value taken by $D(p)$ in (A.3) when, in that formula, we replace the pair (p,t) by $(u,\bar{F}^{-1}(p))$. Let $u = p_1$, which can be interpreted as representing a particular value of $\hat{p}$ (a stochastic function of p), denote the solution of the equation $D_1(u,p)=0$, and put $\delta = N^{1/2}\,(p_1-p)$. The quantity p_1 is defined uniquely since $D(p)$ is a decreasing function of p. Define $X_i[p] = X_i\{\bar{F}^{-1}(p)\}$. Recall that $N=\sum_i N_i$ and that $c_N(p)$ is as defined at (2.12). Note that, by Taylor expansion,

$$a_i(p_1) = a_i(p) + (p_1-p)\,b_i(p) + O\big(|p_1-p|^2\big)\,,$$

uniformly in i, in p in any compact subinterval of the open subset $\mathcal{O}$ of $[0,1]$, and in $|p_1-p|\le\epsilon$, where $\epsilon>0$ does not depend on p, and the function b_i is defined by

$$b_i = a_i' = \frac{h_i''\,h_i\,(1-h_i) + 2\,(h_i')^2\,h_i - (h_i')^2}{\{h_i\,(1-h_i)\}^2}\,. \tag{A.11}$$

Define

$$W_N(p) = \frac{1}{N^{1/2}\,c_N(p)}\sum_{i=1}^m b_i(p)\,\{X_i(t) - N_i\,h_i(p)\}\,; \tag{A.12}$$

this is the same as the definition of $V_N(p)$ at (A.6), except that we replace a_i on the right-hand side there by b_i here. Recalling that the functions h_i, being polynomials, and the functions a_i, which are rational functions of polynomials, are smooth, we deduce by Taylor expansion that:

$$\begin{aligned}
D_1(p_1,p) &= \sum_{i=1}^m a_i(p_1)\Big[X_i[p]\,\{1-h_i(p_1)\} - \{N_i - X_i[p]\}\,h_i(p_1)\Big]\\
&= \sum_{i=1}^m \{a_i(p) + (p_1-p)\,b_i(p)\}\Big\{X_i[p] - N_i\,h_i(p) - N_i\,(p_1-p)\,h_i'(p)\Big\}\\
&\quad + O_p\big\{N\,(p_1-p)^2\big\}\\
&= \sum_{i=1}^m a_i(p)\,\{X_i[p] - N_i\,h_i(p)\} - (p_1-p)\sum_{i=1}^m \Big[N_i\,a_i(p)\,h_i'(p)\\
&\quad - b_i(p)\,\{X_i[p] - N_i\,h_i(p)\}\Big] + O_p\big\{N\,(p_1-p)^2\big\}\\
&= N^{1/2}\,c_N(p_1)\,V_N(p) - (p_1-p)\,\Bigg\{\sum_{i=1}^m N_i\,a_i(p)\,h_i'(p)
\end{aligned}$$

$$- N^{1/2}\, c_N(p_1)\, W_N(p)\Big\} + O_p\big\{N\,(p_1-p)^2\big\}$$
$$= N^{1/2}\, c_N(p_1)\, V_N(p) - N\, c_N(p)\,(p_1-p)$$
$$+ O_p\Big\{N\,(p_1-p)^2 + (N\,\log N)^{1/2}\,|p_1-p|\Big\}\,, \qquad \text{(A.13)}$$

uniformly in p in any compact subinterval of the open subset $\mathcal{O}$. (The appearance of the $\log N$ factor, on the last right-hand side, will be addressed in the subsequent five paragraphs; see the argument below (A.15) and (A.16).) Dividing both sides of (A.13) by $N^{1/2}\, c_N(p_1)$, noting that p_1 solves $D_1(p_1,p)=0$, and putting $\delta = N^{1/2}\,(p_1-p)$, we deduce that

$$0 = \frac{D_1(p_1,p)}{N^{1/2}\, c_N(p_1)} = V_N(p) - \delta + O_p\Big[\big\{N^{1/2}\, c_N(p_1)\big\}^{-1}\Big\{N\,(p_1-p)^2 + (N\,\log N)^{1/2}\,|p_1-p|\Big\}\Big]\,.$$

Hence, since $c_N(p)$ is bounded away from zero and infinity as $N\to\infty$,

$$\big|V_N(p)-\delta\big| = O_p\big\{(p_1-p)^2+(\log N)^{1/2}\,|p_1-p|\big\} = O_p\big\{N^{-1/2}\,(\log N)^{1/2}\big\}\,, \qquad \text{(A.14)}$$

which implies part (b) of Theorem A.1.

The $\log N$ factor, noted in the previous paragraph, derives from the fact that

$$\sup_{a\le p\le b} |W_N(p)| = O_p\big\{(\log N)^{1/2}\big\}\,. \qquad \text{(A.15)}$$

To derive (A.15) observe first that, using (2.11.3) [that is, assumption (2.11.3) in condition (2.11)], we can deduce that if p' is the point in $\mathcal{P}'$ nearest to $p\in\mathcal{P}$, then, for a constant $C_1>0$ not depending on p, it can be shown as in the next two paragraphs that

$$|W_N(p) - W_N(p')| \le C_1\left\{N^{-1} + N^{-1/2}\sum_{j=i}^{m}\big|X_i[p]-X_i[p']\big|\right\}$$
$$= O_p\big(N^{-1/2}\big)\,, \qquad \text{(A.16)}$$

for all $p\in\mathcal{P}$, using the above definite of p' as a function of p.

To derive (A.16) in detail, go back to the definition (A.12) of $W_N(p)$, from which it follows that

$$N^{1/2}\,|W_N(p) - W_N(p')|$$
$$= \left|\frac{1}{c_N(p)}\sum_{i=1}^{m} b_i(p)\,\{X_i(t) - N_i\,h_i(p)\} - \frac{1}{c_N(p')}\sum_{i=1}^{m} b_i(p')\,\{X_i(t) - N_i\,h_i(p')\}\right|$$

$$\leq \left|c_N(p)^{-1} - c_N(p')^{-1}\right| \sum_{i=1}^{m} |b_i(p')|\,|X_i(t) - N_i\,h_i(p')|$$
$$+ c_N(p)^{-1} \sum_{i=1}^{m} |b_i(p) - b_i(p')|\,|X_i(t) - N_i\,h_i(p')|$$
$$+ c_N(p)^{-1} \sum_{i=1}^{m} |b_i(p)|\,N_i\,|h_i(p) - h_i(p')|\,. \tag{A.17}$$

Exploiting the smoothness of the polynomials h_i, and of b_i and c_N in terms of h_i (see (A.11) and (2.12)), we deduce that if $C > 0$ is given then k, in the first paragraph of this proof, can be chosen sufficiently large to ensure that

$$\sup_{p\in\mathcal{P}} \left[\left|c_N(p)^{-1} - c_N(p')^{-1}\right| + \max_{1\leq i\leq m}\Big\{|b_i(p) - b_i(p')| + |h_i(p) - h_i(p')|\Big\}\right] = O\big(N^{-C}\big)\,. \tag{A.18}$$

Let S_i denote the smallest spacing between adjacent values of T_{ij}, for indices j in the range $1 \leq j \leq N_i$ such that $T_{ij} \in F^{-1}(\mathcal{P})$. Using (2.11.3) it can be shown, as we do below, that, for a constant $C_2 > 0$,

$$P\Big\{|X_i[p] - X_i[p']| \leq 1 \text{ for all } p \in \mathcal{P}\Big\} \geq P\big(S_i > C_2\,N^{-k}\big) = 1 - O\big(N^{-k}\,N_i\big)$$

as $N \to \infty$. This result is derived by, first, using the probability transform to switch from spacings of the distribution of the ith system lifetimes to spacings of the exponential distribution on $[0,\infty)$, and then applying a standard result on properties of exponential spacings (see e.g. Sections I.6 and III.3 of Feller, 1966, or Sections A.2–4.4 of Pyke, 1965). Therefore,

$$\begin{aligned} P\Big\{|X_i[p] - X_i[p']| \geq 2 \text{ for some } i = 1,\ldots,m \text{ and some } p \in \mathcal{P}\Big\} &\leq \sum_{i=1}^{m} P\Big\{|X_i[p] - X_i[p']| \geq 2 \text{ for some } p \in \mathcal{P}\Big\} \\ &= O\bigg(N^{-k}\sum_{i=1}^{m} N_i\bigg) = O\big(N^{-k+1}\big)\,. \end{aligned}$$

It follows that

$$P\Big\{\max_{1\leq i\leq m}\,\sup_{p\in\mathcal{P}}\,\big|X_i[p] - X_i[p']\big| \leq 1\Big\} \to 1\,. \tag{A.19}$$

Result (A.16) follows from (A.17), (A.18) and (A.19).

Bernstein's inequality implies directly that, for all $p \in \mathcal{P}$ and all sufficiently large $C_3 > 0$,

$$P\Big\{|W_N(p)| > C_3\,(\log N)^{1/2}\Big\} \le 2\,\exp\big(-C_4\,C_3^2\,\log N\big) = 2\,N^{-C_4\,C_3^2}\,,$$

where $C_4 > 0$ does not depend on C_3. Therefore,

$$\sup_{p\in\mathcal{P}} P\Big\{|W_N(p)| > C_3\,(\log N)^{1/2}\Big\} \le 2\,N^{-C_4\,C_3^2}\,. \qquad \text{(A.20)}$$

Choosing C_3 so large that $C_4\,C_3^2 > k$, we deduce from (A.20) that

$$\sup_{p\in\mathcal{P}} P\Big\{|W_N(p)| > C_3\,(\log N)^{1/2}\Big\} = o\big(N^{-k}\big)\,. \qquad \text{(A.21)}$$

Since the grid $\mathcal{P}'$ has edge width asymptotic to a constant multiple of N^{-k} then $\#\,\mathcal{P}' = O(N^k)$, and so the bound (A.21) implies that

$$P\bigg\{\sup_{p\in\mathcal{P}'} |W_N(p)| > C_3\,(\log N)^{1/2}\bigg\} = o(1)\,. \qquad \text{(A.22)}$$

The desired result (A.15) follows from (A.16) and (A.22).

A.5. Proof of Theorem 2

It can be shown from (A.14) that

$$\begin{aligned}\hat{f}(t) &= -\frac{1}{w}\int K'(u)\,\hat{F}(t-wu)\,du\\ &= f(t) + \tfrac{1}{2}\,\kappa_2\,w^2\,f''(t) - Q_N(t) + O_p\big\{(Nw)^{-1}\,\log N\big\} + o\big(w^2\big)\,, \qquad \text{(A.23)}\end{aligned}$$

where

$$Q_N(t) = \frac{1}{N^{1/2}\,w}\int K'(u)\,V_N\{\bar{F}(t-wu)\}\,du; \qquad \text{(A.24)}$$

to derive (A.23), we use the compactness of the support of K. In view of (2.1) and the definition of $V_N(p)$, formula (A.24) defines $Q_N(t)$ as a sum of independent random variables, and so $\mathrm{var}\{Q_N(t)\}$ can be computed explicitly. Starting from that formula, after lengthy calculations it can be proved that

$$(Nw)^2\,\mathrm{var}\{Q_N(t)\} = \kappa\,w\,f(t)\sum_{i=1}^{m} N_i\,\frac{a_i\{F(t)\}^2}{c_N\{F(t)\}^2}\,h_i'\{F(t)\} + o(w)\,. \qquad \text{(A.25)}$$

Asymptotic normality of $Q_N(t)$, with mean zero and asymptotic variance given by (A.25), can be proved using Lyapounov's theorem. This result and (A.23) imply Theorem 2.

A.6. Proof of Theorem 3

The result of Theorem 3 follows from (2.9) and Theorem 2.

References

Balakrishnan, N., Ng, H. K. T. and Navarro, J. (2011a). Exact nonparametric inference for component lifetime distribution based on lifetime data from systems with known signatures. *J. Nonparametr. Stat.* **23**, 741-752.

Balakrishnan, N., Ng, H. K. T. and Navarro, J. (2011b). Linear inference for type-II censored lifetime data of reliability systems with known signatures. *IEEE Trans. Reliab.* **TR 60**, 426-440.

Bhattacharya, D. and Samaniego, F. J. (2010). Estimating component characteristics from system failure-time data. *Nav. Res. Logist.* **57**, 380-389.

Boyles, R. A. and Samaniego, F. J. (1986). Estimating a distribution function based on nomination sampling. *J. Amer. Statist. Assoc.* **81**, 1039-1045.

Boyles, R. A. and Samaniego, F. J. (1987). On estimating component reliability for systems with random redundancy levels. *IEEE Trans. Reliab.* **R-36**, 403-407.

Bueno, V. C. (1988). A note on the component lifetime estimation of a multistate monotone system through the System lifetime. *Adv. Appl. Probab.* **20**, 686-689.

Chahkandi, M., Ahmadi, J. and Baratpour, S. (2014). Nonparametric prediction intervals for the lifetimes of coherent systems. *Statist. Papers*, to appear.

Eryilmaz, S. (2011) Estimation in coherent reliability systems through copulas. *Reliability Engineering and System Safety* **96**, 564-568.

Feller, W. (1966). *An Introduction to Probability Theory and its Applications.* Vol. 2. Wiley, New York.

Guess, F. M., Usher, J. S. and Hodgson, T. J. (1991). Estimating system and component reliabilities under partial information on cause of failure. *J. Statist. Plann. Inference* **29**, 75-85.

Kvam P. H. and Samaniego, F. J. (1993a). On the inadmissibility of empirical averages as estimators in ranked set sampling. *J. Statist. Plann. Inference* **36**, 39-55.

Kvam P. H. and Samaniego, F. J. (1993b). On maximum likelihood estimation based on ranked set samples, with applications to reliability. In *Advances in Reliability*, (Edited by A. Basu), 215-229. North Holland, Amsterdam.

Kvam P. H. and Samaniego, F. J. (1994). Nonparametric maximum likelihood estimation based on ranked set samples. *J. Amer. Statist. Assoc.* **89**, 526-537.

Meilijson, I. (1981). Estimation of the lifetime distribution of the parts from the autopsy statistics of the machine. *J. Appl. Probab.* **18**, 829-838.

Miyakawa, M. (1984). Analysis of incomplete data in competing risks model. *IEEE Trans. Reliab.* **R-33**, 293-296.

Moeschberger, M. L. and David, H. A. (1971). Life tests under competing cause of failure and the theory of competing risks. *Biometrics* **27**, 909-944.

Navarro, J., Samaniego, F. J. and Balakrishnan, N. (2011). Signature-based representations for the reliability of systems with heterogeneous components. *J. Appl. Probab.* **48**, 856-67.

Patterson, D. A., Gibson, G. and Katz, R. H. (1988). A case for redundant arrays of inexpensive disks (RAID). *SIGMOD Conference Proceedings*, 109-116.

Pyke, R. (1965). Spacings. *J. Roy. Statist. Soc. Ser. B* **27**, 395-449.

Samaniego, F. J. (2007). *System Signatures and Their Applications in Engineering Reliability.* Springer, New York.

Satyanarayana, A. and Prabhakar, A. (1978). A new topological formula and rapid algorithm for reliability analysis of complex networks. *IEEE Trans. Reliab.* **R-27**, 82-100.

Silverman, B. W. (1986). *Density Estimation for Statistics and Data Analysis.* Chapman and Hall, London.

Stokes, S. L. and Sager, T. W. (1988). Characterization of a ranked-set sample with application to estimating distribution functions. *J. Amer. Statist. Assoc.* **83**, 374-381.

Usher, J. and Hodgson, T. (1988). Maximum likelihood analysis of component reliability using masked system life-test data. *IEEE Trans. Reliab.* **37**, 550-555.

Department of Mathematics and Statistics, University of Melbourne, Victoria 3010, Australia.

Department of Statistics, University of California, Davis, California 95616, U.S.A.

E-mail: halpstat@ms.unimelb.edu.au

Department of Statistics, University of California, Davis, California 95616, U.S.A.

E-mail: golden.ucd@gmail.com

Department of Statistics, University of California, Davis, California 95616, U.S.A.

E-mail: fjsamaniego@ucdavis.edu

(Received March 2014; accepted October 2014)