The Annals of Statistics
2000, Vol. 28, No. 1, 116–150

ON THE ASYMPTOTIC DISTRIBUTION THEORY OF A CLASS OF CONSISTENT ESTIMATORS OF A DISTRIBUTION SATISFYING A UNIFORM STOCHASTIC ORDERING CONSTRAINT

BY MIGUEL A. ARCONES AND FRANCISCO J. SAMANIEGO

Binghamton University and University of California, Davis

We identify the asymptotic behavior of the estimators proposed by Rojo and Samaniego and Mukerjee of a distribution F assumed to be uniformly stochastically smaller than a known baseline distribution G. We show that these estimators are $\sqrt{n}$-convergent to a limit distribution with mean squared error smaller than or equal to the mean squared error of the empirical survival function. By examining the mean squared error of the limit distribution, we are able to identify the optimal estimator within Mukerjee's class under a variety of different assumptions on F and G. Similar theoretical results are developed for the two-sample problem where F and G are both unknown. The asymptotic distribution theory is applied to obtain confidence bands for the survival function $\overline{F}$ based on published data from an accelerated life testing experiment.

1. Introduction. Uniform stochastic ordering (USO) is a relation between random variables, or between their cumulative distribution functions, which quantifies the idea that the value taken by one variable tends to be smaller than the value taken by the other. There are many ways to model this idea. Uniform stochastic ordering, which we will denote by $X \leq_{(+)} Y$ or as $F \leq_{(+)} G$, where $X \sim F$ and $Y \sim G$, is characterized by the monotonicity of the ratio $\overline{F}(x)/\overline{G}(x)$ over the support set $\mathscr{S}_G$ of the distribution G, where $\overline{F} = 1 - F$ and $\overline{G} = 1 - G$. We say that $F \leq_{(+)} G$ if and only if $\overline{F}(x)/\overline{G}(x)$ is nonincreasing on $\mathscr{S}_G$. Among the many extant notions of ordering between distributions [see Shaked and Shanthikumar (1994) for a comprehensive discussion], uniform stochastic ordering has a particular relevance in the fields of reliability and survival analysis. It is easily shown that, in the absolutely continuous case where F and G have densities f and g, $F \leq_{(+)} G$ is equivalent to the ordering of the respective failure rates, that is, $F \leq_{(+)} G \Leftrightarrow$

$$\frac{f(x)}{\overline{F}(x)} \geq \frac{g(x)}{\overline{G}(x)} \qquad \forall\ x \in \mathscr{S}_F \cup \mathscr{S}_G. \tag{1.1}$$

Thus, when $F \leq_{(+)} G$ is known, we have a circumstance in which items having distribution F have a higher propensity to fail than items having distribution G. Because of the equivalence of (1.1) and the basic USO definition in the absolutely continuous case, the ordering is often called "the hazard rate or-

Received October 1998; revised October 1999.
AMS 1991 *subject classifications*. Primary 62G05; secondary 62G10, 62E20.
Key words and phrases. Uniform stochastic ordering, hazard rate ordering, Brownian motion, empirical processes.

dering." It is known [see Ross (1983)] to be a stronger form of ordering than ordinary stochastic ordering [$F \leq^{st} G$ iff $\overline{F}(x) \leq \overline{G}(x) \ \forall\ x \in \mathscr{S}_F \cup \mathscr{S}_G$] and to be a weaker ordering than likelihood ratio ordering [$F \leq^{lr} G$ iff $f(x)/g(x)$ is nonincreasing for $x \in \mathscr{S}_G \cup \mathscr{S}_F$].

There has been considerable recent interest in uniform stochastic ordering, both from the mathematical and the statistical perspectives. The relationship between USO and other ordering notions, and various implications of a USO assumption, have been explored in papers by Yanagimoto and Sibuya (1972), Whitt (1980), Keilson and Sumita (1982), Bagai and Kochar (1986) and Boland, El-Newehi and Proschan (1994). An excellent exposition of these and related results may be found in Shaked and Shanthikumar (1994). Statistical papers include Caperaa's (1988) treatment of a nonparametric testing problem, Dykstra, Kochar and Robertson's (1991) derivation of the likelihood ratio test for equality of distributions against a USO alternative, Rojo and Samaniego's (1991, 1993) studies of consistent estimation of a distribution F subject to a USO constraint and Mukerjee's (1996) treatment of the estimation of the pair (F, G), with $F \leq_{(+)} G$, based on independent samples from F and G. The latter four papers deal with estimators of distributions which obey USO constraints; while certain convergence questions are treated there, none of the cited work considers the delicate matter of an asymptotic distribution theory for the estimators in question. The purpose of this paper is to tackle these open questions in the contexts studied by Rojo and Samaniego (1993) and Mukerjee (1996).

The particular problem of interest in Section 2 may be described as follows. A random sample $X_1, \ldots, X_n$ is drawn from a distribution F on $[0, \infty)$, where F is unknown, but is subject to the constraint $F \leq_{(+)} G$, with G a known distribution on $[0, \infty)$. This problem arises in situations in which testing is done under two sets of conditions, one carefully controlled (in the laboratory, for example), yielding effectively unlimited data governed by G, and the other uncontrolled (in the field, e.g.) yielding a data set of modest size governed by F. In that context, Rojo and Samaniego proposed, as an estimator of the survival function $\overline{F} \equiv 1 - F$, the function

$$\widehat{\overline{F}}_n(x) = \inf_{\{y:\ 0 \leq y \leq x\}} \frac{\overline{G}(x)\overline{F}_n(y)}{\overline{G}(y)}, \tag{1.2}$$

where $\overline{F}_n$ is the empirical survival function (esf). They showed that $\widehat{F}_n \leq_{(+)} G$ and showed that strong uniform consistency of $\widehat{\overline{F}}_n$ as an estimator of $\overline{F}$ follows. They also showed that

$$\sup_{0 \leq x} \left|\widehat{\overline{F}}_n(x) - \overline{F}(x)\right| \leq \sup_{0 \leq x} \left|\overline{F}_n(x) - \overline{F}(x)\right|. \tag{1.3}$$

We prove that, except in a trivial case, we have strict inequality with positive probability. It is not true that for each x, $|\widehat{\overline{F}}_n(x) - \overline{F}(x)| \leq |\overline{F}_n(x) - \overline{F}(x)|$

(see the example after Theorem 2.1). However, we are able to prove that, asymptotically, $\hat{\hat{F}}_n$ is more accurate than $\overline{F}_n$. Next, we prove that

$$\left\{n^{1/2}\left(\hat{\hat{F}}_n(x) - \overline{F}(x)\right): x \geq 0\right\}$$

converges weakly to a certain stochastic process $\{L(x): x \geq 0\}$. As is well known, $\{n^{1/2}(\overline{F}_n(x) - \overline{F}(x)): x \geq 0\}$ converges weakly to $\{-W(F(x)): x \geq 0\}$, where $\{W(u) : 0 \leq u \leq 1\}$ is a Brownian bridge. We show that for each $x > 0$ and each $t > 0$,

$$\Pr\{|L(x)| \geq t\} \leq \Pr\{|W(F(x))| \geq t\}.$$

This enables us to provide conservative approximate confidence bands around the estimator $\hat{\hat{F}}_n$, that is, confidence regions which tend to achieve confidence levels somewhat higher than the nominal levels employed in their construction. If both $\overline{F}$ and $\overline{G}$ are discrete distributions, then in some cases we obtain that for each $t > 0$,

$$\Pr\{|L(x)| \geq t\} < \Pr\{|W(F(x))| \geq t\}.$$

The MSE of the limit distribution can be reduced by as much as 50% of that of $\overline{F}_n$. On the other hand, if both $\overline{F}$ and $\overline{G}$ are continuous, then for each $x > 0$ and each $t > 0$,

$$\Pr\{|L(x)| \geq t\} = \Pr\{|W(F(x))| \geq t\}.$$

However, $L(x)$ could have nonzero mean. Thus, $\hat{\hat{F}}_n$ has the same asymptotic MSE as $\overline{F}_n$, but it is not necessarily asymptotically unbiased.

Mukerjee (1996) proposed an estimator in the two-sample problem that also can be used in the one-sample case, using G itself instead of an estimate of G. Given $\alpha \in [0, 1]$, let

$$\hat{\hat{F}}_{\alpha,n}(x) = \inf_{y:\ 0 \leq y \leq x} \frac{\bar{F}_n(y)((1-\alpha)\overline{F}_n(x) + \alpha\overline{G}(x))}{(1-\alpha)\overline{F}_n(y) + \alpha\overline{G}(y)}. \tag{1.4}$$

When $\alpha = 0$, we have $\overline{F}_n(x)$; $\hat{\hat{F}}_{1,n}$ is the Rojo and Samaniego estimator. We give the asymptotic distribution of this estimator for general α. We also find the α giving a limit distribution with the smallest MSE. If both $\overline{F}$ and $\overline{G}$ are discrete distributions, this estimator does not improve upon the Rojo and Samaniego estimator. In one case, the MSE of the limit distribution has a strict minimum at $\alpha = 1$. However, when both $\overline{F}$ and $\overline{G}$ are continuous distributions, the MSE is minimized for $\alpha = \overline{F}(x)/\overline{F}(x) + \overline{G}(x)$. Thus, in this situation the preferred estimator is

$$\inf_{y:\ 0 \leq y \leq x} \frac{2\overline{F}_n(y)\overline{G}(x)\overline{F}_n(x)}{\overline{G}(x)\overline{F}_n(y) + \overline{G}(y)\overline{F}_n(x)}. \tag{1.5}$$

This estimator can reduce the MSE of the limit distribution by as much as 25%.

In Section 3, we turn our attention to the asymptotic distribution theory associated with Mukerjee's (1996) estimators of F and G when it is known that $F \leq_{(+)} G$ and independent random samples are available from both F and G. After obtaining the relevant asymptotic distributions, we derive the general form of the optimal estimator within Mukerjee's class and identify conditions under which Mukerjee's recommended estimator within this class is in fact optimal.

In the final section, we put the asymptotic distribution theory developed in Section 2 to use in providing confidence bands for the survival function $\overline{F}$ based on data from an accelerated life test on Kevlan/epoxy pressure vessels. We close by making a few remarks about the extent to which similar asymptotic results obtain under random censoring and by discussing the application of our asymptotic results to testing problems involving the constraint $F \leq_{(+)} G$.

The proofs of all theoretical results in the sequel have been relegated to the Appendix.

2. Asymptotic distribution theory in the one sample problem. In this section, we will derive the asymptotic distribution of the Rojo–Samaniego estimator $\widehat{\overline{F}}_n$ in (1.2) based on a random sample drawn from F, where $F \leq_{(+)} G$ with G known. We assume throughout that F and G are supported on $[0, \infty)$, and we denote the least upper bounds of their support by

$$t_F = \sup\left\{t: \overline{F}(t) > 0\right\} \quad \text{and} \quad t_G = \sup\left\{t: \overline{G}(t) > 0\right\}, \tag{2.1}$$

respectively. Since $F \leq_{(+)} G$, we have that $t_F \leq t_G$. We will make repeated use of the following additional notation: let

$$H(x) = \frac{\overline{F}(x)}{\overline{G}(x)} \quad \text{for } x \in [0, t_G). \tag{2.2}$$

The constraint that $F \leq_{(+)} G$ is equivalent to the assumption that H is nonincreasing in $[0, t_G)$. Further, we define the function ℓ as

$$\ell(x) = \inf\left\{y: H(y) = H(x)\right\}, \qquad x \in [0, t_G). \tag{2.3}$$

When H is strictly decreasing, we have $\ell(x) \equiv x$; in this case, we will demonstrate that $\widehat{\overline{F}}_n$ and $\overline{F}_n$ are asymptotically equivalent. Finally, for $x \in [0, t_G)$, define the set $\mathscr{A}(x)$ as

$$\begin{aligned} \mathscr{A}(x) = &\left\{\overline{F}(y): y \leq x,\ H(y) = H(x)\right\} \\ &\cup \left\{\overline{F}(y-): y \leq x,\ H(y-) = H(x)\right\}. \end{aligned} \tag{2.4}$$

As we will see, the character of the set $\mathscr{A}(x)$ for $x \in [0, t_G)$ has a strong influence on the limiting distribution of the process

$$\left\{\sqrt{n}\left(\widehat{\overline{F}}_n(t) - \overline{F}(t)\right): t \in [0, t_G)\right\}.$$

Although the general form of the set $\mathscr{A}(x)$ can be very complicated, we can determine the form of the set $\mathscr{A}$ for continuous and for discrete distributions.

If X and Y have continuous cdf's, then $\{y \le x\colon H(y) = H(x)\} = [l(x), x]$ and $\mathscr{A}(x) = [\overline{F}(x), \overline{F}(\ell(x))]$. If X and Y are discrete r.v.'s with finite support $\{x_1 < \cdots < x_m\}$, where $0 = x_0 < x_1 < \cdots < x_m$, then there are positive integers $i \le k$ such that $H(x_{i-1}) > H(x_i) = H(x)$ and $x_i \le x_k \le x < x_{k+1}$; and

$$\mathscr{A}(x) = \{\overline{F}(x_j)\colon x_j \le x, \bar{H}(x_j) = \bar{H}(x)\}.$$

In studying the asymptotic behavior of $\widehat{\overline{F}}_n(x)$, we will restrict attention to $x \in [0, t_F)$ since, for $x \in [t_F, t_G)$,

$$\widehat{\overline{F}}_n(x) = \overline{F}_n(x) = \overline{F}(x) = 0.$$

Our treatment of the Rojo–Samaniego estimator begins with a refinement of a result on the relative accuracy of $\widehat{\overline{F}}_n$ and $\overline{F}_n$. Rojo and Samaniego (1993) proved that, for fixed known G and for all $F \le_{(+)} G$,

$$\sup_{0 \le x < t_F} \left|\widehat{\overline{F}}_n(x) - \overline{F}(x)\right| \le \sup_{0 \le x < t_F} \left|\overline{F}_n(x) - \overline{F}(x)\right|, \tag{2.5}$$

with strict inequality holding for some F. This result may be strengthened as follows.

THEOREM 2.1. *Let $X_1, \dots, X_n \overset{iid}{\sim} F$, where $F \le_{(+)} G$, with G known. Assume that $t_F > 0$.*

(i) *If $\overline{G}(t_F-) = 1$, then for each $0 \le x < t_F$,*

$$\inf_{y\colon 0 \le y \le x} \frac{\overline{G}(x)\overline{F}_n(y)}{\overline{G}(y)} = \overline{F}_n(x).$$

(ii) *If $\overline{G}(t_F-) < 1$, then, with positive probability,*

$$\sup_{0 \le x < t_F} \left| \inf_{y\colon 0 \le y \le x} \left(\frac{\overline{G}(x)\bar{F}_n(y)}{\overline{G}(y)} - \overline{F}(x) \right) \right| < \sup_{0 \le x < t_F} \left|\overline{F}_n(x) - \overline{F}(x)\right|.$$

It is not true that the inequality $|\widehat{\overline{F}}_n(x) - \overline{F}(x)| \le |\overline{F}_n(x) - \overline{F}(x)|$ holds for every $x \ge 0$. Suppose, for example, that $\overline{F} = \overline{G}$, with each distribution placing mass 1/3 at $x = 1, 2, 3$. Then, $\widehat{\overline{F}}_n(x) = \min(3^{-1}, 2^{-1}\overline{F}_n(1), \overline{F}_n(2))$. If a sample of size $n = 4$ yields $\{1, 1, 1, 3\}$, so that $\overline{F}_n(1) = \overline{F}_n(2) = 2^{-2}$, then $|\widehat{\overline{F}}_n(2) - \overline{F}(2)| = 5/24 > 1/12 = |\overline{F}_n(2) - \overline{F}(2)|$.

Next, we consider the asymptotic theory of the Rojo–Samaniego estimator. As is well known, $\{n^{1/2}(F_n(x) - F(x))\colon x \in \mathbb{R}\}$ converges in distribution to $\{W(F(x))\colon x \in \mathbb{R}\}$, where $\{W(t)\colon 0 \le t \le 1\}$ is a Brownian bridge. As a consequence, $\{n^{1/2}(\overline{F}_n(x) - \overline{F}(x))\colon x \in \mathbb{R}\}$ converges in distribution to

$\{-W(F(x))\colon x \in \mathbb{R}\}$. The convergence of $\widehat{\widehat{F}}_n(x)$ in distribution, at rate $n^{1/2}$, is established in the following:

THEOREM 2.2. *Let* $X_1, \dots, X_n \overset{iid}{\sim} F$, *where* $F \leq_{(+)} G$, *with* G *known. Then the stochastic process*

$$\left\{ n^{1/2} \inf_{y\colon 0 \le y \le x} \left(\frac{\overline{G}(x)\overline{F}_n(y)}{\overline{G}(y)} - \overline{F}(x) \right) : 0 \le x < t_F \right\} \tag{2.6}$$

converges weakly to $\{L(x)\colon 0 \le x < t_F\}$ *where*

$$L(x) := -\sup_{t \in \mathscr{A}(x)} \frac{\overline{F}(x) W(1-t)}{t}. \tag{2.7}$$

As one should expect, when $\overline{F}$ and $\overline{G}$ are close together, $\widehat{\widehat{F}}_n$ is superior to $\overline{F}_n$. Suppose, for example, that $\overline{F} = \overline{G}$ is a continuous survival function. Then,

$$\sup_{x \ge 0} n^{1/2} \left| \overline{F}_n(x) - \overline{F}(x) \right| \overset{d}{\to} \sup_{x \ge 0} |W(F(x))| = \sup_{0 \le t \le 1} \left| W(t) \right|,$$

while

$$\sup_{x \ge 0} n^{1/2} \left| \widehat{\widehat{F}}_n(x) - \overline{F}(x) \right| \overset{d}{\to} \sup_{x \ge 0} \left| \sup_{y\colon 0 \le y \le x} \frac{\overline{F}(x) W(F(x))}{\overline{F}(y)} \right|$$

$$= \sup_{x \ge 0} \sup_{y\colon 0 \le y \le x} \frac{\overline{F}(x) W^+(F(x))}{\overline{F}(y)} = \sup_{0 \le t \le 1} W^+(t).$$

The tails of $\sup_{0 \le t \le 1} W(t)$ and $\sup_{0 \le t \le 1} W^+(t)$ can be found on page 85 of Billingsley (1968).

Theorem 2.2 allows us to discuss the asymptotic bias and asymptotic MSE of $\widehat{\widehat{F}}_n(x)$. It is well known that for each $k > 0$, $E[(n^{1/2} \sup_{x > 0} |\overline{F}_n(x) - \overline{F}(x)|)^k]$ is a bounded sequence. By (2.6), we can assert that the same is true for

$$n^{1/2} \sup_{0 \le x < t_F} \left| \widehat{\widehat{F}}_n(x) - \overline{F}(x) \right|.$$

Hence, by uniform integrability, the moments of $n^{1/2}(\widehat{\widehat{F}}_n(x) - \overline{F}(x))$ converge to those of $L(x)$. If $\mathscr{A}(x) = \{\overline{F}(x)\}$, then $L(x) = W(F(x))$ and $\widehat{\widehat{F}}_n(x)$ is asymptotically equivalent to $\overline{F}_n(x)$. Otherwise, $L(x) \le -W(F(x))$ and $E[L(x)] < -E[W(\overline{F}(x))] = 0$; that is, $\widehat{\widehat{F}}_n(x)$ is asymptotically biased. Observe that if $t \in \mathscr{A}(x)$ and $t \ne \overline{F}(x)$, then

$$E[L(x)] \le -E\left[\max\left(\frac{\overline{F}(x) W(1-t)}{t}, W(F(x)) \right) \right]$$

$$= -E\left[\max\left(0, W(F(x)) - \frac{\overline{F}(x) W(1-t)}{t} \right) \right]$$

$$= -(2\pi)^{-1/2} (\overline{F}(x)(1 - t^{-1}\overline{F}(x)))^{1/2} < 0.$$

As to the accuracy and MSE of the Rojo–Samaniego estimator, we have the following:

THEOREM 2.3. *Consider the conditions:*

(a) $H(l(x)) = H(l(x)-)$ *and* $\overline{F}(l(x)) < \overline{F}(l(x)-)$.
(b) $\{\overline{F}(y)\colon l(x) \le y \le x\} \ne [\overline{F}(l(x)), \overline{F}(x)]$.

(i) *If neither* (a) *nor* (b) *holds, then for each* $t > 0$,

$$\Pr\{|L(x)| \ge t\} = \Pr\{|W(F(x))| \ge t\}. \tag{2.8}$$

(ii) *If either* (a) *or* (b) *holds or both hold, then for each* $t > 0$,

$$\Pr\{|L(x)| \ge t\} < \Pr\{|W(F(x))| \ge t\}. \tag{2.9}$$

It follows from the result above that a conservative $100(1-\alpha)\%$ confidence interval for $\overline{F}(x)$ is given by

$$\hat{\hat{\overline{F}}}_n(x) \pm z_{\alpha/2} n^{-1/2} (\hat{\hat{\overline{F}}}_n(x)(1 - \hat{\hat{\overline{F}}}_n(x)))^{1/2},$$

where $\Pr\{Z \ge z_{\alpha/2}\} = \alpha/2$ and Z is a standard normal r.v.

Suppose that $\overline{F}$ and $\overline{G}$ are discrete distributions giving probability 1 to the points $x_1 < \cdots < x_m$. It follows from the previous theorem that if $H(x_i) > H(x_{i+1})$ for each i, then $n^{1/2}(\hat{\hat{\overline{F}}}_n(x) - \overline{F}_n(x)) \xrightarrow{\Pr} 0$. Otherwise for some x's, the limit distribution of $n^{1/2}(\hat{\hat{\overline{F}}}_n(x) - \overline{F}(x))$ has smaller MSE than the limit distribution of $n^{1/2}(\overline{F}_n(x) - \overline{F}(x))$. If $\overline{F}$ and $\overline{G}$ are continuous and $\ell(x) = x$, the asymptotic distribution of $\hat{\hat{\overline{F}}}_n$ is that of $\overline{F}_n$; if, however, $\ell(x) < x$, then $\hat{\hat{\overline{F}}}_n$ is not asymptotically normal.

We now expand our focus to consider estimators of $\overline{F}$ of the form considered by Mukerjee (1996). Let $\hat{\hat{\overline{F}}}_{\alpha,n}(x)$ be the estimator given in (1.4). By the definition of the Rojo–Samaniego estimator, we have that

$$\sup_{\{y\colon 0 \le y \le x\}} \frac{\overline{G}(y)}{\overline{F}_n(y)} = \frac{\overline{G}(x)}{\hat{\hat{\overline{F}}}_n(x)},$$

which implies that

$$\hat{\hat{\overline{F}}}_{\alpha,n}(x) = \frac{\hat{\hat{\overline{F}}}_n(x)((1-\alpha)\overline{F}_n(x) + \alpha\overline{G}(x))}{(1-\alpha)\hat{\hat{\overline{F}}}_n(x) + \alpha\overline{G}(x)}.$$

Since $\hat{\hat{\overline{F}}}_n(x) \le \overline{F}_n(x)$, we have that

$$\hat{\hat{\overline{F}}}_n(x) \le \hat{\hat{\overline{F}}}_{\alpha,n}(x) \le \overline{F}_n(x).$$

This implies that $\hat{\hat{\overline{F}}}_{\alpha,n}$ approaches $\overline{F}$ at least as efficiently as $\overline{F}_n$ does; that is,

$$\sup_{0 \le x} \left|\hat{\hat{\overline{F}}}_{\alpha,n}(x) - \overline{F}(x)\right| \le \sup_{0 \le x} \left|\overline{F}_n(x) - \overline{F}(x)\right|.$$

Among other things, this latter inequality shows that every estimator in the Mukerjee class has the property established in Theorem 2.1 for $\widehat{\widehat{F}}_n$. The asymptotic distribution of $\widehat{\widehat{F}}_{\alpha,n}(x)$ is identified in the following result.

THEOREM 2.4. *Let $X_1, \ldots, X_n \overset{iid}{\sim} F$, where $F \leq_{(+)} G$, with G known. Then*

$$\left\{ n^{1/2} \inf_{y:\ 0\le y\le x} \left(\frac{\overline{F}_n(y)((1-\alpha)\overline{F}_n(x) + \alpha\overline{G}(x))}{(1-\alpha)\overline{F}_n(y) + \alpha\overline{G}(y)} - \overline{F}(x) \right) : 0 \le x < t_F \right\} \tag{2.10}$$

converges weakly to $\{L_\alpha(x)\colon 0 \le x < t_F\}$ where

$$L_\alpha(x) := -\sup_{t\in\mathscr{A}(x)} \frac{(1-\alpha)tW(F(x)) + \alpha\overline{G}(x)W(1-t)}{(1-\alpha)t + \alpha t(\overline{F}(x))^{-1}\overline{G}(x)}. \tag{2.11}$$

As with the Rojo–Samaniego estimator, the moments of estimators in the Mukerjee class $\{\widehat{\widehat{F}}_{\alpha,n}(x)\colon \alpha \in [0,1]\}$ converge to those of the limiting distribution $L_\alpha(x)$. It will be both feasible and useful to examine the limiting MSE of estimators in this class. In Theorem 2.5, we identify the value of α which minimizes this MSE. The optimal value α_0 depends, in general, on the set $\mathscr{A}(x)$ in a complicated way. We will limit ourselves to examining three specific examples in which α_0 can be identified explicitly.

THEOREM 2.5. *Suppose that $\mathscr{A}(x)$ contains more than one element. Then $E[(L_\alpha(x))^2]$ is minimized at*

$$\alpha_0 = \alpha_0(\overline{F}) = \frac{(1-a_1)\overline{F}(x)}{(1-a_1)\overline{F}(x) + \overline{G}(x)(a_2 - a_1)}, \tag{2.12}$$

where

$$a_1 := \left(\frac{1}{\overline{F}(x)} - \frac{1}{m(x)}\right)^{-1} E\left[B\left(\frac{1}{\overline{F}(x)} - \frac{1}{m(x)}\right) \sup_{t\in\mathscr{A}(x)} B\left(\frac{1}{t} - \frac{1}{m(x)}\right)\right],$$

$$a_2 := \left(\frac{1}{\overline{F}(x)} - \frac{1}{m(x)}\right)^{-1} E\left[\left(\sup_{t\in\mathscr{A}(x)} B\left(\frac{1}{t} - \frac{1}{m(x)}\right)\right)^2\right]$$

and $m(x)$ is the largest element of $\mathscr{A}(x)$. The estimator corresponding to this choice of α is

$$\widehat{\widehat{F}}_{\hat\alpha_0,n}(x) = \inf_{y:\ 0\le y\le x} \frac{\overline{F}_n(y)((a_2-a_1)\overline{G}(x)\overline{F}_n(x) + (1-a_1)\overline{F}_n(x)\overline{G}(x))}{(a_2-a_1)\overline{G}(x)\overline{F}_n(y) + (1-a_1)\overline{F}_n(x)\overline{G}(y)}, \tag{2.13}$$

where $\hat\alpha_0 = \alpha_0(\overline{F}_n)$. Moreover, for this choice of α,

$$E[(L_{\alpha_0}(x))^2] = \overline{F}(x)\left(A + \frac{\overline{F}(x)}{m(x)}(1-A)\right) - (\bar{F}(x))^2 < E[(W(F(x))^2], \tag{2.14}$$

where

$$0 < A := \frac{a_2 - a_1^2}{a_2 - 2a_1 + 1} < 1.$$

The optimal choice of α depends on several parameters, mainly on $\mathscr{A}(x)$. We close this section with some optimality results for three possible types of $\mathscr{A}(x)$.

THEOREM 2.6. *Suppose that $l(x) \neq x$ and $\mathscr{A}(x) = [\overline{F}(x), \overline{F}(l(x))]$. Then, $E[(L_\alpha(x))^2]$ is minimized at*

$$\alpha_0 = \alpha_0(\overline{F}) = \frac{\overline{F}(x)}{\overline{F}(x) + \overline{G}(x)}, \tag{2.15}$$

$$E[L_{\alpha_0}^2(x)] = \overline{F}(x)\left(\frac{3}{4} + \frac{\overline{F}(x)}{4\overline{F}(l(x))}\right) - (\overline{F}(x))^2 \tag{2.16}$$

and the estimator with this choice α_0 is

$$\widehat{\overline{F}}_{\hat{\alpha}_0,n}(x) = \inf_{y:\ 0\le y\le x} \frac{2\overline{F}_n(y)\overline{G}(x)\overline{F}_n(x)}{\overline{G}(x)\overline{F}_n(y) + \bar{F}_n(x)\overline{G}(y)}. \tag{2.17}$$

where $\hat{\alpha}_0 = \alpha_0(\overline{F}_n)$.

It is easy to see that in the previous case, $E[L_{\alpha_0}^2(x)]/E[W^2(F(x))] \ge 3/4$, with the lower bound attained if $\overline{F}(\ell(x)) = 1$. Although the reduction in the MSE is not dramatic, neither is it negligible. If $\overline{G}$ and $\overline{F}$ are both continuous, either $L(x) = W(F(x))$ or we have the case in Theorem 2.6. Thus, in the continuous case, the estimate in (2.17) is either equivalent to the Rojo–Samaniego estimator or superior to it.

Next, we consider two special cases in which both $\overline{G}$ and $\overline{F}$ are discrete distributions.

THEOREM 2.7. *Suppose that $\mathscr{A}(x) = \{\overline{F}(x), m(x)\}$, for some $m(x) \neq \overline{F}(x)$, then $E[(L_\alpha(x))^2]$ is minimized at $\alpha_0 = 1$, that is, the Rojo–Samaniego estimator minimizes the MSE among estimators of the form (1.4). The asymptotic MSE of this estimator is given by*

$$E[(L_{\alpha_0}(x))^2] = 2^{-1}\overline{F}(x)\left(1 + \frac{\overline{F}(x)}{m(x)}\right) - (\overline{F}(x))^2. \tag{2.18}$$

In the previous theorem $E[L_{\alpha_0}^2(x)]/E[W^2(F(x))] \ge 1/2$; the lower bound is attained if $m(x) = 1$.

THEOREM 2.8. *Suppose that* $\mathscr{A}(x) = \{\overline{F}(x), y_1, y_2\}$ *where* $\overline{F}(x) < y_1 < y_2$, *then* $E[(L_\alpha(x))^2]$ *is minimized at*

$$\alpha_0 = \frac{(1-a_1)\overline{F}(x)}{(1-a_1)\overline{F}(x) + \overline{G}(x)(a_2 - a_1)},$$

where

$$a_1 = 2^{-1} - t2^{-2} - (2\pi)^{-1} \arctan((t/(1-t))^{1/2}),$$
$$a_2 = 2^{-1} + t2^{-2} + (2\pi)^{-1}(t(1-t))^{1/2} - (2\pi)^{-1} \arctan((t/(1-t))^{1/2})$$

and

$$t = \frac{1/y_1 - 1/y_2}{1/\overline{F}(x) - 1/y_2}.$$

In order to construct an asymptotic confidence interval, we need to have bounds on the tail of the limit distribution. Since the only case in which the general Mukerjee estimator is recommended is the case of continuous distributions, we restrict consideration to this case.

THEOREM 2.9. *If the assumptions of Theorem* 2.4 *hold, and if* F *and* G *are continuous distributions, then*

(i) *If either* $\alpha \in \{0, 1\}$ *or* $\mathscr{A}(x) = \{\overline{F}(x)\}$, *then for each* $t > 0$,

$$\Pr\{|L_\alpha(x)| \geq t\} = \Pr\{|W(F(x))| \geq t\}.$$

(ii) *If* $0 < \alpha < 1$ *and* $\mathscr{A}(x) \neq \{\overline{F}(x)\}$, *then for each* $t > 0$,

$$\Pr\{|L_\alpha(x)| \geq t\} < \Pr\{|W(F(x))| \geq t\}.$$

The theorem above demonstrates that, for continuous distributions, both the empirical survival function and the Rojo–Samaniego estimator of survival are asymptotically dominated by estimators $\widehat{\widehat{F}}_\alpha$ in the Mukerjee class with $\alpha \in (0, 1)$. Further, we have that if $\overline{F}$ and $\overline{G}$ are continuous distributions, a conservative 100(1–α%) confidence interval for $\overline{F}(x)$ is given by

$$\widehat{\widehat{F}}_{\hat{\alpha}_0,n}(x) \pm z_{\alpha/2} n^{-1/2} (\widehat{\widehat{F}}_{\hat{\alpha}_0,n}(x)(1 - \widehat{\widehat{F}}_{\hat{\alpha}_0,n}(x)))^{1/2},$$

where $\widehat{\widehat{F}}_{\hat{\alpha}_0,n}(x)$ is given in (2.13) and $z_{\alpha/2}$ is the $1-\alpha/2$ quantile of the standard normal distribution.

It is worth noting that, except when $\alpha = 1$, the estimator above cannot be guaranteed to satisfy the USO constraint with respect to G. A consistent estimator which does satisfy the USO constraint can be constructed by applying the transformation in (2.4) of Rojo and Samaniego (1993) to $\widehat{\widehat{F}}_{\alpha,n}(x)$.

3. Asymptotic distribution theory for the two sample problem. Suppose now that independent random samples are available from the distributions F and G, where $F \leq_{(+)} G$, and we wish to estimate F (or F and G) from these data. Mukerjee (1996) proposed estimators of the form

$$\widehat{\widehat{\overline{F}}}_{\alpha,n}(x) = \inf_{y:\ 0\le y\le x} \frac{\overline{F}_n(y)((1-\alpha)\overline{F}_n(x)+\alpha\overline{G}_m(x))}{(1-\alpha)\overline{F}_n(y)+\alpha\overline{G}_m(y)} \tag{3.1}$$

and

$$\widehat{\widehat{\overline{G}}}_{\alpha,m}(x) = \sup_{y:\ 0\le y\le x} \frac{\bar{G}_m(y)((1-\alpha)\overline{F}_n(x)+\alpha\overline{G}_m(x))}{(1-\alpha)\overline{F}_n(y)+\alpha\overline{G}_m(y)}, \tag{3.2}$$

where $\alpha \in [0, 1]$, $(\overline{F}_n, \overline{G}_m)$ are the empirical survival functions based on samples of size n and m from F and G, respectively, and x is such that $\max(\overline{F}_n(x), \overline{G}_m(x)) > 0$. He demonstrated the strong consistency of these estimators for arbitrary α and suggested the use of

$$\hat{\alpha} = \frac{m}{m+n}, \tag{3.3}$$

based on a heuristic argument related to maximizing a nonparametric likelihood.

The aim of this section is to develop the asymptotic distribution theory of the estimators $\widehat{\widehat{\overline{F}}}_{\alpha,n}$ and $\widehat{\widehat{\overline{G}}}_{\alpha,m}$. We provide detailed derivations only for $\widehat{\widehat{\overline{F}}}_{\alpha,n}$ since the arguments for $\widehat{\widehat{\overline{G}}}_{\alpha,m}$ are similar. Our results include asymptotic distributions for arbitrary α and the identification of the general form of the value α_0 of α which minimizes the asymptotic MSE of $\widehat{\widehat{\overline{F}}}_{\alpha}$. We will also identify circumstances under which Mukerjee's recommended $\hat{\alpha}$ is asymptotically optimal, but also demonstrate that other choices of α are superior in alternative situations. The notation established for t_F, t_G, $H(x)$, $\ell(x)$ and $\mathscr{A}(x)$ in (2.1)–(2.4) carry over to this section without alteration. In addition, we will use the notation

$$\{n^{1/2}(F_n(x)-F(x))\colon x\in\mathbb{R}\} \xrightarrow{d} \{W_1(F(x))\colon x\in R\}$$

and

$$\{m^{1/2}(G_m(x)-G(x))\colon x\in\mathbb{R}\} \xrightarrow{d} \{W_2(G(x))\colon x\in R\},$$

where $\{W_1(t)\colon 0\le t\le 1\}$ and $\{W_2(t)\colon 0\le t\le 1\}$ are two independent Brownian bridges. As a consequence, we have that

$$\{n^{1/2}(\overline{F}_n(x)-\overline{F}(x))\colon x\in\mathbb{R}\} \xrightarrow{d} \{-W_1(F(x))\colon x\in R\}$$

and

$$\{m^{1/2}(\overline{G}_m(x)-\overline{G}(x))\colon x\in\mathbb{R}\} \xrightarrow{d} \{-W_2(G(x))\colon x\in R\}.$$

Since the estimators are related to the empirical processes $n^{1/2}(\overline{F}_n(x)-\overline{F}(x))$ and $m^{1/2}(\overline{G}_m(x)-\overline{G}(x))$, we have joint convergence of the estimators of $\overline{F}$ and $\overline{G}$.

THEOREM 3.1. *Let* $X_1, \ldots, X_n \overset{iid}{\sim} F$ *and* $Y_1, \ldots, Y_m \overset{iid}{\sim} G$, *where* $F \leq_{(+)} G$.

(i) *If* $m/n \to \infty$, *then*

$$\left\{ n^{1/2} \inf_{y:\ 0 \leq y \leq x} \left(\frac{\overline{F}_n(y)((1-\alpha)\overline{F}_n(x) + \alpha\overline{G}_m(x))}{(1-\alpha)\overline{F}_n(y) + \alpha\bar{G}_m(y)} - \overline{F}(x) \right) : 0 \leq x < t_F \right\}$$

converges weakly to $\{L_\alpha(x)\colon 0 \leq x < t_F\}$ *where*

$$L_\alpha(x) := -\sup_{t \in \mathscr{A}(x)} \frac{(1-\alpha)tW_1(F(x)) + \alpha\overline{G}(x)W_1(1-t)}{(1-\alpha)t + \alpha t(\overline{F}(x))^{-1}\overline{G}(x)}. \tag{3.4}$$

(ii) *If* $m/n \to c$, *where* $0 < c < \infty$, *then*

$$\left\{ n^{1/2} \inf_{y:\ 0 \leq y \leq x} \left(\frac{\overline{F}_n(y)((1-\alpha)\overline{F}_n(x) + \alpha\overline{G}_m(x))}{(1-\alpha)\overline{F}_n(y) + \alpha\bar{G}_m(y)} - \overline{F}(x) \right) : 0 \leq x < t_F \right\} \tag{3.5}$$

converges weakly to $\{L_{\alpha,c}(x)\colon 0 \leq x < t_F\}$ *where*

$$L_{\alpha,c}(x) := -\sup_{t \in \mathscr{A}(x)} \frac{S(t,\alpha,x)}{(1-\alpha)t + \alpha t(\overline{F}(x))^{-1}\overline{G}(x)},$$

with $S(t, \alpha, x)$ *given by*

$$\begin{aligned}(1-\alpha)tW_1(F(x)) + \alpha\overline{G}(x)W_1(1-t) + c^{-1/2}\alpha t W_2(G(x))\\ - c^{-1/2}\alpha\overline{F}(x)W_2(1 - t\overline{G}(x)(\overline{F}(x))^{-1}).\end{aligned}$$

(iii) *If* $m/n \to 0$, *then*

$$\left\{ m^{1/2} \inf_{y:\ 0 \leq y \leq x} \left(\frac{\overline{F}_n(y)((1-\alpha)\overline{F}_n(x) + \alpha\overline{G}_m(x))}{(1-\alpha)\overline{F}_n(y) + \alpha\overline{G}_m(y)} - \overline{F}(x) \right) : 0 \leq x < t_F \right\}$$

converges weakly to $\{L_{\alpha,0}(x)\colon 0 \leq x < t_F\}$ *where*

$$L_{\alpha,0}(x) := -\sup_{t \in \mathscr{A}(x)} \frac{\alpha t W_2(G(x)) - \alpha\overline{F}(x)W_2(t\overline{G}(x)(\overline{F}(x))^{-1})}{(1-\alpha)t + \alpha t(\overline{F}(x))^{-1}\overline{G}(x)}. \tag{3.6}$$

In the case $m/n \to \infty$, we are able to estimate $\overline{G}$ very well, and the asymptotic distribution is the same as in the case when $\overline{G}$ is known. Thus, the comments in Section 2 apply to this case. In the case $m/n \to 0$, with $\alpha > 0$, the rate of convergence is $m^{1/2}$ which is a slower rate than the rate of convergence of the empirical survival function. In this situation, we cannot estimate $\overline{G}$ with sufficient precision, and the information that H in nonincreasing is not very useful. The empirical survival function is the more reliable estimator of F in such situations.

When $m/n \to c$, with $0 < c < \infty$, the best choice of estimator depends on c via an analysis similar to that in Section 2.

THEOREM 3.2. *Suppose that $m/n \to c \in (0, \infty)$ and that $\mathscr{A}(x)$ contains more than one element. Then, $E[(L_{\alpha,c}(x))^2]$ is minimized at*

$$\alpha_0 = \alpha_0(\overline{F}, \overline{G}) = \frac{(1 - a_1)\overline{F}(x)}{(1 - a_1 + c^{-1}(1 - 2a_1 + a_2))\overline{F}(x) + (a_2 - a_1)\overline{G}(x)}, \tag{3.7}$$

where

$$a_1 := \left(\frac{1}{\overline{F}(x)} - \frac{1}{m(x)}\right)^{-1} E\left[B\left(\frac{1}{\overline{F}(x)} - \frac{1}{m(x)}\right) \sup_{t \in \mathscr{A}(x)} B\left(\frac{1}{t} - \frac{1}{m(x)}\right)\right], \tag{3.8}$$

$$a_2 := \left(\frac{1}{\overline{F}(x)} - \frac{1}{m(x)}\right)^{-1} E\left[\left(\sup_{t \in \mathscr{A}(x)} B\left(\frac{1}{t} - \frac{1}{m(x)}\right)\right)^2\right], \tag{3.9}$$

and $m(x)$ is the largest element in $\mathscr{A}(x)$.

Moreover, for this choice of α,

$$E[(L_{\alpha_0,c}(x))^2] = \overline{F}(x)\left(A + \frac{\overline{F}(x)}{m(x)}(1 - A)\right) - (\overline{F}(x))^2,$$

where

$$0 \le A := \frac{\overline{F}(x)(1 + a_2 - 2a_1) + c\overline{G}(x)(a_2 - a_1^2)}{(1 + a_2 - 2a_1)(\overline{F}(x) + c\overline{G}(x))} \le 1.$$

Since $\overline{F}(x)$, $\overline{G}(x)$ and c are unknown, it is natural to estimate them by $\overline{F}_n(x)$, $\overline{G}_m(x)$ and m/n, respectively. In particular, the best choice for α is well approximated by

$$\hat{\alpha}_0(\overline{F}_n, \overline{G}_m) = \frac{(1 - a_1)\overline{F}_n(x)}{(1 - a_1 + nm^{-1}(1 - 2a_1 + a_2))\overline{F}_n(x) + \overline{G}_m(x)(a_2 - a_1)},$$

and the asymptotic behavior of the estimator is unaffected by this substitution. It is easy to see that $E[(L_{\alpha_0,c}(x))^2]$ decreases with c. This corresponds to the intuitive notion that the precision in estimating G increases with m.

THEOREM 3.3. *Suppose that $m/n \to c \in (0, \infty)$, $l(x) \ne x$ and $\mathscr{A}(x) = [\overline{F}(x), \overline{F}(l(x))]$. Then, $E[(L_{\alpha,c}(x))^2]$ is minimized at*

$$\alpha_0 = \alpha_0(\overline{F}, \overline{G}) = \frac{c\overline{F}(x)}{(2 + c)\overline{F}(x) + c\overline{G}(x)}$$

and

$$E[(L_{\alpha_0,c}(x))^2] = \overline{F}(x)\left(A + \frac{\overline{F}(x)}{m(x)}(1 - A)\right) - (\overline{F}(x))^2,$$

where

$$A = \frac{2(\overline{F}(x))^2 + 7\overline{F}(x)c\overline{G}(x) + 3c^2(\overline{G}(x))^2}{4(\overline{F}(x) + c\overline{G}(x))^2}.$$

The estimator of $\overline{F}$ *with this choice* α_0 *is*

$$\overline{F}_{\hat{\alpha}_0,n}(x) = \inf_{y:\ 0\le y\le x} \frac{\overline{F}_n(y)((2n\overline{F}_n(x)+m\overline{G}_m(x))\overline{F}_n(x)+m\overline{F}_n(x)\overline{G}_m(x))}{(2n\overline{F}_n(x)+m\overline{G}_m(x))\overline{F}_n(y)+m\overline{F}_n(x)\overline{G}_m(y)},$$

where $\hat{\alpha}_0 = \alpha_0(\overline{F}_n, \overline{G}_m)$.

Next, we consider two cases in which both $\overline{G}$ and $\overline{F}$ are discrete distributions. The proofs of these results are similar to those of Theorems 2.8 and 2.10 and are omitted.

THEOREM 3.4. *Suppose that* $m/n \to c \in (0,\infty)$ *and that* $\mathscr{A}(x) = \{\overline{F}(x), m(x)\}$ *for some* $m(x) \ne \overline{F}(x)$, *then* $E[(L_{\alpha,c}(x))^2]$ *is minimized at* $\alpha_0(c) = c/(c+1)$, *and*

$$E[(L_{\alpha_0,c}(x))^2] = \overline{F}(x)\left(A + \frac{\overline{F}(x)}{m(x)}(1-A)\right) - (\overline{F}(x))^2,$$

where

$$A = \frac{\overline{F}(x)+(3/2)c\overline{G}(x)}{\overline{F}(x)+c\overline{G}(x)}.$$

The estimator of $\overline{F}$ *with this choice* α_0 *is*

$$\widehat{\overline{F}}_{\hat{\alpha}_0,n}(x) = \inf_{y:\ 0\le y\le x} \frac{\overline{F}_n(y)(n\overline{F}_n(x)+m\overline{G}_m(x))}{n\overline{F}_n(y)+m\overline{G}_m(y)}, \tag{3.10}$$

where $\hat{\alpha}_0 = \alpha_0(m/n)$.

REMARK. Theorem 3.4 identifies sufficient conditions for Mukerjee's recommended α in (3.3) to be asymptotically optimal. Since $m/n \to c$, we have

$$\lim_{n\to\infty} \frac{m}{m+n} = \lim_{n\to\infty} \frac{m/n}{m/n+1} = \frac{c}{c+1} = \alpha_0.$$

THEOREM 3.5. *Suppose that* $\mathscr{A}(x) = \{\overline{F}(x), y_1, y_2\}$ *where* $\overline{F}(x) < y_1 < y_2$, *then* $E[(L_{\alpha,c}(x))^2]$ *is minimized at*

$$\alpha_0 = \frac{(1-a_1)\overline{F}(x)}{(1-a_1+c^{-1}(1-2a_1+a_2))\overline{F}(x)+\overline{G}(x)(a_2-a_1)},$$

where

$$a_1 = 2^{-1} + t2^{-2} - (2\pi)^{-1}\arctan((t/(1-t))^{1/2})$$

and

$$a_2 = 2^{-1} + t2^{-2} + (2\pi)^{-1}(t(1-t))^{1/2} - (2\pi)^{-1}\arctan((t/(1-t))^{1/2}),$$

with

$$t = \frac{1/y_1 - 1/y_2}{1/\overline{F}(x) - 1/y_2}.$$

Next, we present the asymptotics of the estimator in (3.2). Since the proofs of these results are similar to those of the asymptotics of (3.1), they are omitted. The limit distribution will be given in terms of $\{W_1(t)\colon 0 \le t \le 1\}$ and $\{W_2(t)\colon 0 \le t \le 1\}$. Since in both cases, the limit distributions are obtained from

$$\{n^{1/2}(F_n(x) - F(x))\colon x \in \mathbb{R}\} \xrightarrow{w} \{W_1(F(x))\colon x \in R\}$$

and

$$\{m^{1/2}(G_m(x) - G(x))\colon x \in \mathbb{R}\} \xrightarrow{w} \{W_2(G(x))\colon x \in R\},$$

the processes in (3.1) and (3.2) converge jointly.

THEOREM 3.6. *Under the conditions of Theorem* 3.1 :

(i) *If* $m/n \to \infty$, *then*

$$\left\{ m^{1/2} \sup_{y\colon 0 \le y \le x} \left(\frac{\overline{G}_m(y)((1-\alpha)\overline{F}_n(x) + \alpha\overline{G}_m(x))}{(1-\alpha)\overline{F}_n(y) + \alpha\overline{G}_m(y)} - \overline{G}(x) \right) : 0 \le x < t_F \right\}$$

converges weakly to $\{L^*_{\alpha,\infty}(x)\colon 0 \le x < t_F\}$ *where*

$$L^*_{\alpha,\infty}(x) := -\inf_{t \in \mathscr{A}^*(x)} \frac{\alpha t W_2(G(x)) + (1-\alpha)\overline{F}(x) W_2(1-t)}{\alpha t + (1-\alpha) t (\overline{G}(x))^{-1} \overline{F}(x)}$$

and

$$\mathscr{A}^*(x) = \{\overline{G}(y)\colon y \le x, H(y) = H(x)\} \cup \{\overline{G}(y-)\colon y \le x, H(y-) = H(x)\}.$$

(ii) *If* $m/n \to c$, *for some* $0 < c < \infty$, *then*

$$\left\{ m^{1/2} \sup_{y\colon 0 \le y \le x} \left(\frac{\overline{G}_m(y)((1-\alpha)\overline{F}_n(x) + \alpha\overline{G}_m(x))}{(1-\alpha)\overline{F}_n(y) + \alpha\overline{G}_m(y)} - \overline{G}(x) \right) : 0 \le x < t_F \right\}$$

converges weakly to $\{L^*_{\alpha,c}(x)\colon 0 \le x < t_F\}$ *where*

$$L^*_{\alpha,c}(x) := -\inf_{t \in \mathscr{A}_{\overline{G}}(x)} \frac{S^*(t,\alpha,x)}{\alpha t + (1-\alpha) t (\overline{G}(x))^{-1} \overline{F}(x)},$$

with $S^*(t, \alpha, x)$ *given by*

$$\alpha t W_2(G(x)) + (1-\alpha)\overline{F}(x) W_2(1-t) + c^{1/2}(1-\alpha) t W_1(F(x)) \\ - c^{1/2}(1-\alpha)\overline{G}(x) W_1(1 - t\overline{F}(x)(\overline{G}(x))^{-1}).$$

(iii) *If $m/n \to 0$, then*

$$\left\{ n^{1/2} \sup_{y:\ 0 \le y \le x} \left(\frac{\overline{G}_m(y)((1-\alpha)\overline{F}_n(x) + \alpha \overline{G}_m(x))}{(1-\alpha)\overline{F}_n(y) + \alpha \overline{G}_m(y)} - \overline{G}(x) \right) : 0 \le x < t_F \right\}$$

converges weakly to $\{L^*_{\alpha,0}(x): 0 \le x < t_F\}$ *where*

$$L^*_{\alpha,0}(x) := - \inf_{t \in \mathscr{A}_{\overline{G}}(x)} \frac{(1-\alpha)tW_1(F(x)) - (1-\alpha)\overline{G}(x)W_1(1 - t\overline{F}(x)(\overline{G}(x))^{-1})}{\alpha t + (1-\alpha)t(\overline{G}(x))^{-1}\overline{F}(x)}.$$

REMARK. Suppose that $m/n \to c$, where $0 < c < \infty$. The result above is proved by considering the cases in Theorem 2.2. In Cases 1 and 2 of the proof of Theorem 2.2, the limit distributions of (3.1) and (3.2) are the same as the limit distributions of $n^{1/2}(F_n(x) - F(x))$ and $m^{1/2}(G_m(x) - G(x))$, so the joint distribution of (3.1) and (3.2) is that of a multivariate normal random vector with independent components. In other cases, the limiting process is not normal and has correlated components.

THEOREM 3.7. *Suppose that $\mathscr{A}^*(x)$ contains more than one element. Then, $E[(L_{\alpha,c}(x))^2]$ is minimized at*

$$\alpha_0 = \frac{c(1 - 2a_1 + a_2)\overline{G}(x) + (a_2 - a_1)\overline{F}(x)}{(1 - a_1 + c(1 - 2a_1 + a_2))\overline{G}(x) + (a_2 - a_1)\overline{F}(x)}.$$

Moreover, for this choice of α,

$$E[(L_{\alpha_0,c}(x))^2] = \overline{G}(x)\left(A + \frac{\overline{G}(x)}{m(x)}(1 - A) \right) - (\overline{G}(x))^2,$$

where

$$0 \le A := \frac{c\overline{G}(x)(1 + a_2 - 2a_1) + \overline{F}(x)(a_2 - a_1^2)}{(1 + a_2 - 2a_1)(\overline{F}(x) + c\overline{G}(x))} \le 1.$$

and $m(x)$ is the largest member of the set $\mathscr{A}^*(x)$.

Although the estimators, in (3.1) and (3.2) satisfy, $\widehat{F}_{\alpha,n} \le_+ \widehat{G}_{\alpha,m}$, for each α, this USO condition is not satisfied for the estimators featured in Theorems 3.3 and 3.6. When the choice of alpha is allowed to vary with x and take on different values within $\hat{F}$ and $\hat{G}$, the ordering of the resulting estimators is sacrificed. These estimators each satisfy the USO constraint with respect to a consistent, though suboptimal, estimator of the complementary distribution function. For example, if $\alpha(x)$ is chosen minimizing the asymptotic MSE of $\widehat{F}_{\alpha,n}$, then $\widehat{F}_{\alpha(x),n} \le_+ \widehat{G}_{\alpha(x),m}$.

4. Discussion. The domination of the hazard rate of one population of items by the hazard rate of another or, more generally, the uniform stochastic ordering between two lifetime distributions, is often the most natural way to model the superiority of one population over another in a given reliability study. Two recent investigations have provided approaches to the estimation of underlying survival functions under such ordering constraints. While these studies have produced closed form estimators with good fixed-sample size and asymptotic properties, neither provided insight into the theoretical variability of the estimators. The purpose of the present study is to fill this gap by providing a complete asymptotic distribution theory for the estimators. The achievement of this goal will facilitate the development of confidence bands for the true survival curves based on these estimators and has also served the important purpose of facilitating the comparison of competing estimators in terms of their asymptotic precision.

In Section 2, the asymptotic distribution of the estimator proposed by Rojo and Samaniego (1993) is derived, as is that of the one-sample version of the class of estimators introduced by Mukerjee (1996). Among the notable findings of that section is the demonstration of the asymptotic inadmissibility of the empirical survival curve under a squared error criterion, the identification of conditions under which the Rojo–Samaniego estimator is optimal within the Mukerjee class and the identification of conditions under which alternative estimators within the Mukerjee class are optimal.

With the asymptotic distributions of estimators in (1.4) and (3.1) in hand, we are now able to formulate confidence statements regarding the underlying distribution F based on the available (one or two) samples. The argument for using these estimators in place of the empirical survival function (esf) can be made on either of two bases, one logical and one practical. The logical basis stems from the fact that, when a postulated model satisfies a known constraint, one should estimate that model from within the constrained class. For instance, when one is satisfied that the available data is derived from a normal population, it would seem unreasonable to estimate the density governing the data as a double exponential.

On practical grounds, one can justify the constrained estimators in this paper on the basis of their relative performance. We have shown that the asymptotic performance of these estimators is never worse than that of the esf, and is superior to it in certain specific circumstances. In addition, evidence is presented in Rojo and Samaniego (1993) and in Mukerjee (1996) that, even in cases in which their estimators are asymptotically equivalent to the esf, the constrained estimators have uniformly smaller mean squared errors in an array of fixed-sample size problems. Given this, it is plausible to suggest that approximate confidence bands in the latter cases using the identical asymptotic variances will tend to yield conservative bands when based on the constrained estimator, that is, will provide, in conjunction with the constrained estimator, confidence bands with coverage probability exceeding the nominal level. This would provide good reason, even in the presence of asymptotic equivalence, to favor confidence statements based on the constrained estimator. Of course the strongest argument for the constrained estimates and their associated confidence bands arises in cases (for example, where F and G are discrete distributions) in which the bands around the constrained estimators are narrower that those around the esf. While the examples below are not of that type, they do illustrate that, having settled on a specific constrained estimator for $\overline{F}_n$ in the one- or two-sample problem, one can, using results developed here, complete the inference concerning $\overline{F}_n$ by factoring in an estimate of variability.

Rojo and Samaniego (1993) employed their one-sample estimator in analyzing the results of an accelerated life testing experiment in which a sample of 39 failure times of Kevlar/Epoxy Pressure Vessels under “86% stress” were assumed to be drawn from a distribution F satisfying the constraint $F <_{(+)} \mathrm{Wei}(0.875, 0.00127)$, where $\mathrm{Wei}(\alpha, \lambda)$ represents the Weibull distribution with shape parameter α and scale parameter $1/\lambda$. These data, drawn from Barlow, Toland and Freeman (1988), are displayed in Table 1.

Using the fact that for continuous distributions F and G and decreasing H, the Rojo–Samaniego estimator of $F(x)$ has standard error $(F(x)(1 - F(x))/n)^{1/2}$ (see the proof of Theorem 2.2, case 1), a 95 confidence band for $\overline{F}$ is displayed in Figure 1.

The assumption of known G in the example above might well be considered heroic. As explained in Rojo and Samaniego, the assumption was motivated by the fact that the particular Weibull curve used above was the model fitted to data from an auxiliary test at a lower stress level and thus represented a plausible approximation of a dominating G. As an alternative to an assumed known G, we have utilized the Mukerjee two-sample estimator for $\overline{F}$ with the 24 failure times from the auxiliary experiment serving as the second sample. These latter data, also from Barlow, Toland and Freeman (1988), are displayed in Table 2.

The Mukerjee estimator $\hat{\hat{\overline{F}}}_{\hat{\alpha}_{0,n}}$ of (3.1), with $\hat{\alpha} = 24/(24 + 39) = 0.381$, is shown in Figure 2, together with an approximate 95% confidence band for $\overline{F}$. The band utilizes the standard error $(F(x)(1 - F(x))/n)^{1/2}$ for $\hat{\hat{\overline{F}}}_{\hat{\alpha}_{0,n}}$ [see Theorem 3.1, part (ii), case 1], applicable for continuous F and G and decreasing H.

Figures 1 and 2 show that the estimators $\hat{\hat{\overline{F}}}_n$ and $\hat{\hat{\overline{F}}}_{\hat{\alpha}_{0,n}}$ differ rather little over the major portion of the range of observed failure times; for $x \in [300, 600]$; however, $\hat{\hat{\overline{F}}}_{\hat{\alpha}_{0,n}}(x)$ is noticeable larger than $\hat{\hat{\overline{F}}}_n(x)$.

We turn to a brief discussion of the natural extension of our results to the treatment of censored data. For simplicity, we confine our remarks to the

TABLE 1

Failure times (in hours) of Kevlar/Epoxy pressure vessels at 86% stress

2.2	8.5	18.7	111.4	755.2	4.0	9.1	22.1	144.0	952.2
4.0	10.2	45.9	158.7	1108.2	4.6	12.5	55.4	243.9	1148.5
6.1	13.3	61.2	254.1	1569.3	6.7	14.0	87.5	444.4	1750.6
7.9	14.6	98.2	590.4	1802.1	8.3	15.0	101.0	638.2	

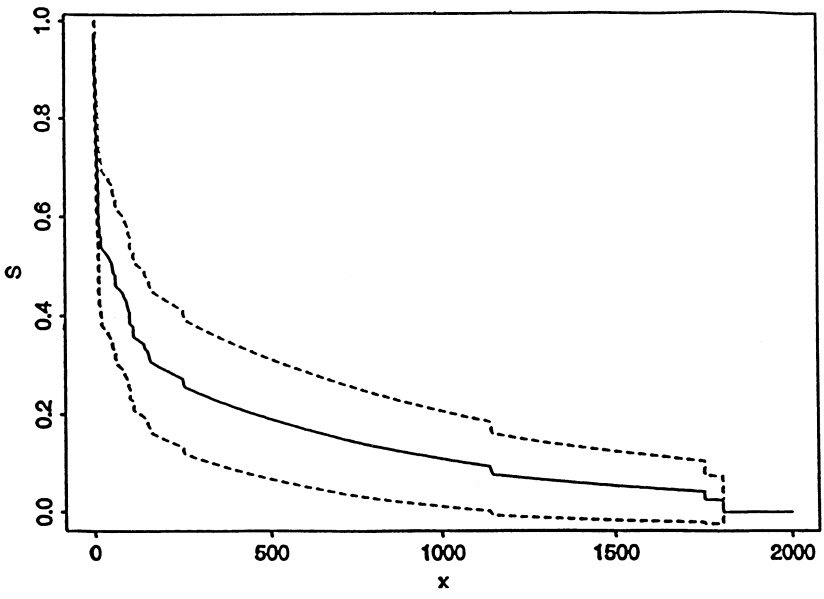

FIG. 1. *Estimated survival function* $\widehat{\overline{F}}_n$ *and approximate* 95% *confidence band for* $\overline{F}$.

one-sample problem. Suppose we observe $(Z_1, \delta_1), \dots, (Z_n, \delta_n)$ where

$$Z_i = \min(X_i, Y_i), \qquad \delta_i = I\{X_i \le Y_i\},$$

and $X_1, \dots, X_n \overset{iid}{\sim} F$ and $Y_1, \dots, Y_n \overset{iid}{\sim} K$ are independent samples from a lifetime distribution F and a censoring distribution K, respectively. Suppose further that $F \le_{(+)} G$, where G is known. If S_n represents the Kaplan–Meier estimator of $\overline{F}$, then it is well known [see, e.g., Breslow and Crowley (1974), Gill (1981)] that

$$\sup_{t>0} \left| S_n(t) - \overline{F}(t) \right| \to 0 \quad \text{a.s.}$$

and that $\{n^{1/2}(S_n(t) - \overline{F}(t)): t \ge 0\}$ converges weakly to a Gaussian process $\{Z(t): t \ge 0\}$ with mean zero and covariance

$$E\left(Z(s)Z(t)\right) = C(s)\overline{F}(s)\overline{F}(t),$$

TABLE 2

Failure times (in hours) of Kevlar/Epoxy pressure vessels at 80% *stress*

19.1	199.1	514.2	694.1	1536.8	24.3	403.7	541.6	876.7	1755.5
69.8	432.2	544.9	930.4	2046.2	71.2	453.4	554.2	1254.9	6177.5
136.0	514.1	664.5	1275.6						

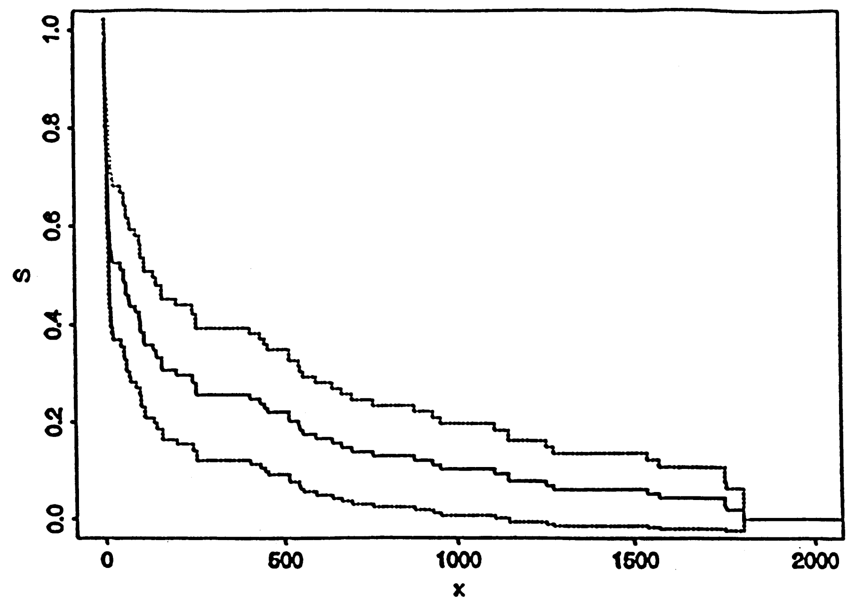

FIG. 2. *Estimated survival function* $\widehat{\overline{F}}_{\alpha_{0,n}}$ *and approximate* 95% *confidence band for* $\overline{F}$.

where $s < t$ and

$$C(s) = \int_0^s \frac{1}{\overline{F}^2(t)K(t)}\, dF(t).$$

Thus, the same arguments made in proving Theorem 2.2 apply when F and G are continuous. In particular, we have that

$$\left\{ n^{1/2} \inf_{y:\ 0\le y\le x} \left(\frac{S_n(y)\big((1-\alpha)S_n(x) + \alpha\overline{G}(x)\big)}{(1-\alpha)S_n(y) + \alpha\overline{G}(y)} - \overline{F}(x) \right) : 0 \le x < t_F \right\}$$

converges weakly to $\{L_\alpha(x)\colon 0 < x < t_F\}$ where

$$L_\alpha(x) = \inf_{y:\ \ell(x)\le y\le x} \frac{(1-\alpha)\overline{F}(x)Z(x) + \alpha\overline{G}(x)Z(y)}{(1-\alpha)\overline{F}(y) + \alpha\overline{G}(y)}$$

and $\ell(x) = \inf\{y\colon H(y) = H(x)\}$.

Rojo (1998) considers asymptotics of estimators similar to those in the present paper in the context of estimating the quantile function F^{-1} under the restriction that $F^{-1}(x)/G^{-1}(x)$ is nonincreasing on $(0, 1)$. In his Theorem 7, he shows that the estimator is asymptotically equivalent to the empirical distribution function, assuming that $F^{-1}(x)/G^{-1}(x)$ is strictly decreasing and F is absolutely continuous. This result parallels our result under comparable conditions on F, G and H (though we eschew the absolute continuity assumption). We believe the methods in the present paper can be employed to provide more general asymptotics for the estimators in Rojo (1998).

Finally, we consider briefly how the asymptotics developed in this paper may be used to address testing questions regarding the constraint $F \le_{(+)} G$. We first examine the one sample problem: let $X_1, \dots, X_n$ be i.i.d. r.v.'s with unknown distribution F, and let G be a known distribution. We wish to test H_0: $F \le_{(+)} G$ versus H_1: $F \not\le_{(+)} G$. Let $D = \sup_{y \le x}(\overline{G}(y)\overline{F}(x) - \overline{G}(x)\overline{F}(y))$. It is easy to see that $D = 0$ iff $F \le_{(+)} G$. A natural statistic for this testing problem is

$$D_n = \sup_{0 \le y \le x} (\overline{G}(y)\overline{F}_n(x) - \overline{G}(x)\overline{F}_n(y)).$$

An elementary computation yields

$$\left|D_n - D\right| \le 2 \sup_{x \ge 0} \left|\overline{F}_n(x) - \overline{F}(x)\right|.$$

Thus, if $F \le_{(+)} G$, then $D_n \to 0$ a.s.; on the other hand, if $F \not\le_{(+)} G$, then $D_n \to D > 0$ almost surely. The methods of Theorem 2.2 will show that if $F \le_{(+)} G$, then

$$n^{1/2} D_n \xrightarrow{d} \sup_{x \ge 0} \sup_{t \in \mathscr{A}(x)} (\overline{G}(x)W(1-t) - t\overline{G}(x)(\overline{F}(x))^{-1}W(F(x))).$$

It is easy to see that

$$0 \le \sup_{x \ge 0} \sup_{t \in \mathscr{A}(x)} (\bar{G}(x)W(1-t) - t\overline{G}(x)(\overline{F}(x))^{-1}W(F(x))) \le 2 \sup_{0 \le x \le 1} |W(x)|.$$

Given $0 < \alpha < 1$, take $c_{\alpha/2}$ such that $\Pr\{\sup_{0 \le x \le 1} |W(x)| \ge c_{\alpha/2}\} = \alpha$. We propose the test that rejects H_0 if $n^{1/2} D_n \ge c_{\alpha/2}$. This test has asymptotic level α: if $F \le_{(+)} G$,

$$\lim_{n \to \infty} \Pr\{n^{1/2} D_n \ge c_{\alpha/2}\} \le \alpha.$$

It also has asymptotic power one: if $F \not\le_{(+)} G$,

$$\lim_{n \to \infty} \Pr\{n^{1/2} D_n \ge c_{\alpha/2}\} = 1.$$

Trivial variations of the argument above give a test when two samples, one for F and another for G, are available. We omit the details.

Another problem of interest is that of testing the equality of two distributions. Let $X_1, \dots, X_n$ be i.i.d. r.v.'s with unknown df F. Let $Y_1, \dots, Y_m$ be i.i.d. r.v.'s with unknown df G. Suppose that $F \le_{(+)} G$. The problem is to test H_0: $F = G$ versus H_1: $F \ne G$. The common test for this situation (ignoring the stochastic ordering assumption) is the Kolgomorov–Smirnov test. This is based on the fact that, under the null hypothesis,

$$\sup_{x \ge 0} \sqrt{\frac{mn}{m+n}} \left|\overline{F}_n(x) - \overline{G}_m(x)\right| \xrightarrow{d} \sup_{x \ge 0} \left|W(F(x))\right|.$$

Consider

$$D_{n,m} = \sup_{x\ge 0}\left|\overline{G}_m(x) - \inf_{y:\ 0\le y\le x}\frac{\overline{G}_m(x)\overline{F}_n(y)}{\overline{G}_m(y)}\right| = \sup_{0\le y\le x}\left(\overline{G}_m(x) - \frac{\overline{G}_m(x)\overline{F}_n(y)}{\overline{G}_m(y)}\right).$$

Let $D = \sup_{x\ge 0}\left|\overline{F}(x) - \overline{G}(x)\right|$. Obviously, we have that $D = 0$ iff $F = G$ and that $D_{n,m} \to D$ almost surely. The methods in the proof of Theorem 3.2 give that if $F = G$, then

$$\sqrt{\frac{mn}{m+n}}D_{n,m} \overset{d}{\to} \sup_{x\ge 0}\left|\sup_{t\in\mathscr{A}(x)}\frac{\overline{F}(x)}{t}W(1-t)\right|,$$

where $\mathscr{A}(x)$ is the set defined when $G = F$. It is easy to see that

$$\sup_{x\ge 0}\left|\sup_{t\in\mathscr{A}(x)}\frac{\overline{F}(x)}{t}W(1-t)\right| \le \sup_{x\ge 0}\ \sup_{t\in\mathscr{A}(x)}\frac{\overline{F}(x)}{t}W(1-t) \le \sup_{0\le u\le 1} W(u),$$

where equality holds for continuous distributions. It is well known [see, e.g., page 85 in Billinsgley (1968)] that $\Pr\{\sup_{0\le u\le 1} W(u) \ge \lambda\} = \exp(-2\lambda^2)$. Thus, the test that rejects H_0 if $\sqrt{mn/(m+n)}D_{n,m} \ge \sqrt{2^{-1}\log\alpha^{-1}}$ has asymptotic level α: if $F = G$,

$$\limsup_{n\to\infty}\Pr\left\{\sqrt{\frac{mn}{m+n}}D_{n,m} \ge \sqrt{2^{-1}\log\alpha^{-1}}\right\} \le \alpha.$$

Of course, the test also has asymptotic power one: if $F \ne G$ (but $F \le_{(+)} G$),

$$\lim_{n\to\infty}\Pr\left\{\sqrt{\frac{mn}{m+n}}D_{n,m} \ge \sqrt{2^{-1}\log\alpha^{-1}}\right\} = 1.$$

The cut-off points of this test are smaller than those of the Kolmogorov–Smirnov test. Instead of using the tails of $\sup_{0\le x\le 1}|W(x)|$, we use the tails of $\sup_{0\le x\le 1} W(x)$. The latter testing problem has been considered by Dykstra, Kochar and Robertson (1991) using the methods of isotonic regression; no asymptotic distribution theory is provided for the test they propose.

APPENDIX

We will make use of the following lemmas:

LEMMA A.1. *Consider the function* $f(x) = (c(1-x)^2 + 2dx(1-x) + ex^2)/(a(1-x)+bx)^2$, *for* $x > 0$, *where* $b > a > 0$. *Suppose that* $a(e-d) > b(d-c)$ *and* $bc > ad$. *Then* f *is minimized over* $[0,\infty)$ *at* $x = (bc-ad)/a(e-d) + b(c-d)$.

PROOF. Since $f'(x) = (2(a(e-d)+b(c-d))x - 2(bc-ad))/(a(1-x)+bx)^3$, the claim is transparent. □

LEMMA A.2. *Let X and Y be two independent standard normal r.v.'s and let $a > 0$; then*

$$E[X \max(0, X, aY)] = 2^{-1} - (2\pi)^{-1} \arctan(a)$$

and

$$E[(\max(0, X, aY))^2] = 2^{-1} + 2^{-2}a^2 + a(2\pi)^{-1} + (a^2 - 1)(2\pi)^{-1} \arctan(a).$$

The proof of Lemma A.2 is omitted since it is a simple calculus exercise.

PROOF OF THEOREM 2.1. The truth of the claim in case (i) is self-evident. As for case (ii), we consider two subcases. Assume first that $0 < \overline{F}(t_F-)$ and $\overline{G}(t_F-) < 1$. Since $0 = \overline{F}(t_F) < \overline{F}(t_F-)$, it follows that $\overline{F}_n(t_F-) - \overline{F}_n(t_F) = 1$ with positive probability. If $\overline{F}_n(t_F-) - \overline{F}_n(t_F) = 1$, then

$$\sup_{x>0} \left|\overline{F}_n(x) - \overline{F}(x)\right| = 1 - \overline{F}(t_F-)$$

and

$$\sup_{0\le x<t_F} \left| \inf_{y:\ 0\le y\le x} \left(\frac{\overline{G}(x)\overline{F}_n(y)}{\overline{G}(y)} - \overline{F}(x) \right) \right| = \sup_{0\le x<t_F} (\overline{G}(x) - \overline{F}(x)).$$

Since $\overline{G}(t_F-) < 1$,

$$\sup_{0\le x<t_F} (\overline{G}(x) - \overline{F}(x)) < 1 - \overline{F}(t_F-).$$

Hence, the claim follows in this subcase.

Assume now that $\overline{F}(t_F-) = 0$ and $\overline{G}(t_F-) < 1$. Let $\delta = 1 - \sup_{0\le x<t_F} (\overline{G}(x) - \overline{F}(x))$ and take $x_0 < t_F$ such that $0 < \overline{F}(x_0) < \min(2^{-1}, \delta)$. Now, $\Pr\{\overline{F}_n(x_0) = 1\} = (\overline{F}(x_0))^n > 0$. Assume that $\overline{F}_n(x_0) = 1$, and let $X_{(1)} = \min_{1\le i\le n} X_i$. For $0 \le x < X_{(1)}$,

$$\left| \inf_{y:\ 0\le y\le x} \left(\frac{\overline{G}(x)\overline{F}_n(y)}{\overline{G}(y)} - \overline{F}(x) \right) \right| = |\overline{G}(x) - \overline{F}(x)| \le 1 - \delta.$$

For $x \ge X_{(1)} > x_0$, we have

$$-\overline{F}(x_0) \le -\overline{F}(x) \le \inf_{y:\ 0\le y\le x} \left(\frac{\overline{G}(x)\overline{F}_n(y)}{\overline{G}(y)} - \overline{F}(x) \right) \le \overline{G}(x) - \overline{F}(x) \le 1 - \delta.$$

Combining the bounds above, we obtain

$$\sup_{0\le x<t_F} \left| \inf_{y:\ 0\le y\le x} \left(\frac{\overline{G}(x)\overline{F}_n(y)}{\overline{G}(y)} - \overline{F}(x) \right) \right| \le \max(1 - \delta, \overline{F}(x_0)) < 1 - \overline{F}(x_0).$$

Since $x_0 < X_{(1)}$, it follows that $\sup_{x>0} |\overline{F}_n(x) - \overline{F}(x)| \ge 1 - \overline{F}(X_{(1)}-) \ge 1 - \overline{F}(x_0)$, establishing the claim in the second subcase. □

PROOF OF THEOREM 2.2. First, we prove convergence for a fixed $x \in [0, t_F)$. Let $H(x)$, $\ell(x)$ and $\mathscr{A}(x)$ be as in (2.2)–(2.4), and let $H_n(y) = \overline{F}_n(y)/\overline{G}(y)$.

We consider six mutually exclusive and exhaustive cases.

Case 1. $l(x) = x, \bar{H}(x-) = \bar{H}(x)$ *and* $\overline{F}(x-) = \overline{F}(x)$. In this case, $\mathscr{A}(x) = \{\overline{F}(x)\}$. Let $\theta_n = \inf\{t\colon H_n(t) \le \inf_{y:\ 0\le y\le x} H_n(y) + n^{-1}\}$. Since H_n is right continuous, $H_n(\theta_n) \le \inf_{y:\ 0\le y\le x} H_n(y) + n^{-1}$. By an elementary inequality and the Glivenko–Cantelli theorem, we have that for each $t_F > z \ge 0$,

$$\left|\inf_{y:\ 0\le y\le z} H_n(y) - \inf_{y:\ 0\le y\le z} H(y)\right| \le \sup_{y:\ 0\le y\le z} \frac{|\overline{F}_n(y) - \overline{F}(y)|}{\overline{G}(y)} \longrightarrow 0 \quad \text{a.s.} \tag{A.1}$$

Thus, for each $\delta > 0$, we have, with probability 1, that

$$\lim_{n\to\infty} \inf_{y:\ 0\le y\le x-\delta} H_n(y) = H(x-\delta) > H(x) = \lim_{n\to\infty} H_n(x).$$

Thus, with probability 1, $\theta_n \to x$. Observe that by the definition of θ_n and the uniform stochastic ordering assumption,

$$\begin{aligned} -n^{-1/2} &\le n^{1/2}(H_n(x) - H_n(\theta_n)) \le n^{1/2}(H_n(x) - H_n(\theta_n) - H(x) + H(\theta_n)) \\ &= (\overline{G}(x))^{-1} n^{1/2}(\overline{F}_n(x) - \overline{F}_n(\theta_n) - \overline{F}(x) + \overline{F}(\theta_n)) \\ &\quad + \left((\overline{G}(x))^{-1} - (\overline{G}(\theta_n))^{-1}\right) n^{1/2}(\overline{F}_n(\theta_n) - \overline{F}(\theta_n)). \end{aligned} \tag{A.2}$$

By the continuity of $\overline{G}$ at x, $\overline{G}(\theta_n) \overset{\Pr}{\to} \overline{G}(x)$. Since $\overline{F}$ is continuous at x, for each $\eta > 0$,

$$\lim_{\delta\to 0} \limsup_{n\to\infty} \Pr\left\{\sup_{|t-x|\le\delta} n^{1/2}\left|\overline{F}_n(t) - \overline{F}(t) - (\overline{F}_n(x) - \overline{F}(x))\right| \ge \eta\right\} = 0.$$

So, $n^{1/2}(\overline{F}_n(\theta_n) - \overline{F}(\theta_n)) \overset{\Pr}{\to} 0$. Hence, the right-hand side in (A.2) converges to zero in probability. This implies that

$$n^{1/2}\left(\inf_{y:\ 0\le y\le x} \frac{\overline{G}(x)\overline{F}_n(y)}{\overline{G}(y)} - \overline{F}_n(x)\right) \overset{\Pr}{\to} 0$$

and the claim of the theorem follows.

Case 2. $l(x) = x$ *and* $\bar{H}(x-) > \bar{H}(x)$. In this case,

$$\lim_{n\to\infty} \inf_{y:\ 0\le y< x} H_n(y) = H(x-) > H(x) = \lim_{n\to\infty} H_n(x).$$

So, $\theta_n = x$ eventually w.p. 1 and the claim follows.

Case 3. $l(x) = x$, $H(x-) = H(x)$ *and* $\overline{F}(x-) > \overline{F}(x)$. In this case $\mathscr{A}(x) = \{\overline{F}(x), \overline{F}(x-)\}$. We claim that

$$n^{1/2} \inf_{y:\ 0\le y<x} \left(\frac{\overline{G}(x)\overline{F}_n(y)}{\overline{G}(y)} - \frac{\overline{G}(x)\overline{F}_n(x-)}{\overline{G}(x-)} \right) \overset{\Pr}{\to} 0. \tag{A.3}$$

Let $\theta_{n,-} = \inf\{t\colon H_n(t) \le \inf_{y:\ 0\le y<x} H_n(y) + n^{-1}\}$. Observe that $\theta_{n,-} < x$ and that $H_n(\theta_{n,-}) \le H_n(x-) + n^{-1}$. It is easy to see that, as in Case 1, $\theta_{n,-} \to x$. In this case, we have that

$$\begin{aligned} -n^{-1/2} &\le n^{1/2}(H_n(x-) - H_n(\theta_{n,-})) \\ &\le n^{1/2}(H_n(x-) - H_n(\theta_{n,-}) - H(x-) + H(\theta_{n,-})) \\ &= (\overline{G}(x-))^{-1} n^{1/2}(\overline{F}_n(x-) - \overline{F}_n(\theta_{n,-}) - \overline{F}(x-) + \overline{F}(\theta_{n,-})) \\ &\quad + \big((\overline{G}(x-))^{-1} - (\overline{G}(\theta_{n,-}))^{-1}\big)\, n^{1/2}(\bar{F}_n(\theta_{n,-}) - \overline{F}(\theta_{n,-})) \overset{\Pr}{\to} 0. \end{aligned}$$

This implies that

$$n^{1/2} \left(\inf_{y:\ 0\le y\le x} \frac{\overline{G}(x)\overline{F}_n(y)}{\overline{G}(y)} - \overline{F}_n(x) \right) \overset{\Pr}{\to} 0.$$

Observe that for any distribution, even a discrete one, for each $\eta > 0$,

$$\lim_{\delta\to 0} \limsup_{n\to\infty} \Pr\left\{ \sup_{t\in(x-\delta,x)} \left| n^{1/2}(\overline{F}_n(x-) - \overline{F}_n(t) - \overline{F}(x-) + \overline{F}(t)) \right| \ge \eta \right\} = 0.$$

So, (A.3) holds.

We also have that

$$\begin{aligned} & n^{1/2} \left(\inf_{y:\ 0\le y\le x} \frac{\overline{G}(x)\overline{F}_n(y)}{\overline{G}(y)} - \overline{F}(x) \right) \\ &\quad = n^{1/2} \min\left(\inf_{y:\ 0\le y<x} \frac{\overline{G}(x)\overline{F}_n(y)}{\overline{G}(y)} - \overline{F}(x), \overline{F}_n(x) - \overline{F}(x) \right) \\ &\quad = n^{1/2} \min\left(\inf_{y:\ 0\le y<x} \frac{\overline{G}(x)\overline{F}_n(y)}{\overline{G}(y)} - \frac{\overline{G}(x)\overline{F}_n(x-)}{\overline{G}(x-)} + \frac{\overline{G}(x)\overline{F}_n(x-)}{\overline{G}(x-)} \right. \\ &\qquad\qquad\qquad \left. - \overline{F}(x), \overline{F}_n(x) - \overline{F}(x) \right) \end{aligned}$$

Since $H(x) = H(x-)$, we get that

$$n^{1/2} \left(\frac{\overline{G}(x)\overline{F}_n(x-)}{\overline{G}(x-)} - \overline{F}(x) \right) = \frac{\overline{F}(x) n^{1/2}(\overline{F}_n(x-) - \overline{F}(x-))}{\overline{F}(x-)}.$$

So, the limit distribution in this case is

$$L(x) = \min\left(\frac{-\overline{F}(x)W(F(x-))}{\overline{F}(x-)}, -W(F(x))\right), \tag{A.4}$$

which is equivalent to (2.7) in the case under consideration.

Case 4. $l(x) < x$, $\overline{F}(l(x)-) = \overline{F}(l(x))$ *and* $\bar{H}(l(x)-) = \bar{H}(l(x))$. We have that

$$\begin{aligned} n^{1/2}&\left(\inf_{y:\ 0\le y\le x} \frac{\overline{G}(x)\overline{F}_n(y)}{\overline{G}(y)} - \overline{F}(x)\right) \\ &= n^{1/2}\min\left(\inf_{y:\ 0\le y\le l(x)} \frac{\overline{G}(x)\overline{F}_n(y)}{\overline{G}(y)} - \overline{F}(x), \inf_{y:\ l(x)\le y\le x} \frac{\overline{G}(x)\overline{F}_n(y)}{\overline{G}(y)} - \overline{F}(x)\right). \end{aligned}$$

In this case, the argument in the proof of Case 1 yields that

$$n^{1/2}\inf_{y:\ 0\le y\le l(x)}\left(\frac{\overline{G}(x)\overline{F}_n(y)}{\overline{G}(y)} - \frac{\overline{G}(x)\overline{F}_n(l(x))}{\overline{G}(l(x))}\right) \xrightarrow{\Pr} 0.$$

Since $H(x) = H(l(x))$, we get that

$$\begin{aligned} n^{1/2}&\inf_{y:\ 0\le y\le l(x)}\left(\frac{\overline{G}(x)\overline{F}_n(y)}{\overline{G}(y)} - \overline{F}(x)\right) \\ &= n^{1/2}\inf_{y:\ 0\le y\le l(x)}\left(\frac{\overline{G}(x)\overline{F}_n(y)}{\overline{G}(y)} - \frac{\overline{G}(x)\overline{F}_n(l(x))}{\overline{G}(l(x))} + \frac{\overline{G}(x)\overline{F}_n(l(x))}{\overline{G}(l(x))} - \overline{F}(x)\right) \\ &= n^{1/2}\inf_{y:\ 0\le y\le l(x)}\left(\frac{\overline{G}(x)\overline{F}_n(y)}{\overline{G}(y)} - \frac{\overline{G}(x)\overline{F}_n(l(x))}{\overline{G}(l(x))}\right) \\ &\quad + \frac{\overline{F}(x)n^{1/2}(\overline{F}_n(l(x)) - \overline{F}(l(x)))}{\overline{F}(l(x))}, \end{aligned}$$

which converges in distribution to $-\overline{F}(x)W(F(l(x)))/\overline{F}(l(x))$. Since H is a constant in $[l(x), x]$, we obtain that

$$n^{1/2}\inf_{y:\ l(x)\le y\le x}\left(\frac{\overline{G}(x)\overline{F}_n(y)}{\overline{G}(y)} - \overline{F}(x)\right) = \inf_{y:\ l(x)\le y\le x}\frac{\overline{F}(x)n^{1/2}(\overline{F}_n(y) - \overline{F}(y))}{\overline{F}(y)}$$

which converges in distribution to

$$L(x) = \inf_{y:\ l(x)\le y\le x}\frac{-\overline{F}(x)W(F(y))}{\overline{F}(y)}. \tag{A.5}$$

Case 5. $l(x) < x$ *and* $H(l(x)-) > \bar{H}(l(x))$. The argument in Case 2 gives that the limit distribution is the same one as in Case 4.

Case 6. $l(x) < x$, $H(l(x)-) = \bar{H}(l(x))$ *and* $\overline{F}(l(x)-) > \overline{F}(l(x))$. Again the same arguments imply that the limit distribution is

$$\text{(A.6)} \qquad L(x) = \min\left(\frac{-\overline{F}(x)W(F(l(x)-))}{\overline{F}(l(x)-)}, \inf_{y:\ l(x)\le y\le x} \frac{-\overline{F}(x)W(F(y))}{\overline{F}(y)} \right).$$

Now, we proceed to prove the weak convergence of the whole process. If $\bar{G}(t_F-) > 0$, then $\overline{G}(y) \ge \overline{G}(t_F-) > 0$ for each $y < t_F$ and the remainder terms in the approximations above go to zero uniformly in x. Thus, the convergence of $\widehat{\overline{F}}_n(x)$ to $L(x)$ above is uniform on $[0, t_F)$. Assume that $\overline{G}(t_F-) = 0$; then for each $k < t_F$, we have convergence of $\{n^{1/2}(\widehat{\overline{F}}_n(x) - \overline{F}(x))\colon 0 \le x \le k\}$. We need to prove that for each $\eta > 0$,

$$\text{(A.7)} \qquad \lim_{k\to t_F-} \limsup_{n\to\infty} \Pr\left\{ \sup_{k\le x<t_F} \left| n^{1/2}(\widehat{\overline{F}}_n(x) - \overline{F}(x)) \right| \ge \eta \right\} = 0.$$

Since for each x $\overline{G}(x) \ge \overline{F}(x), \overline{F}(t_F-) = 0$. Now, for $k \le x < t_F$ and $0 < \alpha < 1/2$, we have that

$$\begin{aligned}
\left| n^{1/2}(\widehat{\overline{F}}_n(x) - \overline{F}(x)) \right| &\le \sup_{y:\ 0\le y<x} \frac{n^{1/2}\overline{G}(x)|\overline{F}_n(y) - \overline{F}(y)|}{\overline{G}(y)} \\
&\le (\overline{G}(x))^{\alpha} \sup_{y:\ 0\le y<x} \frac{n^{1/2}|\overline{F}_n(y) - \overline{F}(y)|}{(\overline{G}(y))^{\alpha}} \\
&\le (\overline{G}(k))^{\alpha} \sup_{y:\ 0\le y<x} \frac{n^{1/2}|\overline{F}_n(y) - \overline{F}(y)|}{(\overline{F}(y))^{\alpha}},
\end{aligned}$$

which converges in distribution to

$$(\overline{G}(k))^{\alpha} \sup_{y:\ 0\le y<x} \frac{|W(F(y))|}{(\overline{F}(y))^{\alpha}}$$

[see, e.g., Theorem 3.7.1 in Shorack and Wellner (1986)]. Hence,

$$\begin{aligned}
&\lim_{k\to t_F-} \limsup_{n\to\infty} \Pr\left\{ \sup_{k\le x<t_F} \left| n^{1/2}(\widehat{\overline{F}}_n(x) - \overline{F}(x)) \right| \ge \eta \right\} \\
&\quad \le \lim_{k\to t_F-} \Pr\left\{ (\overline{G}(k))^{\alpha} \sup_{y:\ 0\le x<t_F} \frac{|W(F(x))|}{(\overline{F}(x))^{\alpha}} \ge \eta \right\} = 0.
\end{aligned}$$

So (5.7) follows. □

PROOF OF THEOREM 2.3. We consider the six cases in the proof of Theorem 2.2. In Cases 1 and 2, the limit distribution is $W(F(x))$ and neither (a) nor

(b) hold. In Case 3, the limit distribution is

$$\begin{aligned} L(x) &= -\max\left(\frac{\overline{F}(x)W(F(x-))}{\overline{F}(x-)}, W(F(x))\right) \\ &= -\frac{\overline{F}(x)W(F(x-))}{\overline{F}(x-)} - \max\left(0, W(F(x)) - \frac{\overline{F}(x)W(F(x-))}{\overline{F}(x-)}\right). \end{aligned} \tag{A.8}$$

Let $U_1 = \overline{F}(x)W(F(x-))/\overline{F}(x-)$ and let

$$U_2 = W(F(x)) - \overline{F}(x)W(F(x-))/\overline{F}(x-).$$

Since U_1 and U_2 are normal with zero means and $E[U_1U_2] = 0, U_1$ and U_2 are independent. It it easy to see that

$$\begin{aligned} &\Pr\{|U_1 + U_2^+| \geq t\} \\ &\quad = 2^{-1}\Pr\{(E[U_1^2])^{1/2}|Z| \geq t\} + 2^{-1}\Pr\{(E[U_1^2 + U_2^2])^{1/2}|Z| \geq t\}, \end{aligned}$$

where Z is a standard normal r.v. Note that

$$E[U_1^2] = \frac{(\overline{F}(x))^2(1 - \overline{F}(x-))}{\overline{F}(x-)} < \overline{F}(x)(1 - \overline{F}(x)) = E[U_1^2 + U_2^2].$$

This shows that $P(|L(x)| \geq t) = P(|U_1 + U_2^+| \geq t) < P(|U_1 + U_2| \geq t) = P(|W(F(x))| \geq t)$, so that the claim follows in this case.

Now, we consider Case 4. Here the limit distribution is

$$\begin{aligned} L(x) &= -\sup_{y:\ l(x) \leq y \leq x} \frac{\overline{F}(x)W(F(y))}{\overline{F}(y)} \\ &= -\frac{\overline{F}(x)W(F(l(x)))}{\overline{F}(l(x))} \\ &\quad - \sup_{y:\ l(x) \leq y \leq x}\left(\frac{\overline{F}(x)W(F(y))}{\overline{F}(y)} - \frac{\overline{F}(x)W(F(l(x)))}{\overline{F}(l(x))}\right). \end{aligned} \tag{A.9}$$

Taking covariances, it is easy to see that

$$\frac{\overline{F}(x)W(F(l(x)))}{\overline{F}(l(x))}$$

is independent of

$$\left\{\frac{W(F(y))}{\overline{F}(y)} - \frac{W(F(l(x)))}{\overline{F}(l(x))} \colon l(x) \leq y \leq x\right\}.$$

We also have that

$$\left\{\frac{\overline{F}(x)W(F(y))}{\overline{F}(y)} - \frac{\overline{F}(x)W(F(l(x)))}{\overline{F}(l(x))} \colon l(x) \leq y \leq x\right\}$$

has the distribution of

$$\left\{B\left(\frac{1}{\overline{F}(y)} - \frac{1}{\overline{F}(l(x))}\right) \colon l(x) \leq y \leq x\right\},$$

where $\{B(u)\colon 0 \le u\}$ is a Brownian motion. Thus, $L(x)$ has the distribution of

$$-\overline{F}(x)\left(\frac{1}{\overline{F}(l(x))}-1\right)^{1/2} Z_1 - \sup_{y\colon\, l(x)\le y\le x} \overline{F}(x)B\left(\frac{1}{\overline{F}(y)}-\frac{1}{\overline{F}(l(x))}\right),$$

where Z_1 is a standard normal r.v. independent of the Brownian motion B.

Now, we consider two subcases. If $\{\overline{F}(y)\colon l(x)\le y\le x\} = [\overline{F}(x), \overline{F}(l(x))]$ then

$$\sup_{y\colon\, l(x)\le y\le x} \overline{F}(x)B\left(\frac{1}{\overline{F}(y)}-\frac{1}{\overline{F}(l(x))}\right) = \sup_{0\le u\le 1/\overline{F}(x)-1/\overline{F}(l(x))} \overline{F}(x)B(u).$$

By Billingsley [(1968), page 72], the random variable

$$\sup_{0\le u\le 1/\overline{F}(x)-1/\overline{F}(l(x))} \overline{F}(x)B(u)$$

has the distribution of

$$\overline{F}(x)\left(\frac{1}{\overline{F}(x)}-\frac{1}{\overline{F}(l(x))}\right)^{1/2}|Z_2|,$$

where Z_2 is a standard normal r.v. independent of Z_1. So $|L(x)|$ has the distribution of

$$\left|\overline{F}(x)\left(\frac{1}{\overline{F}(l(x))}-1\right)^{1/2} Z_1 + \overline{F}(x)\left(\frac{1}{\overline{F}(x)}-\frac{1}{\overline{F}(l(x))}\right)^{1/2}|Z_2|\right|, \tag{A.10}$$

which has the distribution of $|W(F(x))|$.

If $\{\overline{F}(y)\colon l(x)\le y\le x\} \ne [\overline{F}(x), \overline{F}(l(x))]$, then with positive probability,

$$\sup_{y\colon\, l(x)\le y\le x} \overline{F}(x)B\left(\frac{1}{\overline{F}(y)}-\frac{1}{\overline{F}(l(x))}\right) < \sup_{0\le u\le 1/\overline{F}(x)-1/\overline{F}(l(x))} \overline{F}(x)B(u).$$

An elementary computation shows that, for any fixed $t>0$, $\Pr\{|Z+b|\ge t\}$ is increasing in $b>0$, where Z is a standard normal random variable. Thus,

$$\begin{aligned}&\Pr\left\{\left|\overline{F}(x)\left(\frac{1}{\overline{F}(l(x))}-1\right)^{1/2} Z_1 + \sup_{y\colon\, l(x)\le y\le x} \overline{F}(x)B\left(\frac{1}{\overline{F}(y)}-\frac{1}{\overline{F}(l(x))}\right)\right| \ge t\right\}\\ &\quad < \Pr\left\{\left|\overline{F}(x)\left(\frac{1}{\overline{F}(l(x))}-1\right)^{1/2} Z_1 + \sup_{u\colon\, 0\le u\le 1/\overline{F}(x)-1/\overline{F}(l(x))} \overline{F}(x)B(u)\right| \ge t\right\}.\end{aligned}$$

Thus the claim follows in this case. Cases 5 and 6 follow similarly. □

PROOF OF THEOREM 2.4. The arguments in our proof of Theorem 2.2 apply. The only change is that a different expression appears in the supremum in the limit distribution.

We consider the same six cases as in Theorem 2.2. In Cases 1 and 2, we have that $\mathscr{A}(x) = \{\overline{F}(x)\}$,

$$n^{1/2} \inf_{y:\ 0 \le y \le x} \left(\frac{\overline{F}_n(y)((1-\alpha)\overline{F}_n(x) + \alpha\overline{G}(x))}{(1-\alpha)\overline{F}_n(y) + \alpha\overline{G}(y)} - \overline{F}_n(x) \right) \xrightarrow{\Pr} 0.$$

and $L_\alpha(x) = W(F(x))$. In Case 3, we have that

$$n^{1/2} \inf_{y:\ 0 \le y < x} \left(\frac{\overline{F}_n(y)((1-\alpha)\overline{F}_n(x) + \alpha\overline{G}(x))}{(1-\alpha)\overline{F}_n(y) + \alpha\overline{G}(y)} - \frac{\overline{F}_n(x-)((1-\alpha)\overline{F}_n(x) + \alpha\overline{G}(x))}{(1-\alpha)\overline{F}_n(x-) + \alpha\overline{G}(x-)} \right) \xrightarrow{\Pr} 0.$$

The proof follows as in Theorem 2.2, except that since $H(x) = H(x-)$ in this case, we have

$$\begin{aligned} n^{1/2} &\left(\frac{\overline{F}_n(x-)((1-\alpha)\overline{F}_n(x) + \alpha\overline{G}(x))}{(1-\alpha)\overline{F}_n(x-) + \alpha\overline{G}(x-)} - \overline{F}(x) \right) \\ &= \frac{(1-\alpha)\overline{F}_n(x-)n^{1/2}(\overline{F}_n(x) - \overline{F}(x)) + \alpha\overline{G}(x)n^{1/2}(\overline{F}_n(x-) - \overline{F}(x-))}{(1-\alpha)\overline{F}_n(x-) + \alpha\overline{G}(x-)}, \end{aligned}$$

which converges to

$$\frac{-(1-\alpha)\overline{F}(x-)W(F(x)) - \alpha\overline{G}(x)W(F(x-))}{(1-\alpha)\overline{F}(x-) + \alpha\overline{G}(x-)}. \tag{A.11}$$

Thus the limit distribution is that in (A.4) with the term in (A.10) replacing

$$\frac{-\overline{F}(x)W(F(x-))}{\overline{F}(x-)}.$$

Similar arguments apply in Cases 4, 5 and 6. □

PROOF OF THEOREM 2.5. We have that $W(1-t)/t$ has the distribution of $W(1-m(x))/m(x) + B(1/t - 1/m(x))$, where $\{B(u): u \ge 0\}$ is a Brownian motion independent of $W(1-m(x))$. So $L(x)$ has the distribution of

$$-\frac{\overline{F}(x)W(1-m(x))}{m(x)} - \frac{\overline{F}(x)((1-\alpha)U_1 + \alpha U_2)}{(1-\alpha)\overline{F}(x) + \alpha\overline{G}(x)},$$

where $U_1 = \overline{F}(x)B(1/\overline{F}(x) - 1/m(x))$ and $U_2 = \overline{G}(x)\sup_{t \in \mathscr{A}(x)} B(1/t - 1/m(x))$. We have that

$$\begin{aligned} E[L_\alpha^2(x)] &= \frac{(\overline{F}(x))^2(1-m(x))}{m(x)} + (\overline{F}(x))^2 \left(\frac{1}{\overline{F}(x)} - \frac{1}{m(x)} \right) \\ &\quad \times \frac{(1-\alpha)^2(\overline{F}(x))^2 + 2\alpha(1-\alpha)\overline{F}(x)\overline{G}(x)a_1 + \alpha^2(\overline{G}(x))^2 a_2}{((1-\alpha)\overline{F}(x) + \alpha\overline{G}(x))^2}. \end{aligned}$$

By Lemma A.1, the last expression is minimized at

$$\alpha_0 = \frac{(1-a_1)\overline{F}(x)}{(1-a_1)\overline{F}(x) + \overline{G}(x)(a_2 - a_1)}.$$

An elementary computation shows that (A.6) holds. Now, because

$$0 \le \sup_{t \in \mathscr{A}(x)} B(t^{-1} - (m(x))^{-1}) \le \sup_{0 \le u \le (\overline{F}(x))^{-1} - (m(x))^{-1}} B(u),$$

we have $0 \le a_2 \le 1$. Since

$$B((\overline{F}(x))^{-1} - (m(x))^{-1}) \le \sup_{t \in \mathscr{A}(x)} B(t^{-1} - (m(x))^{-1}),$$

we also have that $a_1 \le a_2$. Further, the Cauchy–Schwartz inequality implies that $|a_1| < a_2^{1/2}$. It follows from these three inequalities in a_1 and a_2 that $0 \le A \le 1$. The replacement of α_0 by $\hat{\alpha}_0$, which has no effect on the limiting distribution, is justified by the continuity of $L_\alpha(x)$ in α, the fact that $\hat{\alpha}_0 \xrightarrow{p} \alpha_0$ and the fact that the process $\{L_\alpha(x)\}$ is tight. □

PROOF OF THEOREM 2.6. We apply Theorem 2.5. We need to find a_1 and a_2. By the reflection principle $a_2 = E[(\sup_{0 \le t \le 1} B(t))^2] = 1$ [see Billingsley (1968)]. Let $U = \sup_{0 \le t \le 1} B(t)$ and let $V = B(1)$. It is known that the joint density of (U, V) is given by $f_{U,V}(u, v) = 2(2u - v)\exp(-(v-2u)^2)/\sqrt{2\pi}$, $u \ge v$, $u \ge 0$ [see Equation (11.11) in Billingsley (1968), page 79]. It follows that $a_1 = E[UV] = 1/2$ and $A = 3/4$. The formulas in (2.15)–(2.17) follow from (2.12)–(2.14) upon substituting these values for a_1 and a_2 and A. □

PROOF OF THEOREM 2.7. This follows from Theorem 2.5 upon evaluating $a_1 = E[B(1)\max(0, B(1))] = 1/2$ and $a_2 = E[(\max(0, B(1)))^2] = 1/2$. □

PROOF OF THEOREM 2.8. It suffices to prove that if $0 \le t \le 1$, then

$$\begin{aligned} E[B(1)\max(0, B(t), B(1))] = 2^{-1} - t2^{-2} \\ - (2\pi)^{-1} \arctan((t/(1-t))^{1/2}) \end{aligned} \tag{A.12}$$

and

$$\begin{aligned} E[(\max(0, B(t), B(1)))^2] = 2^{-1} + t2^{-2} + (2\pi)^{-1}(t(1-t))^{1/2} \\ - (2\pi)^{-1} \arctan((t/(1-t))^{1/2}). \end{aligned} \tag{A.13}$$

To prove (A.12) let $X = t^{-1/2}B(t)$ and let $Y = (1-t)^{-1/2}(B(1) - B(t))$. By Lemma A.2, we have

$$\begin{aligned} E[B(1)\max(0, B(t), B(1))] &= E[B(1)(B(t) + \max(-B(t), 0, B(1) - B(t)))] \\ &= t + E[B(t)\max(-B(t), 0, B(1) - B(t))] \\ &\quad + E[(B(1) - B(t))\max(-B(t), 0, B(1) - B(t))] \\ &= t - tE[X\max(0, X, ((1-t)/t)^{1/2}Y)] \\ &\quad + (1-t)E[X\max(0, Y, (t/(1-t))^{1/2}X)] \\ &= t - t(2^{-1} - (2\pi)^{-1}\arctan((1-t)/t)^{1/2}) \\ &\quad + (1-t)(2^{-1} - (2\pi)^{-1}\arctan(t/(1-t))^{1/2}) \\ &= 2^{-1} - t2^{-2} - (2\pi)^{-1}\arctan((t/(1-t))^{1/2}). \end{aligned}$$

Equation (A.13) follows similarly. □

PROOF OF THEOREM 2.9. By Theorem 2.3, we only need to consider the case $0 < \alpha < 1$. If $\mathscr{A}(x) = \{\overline{F}(x)\}$, then $L_\alpha(x) = W(F(x))$ and part (i) follows.

Next, we consider case (ii). In this case, $\mathscr{A}(x) = [\overline{F}(l(x)), \overline{F}(x)]$ for some $l(x) < x$. By the proof of Theorem 2.5, $L_\alpha(x)$ has the distribution of

$$-\frac{\overline{F}(x)W(\overline{F}(l(x)))}{\overline{F}(l(x))} - \overline{F}(x)\left(\frac{1}{\overline{F}(x)} - \frac{1}{\overline{F}(l(x))}\right)^{1/2}\left(\tau B(1) + (1-\tau)\sup_{0\le u\le 1} B(u)\right),$$

where $\tau = \alpha\overline{G}(x)/((1-\alpha)\overline{F}(x) + \alpha\overline{G}(x))$. We claim that for each $t > 0$,

$$\text{(A.14)} \qquad \Pr\left\{\left|\tau B(1) + (1-\tau)\sup_{0\le u\le 1} B(u)\right| \ge t\right\} < \Pr\{|Z| \ge t\},$$

where Z is a standard normal r.v. We have noted earlier that, given $t > 0$ and a standard normal r.v. Z, $\Pr\{|Z + b| \ge t\}$ is increasing in $b > 0$. Thus

$$\begin{aligned} &\Pr\{|L_\alpha(x)| \ge t\} \\ &\quad < \Pr\left\{\left|\frac{\overline{F}(x)W(\overline{F}(l(x))}{\overline{F}(l(x))} + \overline{F}(x)\left(\frac{1}{\overline{F}(x)} - \frac{1}{\overline{F}(l(x))}\right)^{1/2} Z\right| \ge t\right\} \\ &\quad = \Pr\{|W(F(x))| \ge t\}, \end{aligned}$$

where we have used that

$$\frac{\overline{F}(x)W(\overline{F}(l(x))}{\overline{F}(l(x))} + \overline{F}(x)\left(\frac{1}{\overline{F}(x)} - \frac{1}{\overline{F}(l(x))}\right)^{1/2} Z$$

is a normal r.v. with mean zero and the variance of $W(F(x))$.

Let $U = \sup_{0\le t\le 1} B(t)$ and let $V = B(1)$, as in the proof of Theorem 2.6. An elementary computation shows that for each $t > 0$,

$$\Pr\{|\tau V + (1-\tau)U| \ge t\} = \frac{1}{2-\tau}\Pr\left\{|Z| \ge \frac{t}{1-\tau}\right\} + \frac{1-\tau}{2-\tau}\Pr\left\{|Z| \ge \frac{t}{1-\tau}\right\},$$

which implies (A.14). □

PROOF OF THEOREM 3.1. If $m/n \to \infty$, then

$$\begin{aligned}
&\left| \inf_{y:\ 0\le y\le x} \left(\frac{\overline{F}_n(y)((1-\alpha)\overline{F}_n(x)+\alpha\overline{G}_m(x))}{(1-\alpha)\overline{F}_n(y)+\alpha\bar{G}_m(y)} \right) \right. \\
&\qquad \left. - \inf_{y:\ 0\le y\le x} \left(\frac{\overline{F}_n(y)((1-\alpha)\overline{F}_n(x)+\alpha\overline{G}(x))}{(1-\alpha)\overline{F}_n(y)+\alpha\overline{G}(y)} \right) \right| \\
&\qquad \le \sup_{y:\ 0\le y\le x} \left| \frac{\overline{F}_n(y)((1-\alpha)\overline{F}_n(x)+\alpha\overline{G}_m(x))}{(1-\alpha)\overline{F}_n(y)+\alpha\overline{G}_m(y)} \right. \\
&\qquad\qquad \left. - \frac{\overline{F}_n(y)((1-\alpha)\overline{F}_n(x)+\alpha\overline{G}(x))}{(1-\alpha)\overline{F}_n(y)+\alpha\overline{G}(y)} \right|.
\end{aligned}$$

Since $\sup_{x>0} m^{1/2}|\overline{G}_m(x)-\overline{G}(x))|$ is bounded in probability, the expression above is $O_p(m^{-1/2})$. Thus, the asymptotics of the two expressions in (1.4) and (3.2) are the same; these asymptotics are considered in detail in Section 2.

If $m/n \to c$, for some $0 < c < \infty$, we proceed as in Theorem 2.2. In Cases 1 and 2 we have that

$$n^{1/2} \inf_{y:\ 0\le y\le x} \left(\frac{\overline{F}_n(y)((1-\alpha)\overline{F}_n(x)+\alpha\overline{G}_m(x))}{(1-\alpha)\overline{F}_n(y)+\alpha\overline{G}_m(y)} - \overline{F}_n(x) \right) \xrightarrow{\Pr} 0.$$

In Case 3, we have that

$$\begin{aligned}
&n^{1/2} \inf_{y:\ 0\le y< x} \left(\frac{\overline{F}_n(y)((1-\alpha)\overline{F}_n(x)+\alpha\overline{G}_m(x))}{(1-\alpha)\overline{F}_n(y)+\alpha\bar{G}_m(y)} \right. \\
&\qquad\qquad \left. - \frac{\overline{F}_n(x-)((1-\alpha)\overline{F}_n(x)+\alpha\overline{G}_m(x))}{(1-\alpha)\overline{F}_n(x-)+\alpha\overline{G}_m(x-)} \right) \xrightarrow{\Pr} 0.
\end{aligned}$$

The proof follows as in Theorem 2.2, with the obvious difference that

$$\begin{aligned}
&\frac{\overline{F}_n(x-)((1-\alpha)\overline{F}_n(x)+\alpha\overline{G}_m(x))}{(1-\alpha)\overline{F}_n(x-)+\alpha\overline{G}_m(x-)} - \overline{F}(x) \\
&\quad = \frac{(1-\alpha)\overline{F}_n(x-)n^{1/2}(\overline{F}_n(x)-\overline{F}(x))+\alpha\overline{G}(x)n^{1/2}(\overline{F}_n(x-)-\overline{F}(x-))}{(1-\alpha)\overline{F}_n(x-)+\alpha\overline{G}(x-)} \\
&\qquad + \frac{\alpha\overline{F}_n(x-)n^{1/2}(\overline{G}_m(x)-\overline{G}(x))-\alpha\bar{F}(x)n^{1/2}(\overline{G}_m(x-)-\overline{G}(x-))}{(1-\alpha)\overline{F}_n(x-)+\alpha\overline{G}(x-)},
\end{aligned}$$

which converges in distribution to

$$\frac{-(1-\alpha)\overline{F}(x-)W_1(F(x)) - \alpha\overline{G}(x)W_1(F(x-))}{(1-\alpha)\overline{F}(x-)+\alpha\overline{G}(x-)} - \frac{c^{-1/2}\alpha\overline{F}(x-)W_2(G(x)) + c^{-1/2}\alpha\overline{F}(x)W_2(G(x-))}{(1-\alpha)\overline{F}(x-)+\alpha\overline{G}(x-)}.$$

Similar arguments give Cases 4, 5 and 6. The proof of our claim when $m/n \to 0$ is similar and it is omitted. □

PROOF OF THEOREM 3.2. Let $\{B_1(u)\colon u \geq 0\}$ and let $\{B_2(u)\colon u \geq 0\}$ be two independent Brownian motions independent of $W_1(1-m(x))$ and $W_2(1-m(x))$. We have that $W_1(1-t)/t$ has the distribution of

$$\frac{W_1(1-m(x))}{m(x)} + B_1\left(\frac{1}{t} - \frac{1}{m(x)}\right).$$

We also have that

$$\frac{W_2(1-t(\overline{F}(x))^{-1}\overline{G}(x))}{t}$$

has the distribution of

$$\frac{W_2(1-m(x))(\overline{F}(x))^{-1}\overline{G}(x))}{m(x)} + (\overline{F}(x))^{1/2}(\overline{G}(x))^{1/2}B_2\left(\frac{1}{t} - \frac{1}{m(x)}\right).$$

Hence, $L_{\alpha,c}(x)$ has the distribution of

$$-\frac{\overline{F}(x)W(1-m(x))}{m(x)} - \frac{\overline{F}(x)((1-\alpha)U_1 + \alpha U_2)}{(1-\alpha)\overline{F}(x)+\alpha\overline{G}(x)},$$

where $U_1 = \overline{F}(x)B_1\left(1/\overline{F}(x) - 1/m(x)\right)$ and

$$\begin{aligned} U_2 &= c^{-1/2}(\overline{F}(x))^{1/2}(\overline{G}(x))^{1/2}B_2\left(\frac{1}{\overline{F}(x)} - \frac{1}{m(x)}\right) \\ &\quad + \sup_{t\in\mathscr{A}(x)}\left(\overline{G}(x)B_1\left(\frac{1}{t} - \frac{1}{m(x)}\right)\right. \\ &\quad \left. - c^{-1/2}(\overline{F}(x))^{1/2}(\overline{G}(x))^{1/2}B_2\left(\frac{1}{t} - \frac{1}{m(x)}\right)\right). \end{aligned}$$

We have that

$$E[L^2_{\alpha,c}(x)] = \frac{(\overline{F}(x))^2(1-m(x))}{m(x)} + (\overline{F}(x))^2 E\left[\left(\frac{(1-\alpha)U_1+\alpha U_2}{(1-\alpha)\bar{F}(x)+\alpha\overline{G}(x)}\right)^2\right].$$

By Lemma A.1, we get that the last expression is minimized at the claimed α_0. □

Acknowledgments. The authors thank Hari Mukerjee for his many helpful comments on an earlier version of this paper. Thanks are also due to an Associate Editor and to a referee for constructive criticisms that led to improved clarity.

REFERENCES

BAGAI, I. and KOCHAR, S. C. (1986). On tail–ordering and comparison of failure rates. *Comm. Statist. A Theory Methods* **15** 1377–1388.

BARLOW, R. E., TOLAND, R. H. and FREEMAN, T. (1988). A Bayesian analysis of the stress-rupture life of Kevlar/Epoxy spherical pressure vessels. In *Accelerated Life Testing and Experts' Opinions* (C. A. Clarotti and D. V. Lindley, eds.) 203–236. North-Holland, Amsterdam.

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

BOLAND, P. J., EL-NEWEIHI, E. and PROSCHAN, F. (1994). Applications of the hazard rate ordering in reliability and order statistics. *J. Appl. Probab.* **31** 180–192.

BRESLOW, N. and CROWLEY, J. (1974). A large sample study of the life table and product limit estimates under random censorship. *Ann. Statist.* **2** 473–453.

CAPÉRAÀ, P. (1988). Tail ordering and asymptotic efficiency of rank tests. *Ann. Statist.* **16** 470–478.

DYKSTRA, R., KOCHAR, S. and ROBERTSON, T. (1991). Statistical inference for uniform stochastic ordering in several populations. *Ann. Statist.* **19** 870–888.

GILL, R. D. (1981). Testing with replacement and the product limit estimator. *Ann. Statist.* **9** 853–860.

KEILSON, J. and SUMITA, U. (1982). Uniform stochastic ordering and related inequalities. *Canad. J. Statist.* **10** 101–198.

MUKERJEE, H. (1996). Estimation of survival functions under uniform stochastic ordering. *J. Amer. Statist. Assoc.* **91** 1684–1689.

ROJO, J. (1998). Estimating of the quantile function of an IFRA distribution. *Scand. J. Statist.* **25** 293–310.

ROJO, J. and SAMANIEGO, F. J. (1991). On nonparametric maximum likelihood estimation of a distribution uniformly stochastically smaller than a standard. *Statist. Probab. Lett.* **11** 267–271.

ROJO, J. and SAMANIEGO, F. J. (1993). On estimating a survival curve subject to a uniform stochastic ordering constraint. *J. Amer. Statist. Assoc.* **88** 566–572.

ROSS, S. M. (1983). *Stochastic Processes*. Wiley, New York.

SHAKED, M. and SHANTHIKUMAR, J. G. (1994). *Stochastic Orders and Their Applications.* Academic Press, Boston.

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

YANAGIMOTO, T. and SIBUYA, M. (1972). Stochastically larger component of a random vector. *Ann. Inst. Statist. Math.* **24** 259–269.

WHITT, W. (1980). Uniform conditional stochastic order. *J. Appl. Probab.* **17** 112–123.

DEPARTMENT OF MATHEMATICAL SCIENCES
BINGHAMTON UNIVERSITY
BINGHAMTON, NEW YORK 13902
E-MAIL: arcones@math.binghamton.edu

DIVISION OF STATISTICS
UNIVERSITY OF CALIFORNIA
DAVIS, CALIFORNIA 95616
E-MAIL: fjsamaniego@ucdavis.edu