JOURNAL OF MULTIVARIATE ANALYSIS 49, 150–163 (1994)

Uniform Strong Consistent Estimation of an IFRA Distribution Function

JAVIER ROJO*

University of Texas at El Paso

AND

FRANCISCO J. SAMANIEGO†

University of California at Davis

Let $\bar{F}_n$ be an estimator of an IFRA survival function $\bar{F}$ and let A be such that $0 < \bar{F}(A) < 1$. The main result constructs an IFRA estimator by splicing the smallest increasing failure rate on the average majorant and greatest increasing failure rate on the average minorant of the restrictions of $\bar{F}_n$ to the intervals $[0, A]$ and $[A, \infty)$, respectively. The resulting estimator $\hat{\bar{F}}_n$ has the property that $\sup_x |\hat{\bar{F}}_n - \bar{F}| \leqslant k \sup_x |\bar{F}_n - \bar{F}|$, where $k \geqslant 2$, and $k = 2$ if and only if A is the median of F. As a consequence, if $\bar{F}_n$ represents the empirical survival function, or the Kaplan–Meier estimator, the estimator $\hat{\bar{F}}_n$ inherits the strong and uniform convergence properties, as well as the optimal rates of convergence of the empirical survival function and Kaplan–Meier estimator respectively. Simulations show a substantial improvement in mean-squared error when comparing $\hat{\bar{F}}_n$ to those IFRA estimators available in the literature. Under suitable conditions, asymptotic confidence intervals for $\bar{F}(t_0)$ are also provided.

1. INTRODUCTION

Let F be a distribution function on $(0, \infty)$, and let $\bar{F} = 1 - F$ be its corresponding survival function. The empirical distribution function F_n enjoys many good properties as an estimator of F. For example, F_n is

Received October 27, 1992; revised August 6, 1993

AMS 1991 classifications: primary 62N05, secondary 60F15, 62G30, 62G05.

Key words and phrases: hazard function, tails of distributions, regularly varying tails, Kaplan–Meier estimator, censoring, SIFRAM, GIFRAM.

* The partial support of the National Science Foundation, the Air Force Office of Sponsored Research, and the National Institutes of Health through Grant NSF-DMS-86606420, Contract AFOSR F49620-89-C-0085, and Grants SO6 GM08012-21 and 1 G12 RR08124-01, respectively, is gratefully acknowledged.

† Research partially supported by the Air Force Office of Scientific Research under Grant AFOSR 88-308 and by the National Security Agency under Grant MDA 904-90-H-4031.

0047-259X/94 $6.00

unbiased and mean square consistent, as well as almost surely uniformly consistent with optimal rates $O(n^{-1})$ and $O(n^{-1/2}(\log\log n)^{1/2})$ respectively. Moreover, as shown by Dvoretzky *et al.* (1956), F_n is asymptotically minimax among the collection of all continuous distribution functions. When F, however, is know to satisfy certain nonparametric constraints, F_n does not typically belong to the constrained class of distributions in question and then it is of interest to search for estimators belonging to the class and sharing many of the good properties F_n has in the unrestricted case.

Several authors have considered the estimation of a distribution function subject to membership in a constrained class. Grenander (1956) proved that the maximum likelihood estimator of a concave distribution function F is the least concave majorant C_n of the empirical distribution function F_n. Consistent estimators of IFR distributions were obtained by Marshall and Proschan (1965). Boyles and Samaniego (1984) provided estimators for NBU survival functions and proved their strong uniform consistency in compact sets. Barlow and Scheuer (1971), hereafter referred to as BS (1971), considered an estimator for an IFRA distribution and a proof of the strong consistency of the estimator was given there as well as in Barlow *et al.* (1972). When F is convex, Kiefer and Wolfowitz (1976) showed the greatest convex minorant of F_n to be asymptotically minimax. Wang (1988) showed that the greatest star-shaped minorant (GSM) of $\bar{F}_n$ is closer to the true star-shaped F than F_n is in the sup norm.

The focus of this paper is the IFRA class. Recall that F is IFRA if and only if its hazard function is star-shaped. That is, if and only if

$$-\ln\frac{\bar{F}(t)}{t} \text{ is nondecreasing for } t\in(0,\infty). \tag{11}$$

Since the IFRA class of distribution is the smallest class which contains the exponential distribution and is closed under the formation of coherent systems and weak limits (see Birnbaum *et al.* (1966)), it has found an important place in reliability theory. Its importance also derives from the fact that it arises in connection with shock models and wear processes. See, for example, Esary *et al.* (1973), Esary and Marshall (1975), and A-Hameed and Proschan (1973, 1975).

It is clear that any distribution function which is constant on an interval does not satisfy (1.1) and hence F_n is not IFRA. The search for alternatives to F_n which belong to the IFRA class and possess good properties goes back to BS (1971). It is well known (e.g., Boyles *et al.* (1985)) that the maximum likelihood estimator (MLE) in this case converges but to the wrong distribution. For example, when $F(x)=x$, $0\leqslant x\leqslant 1$, the MLE of $\bar{F}$ converges to $[(1-x)/(1+x)]^x$ rather than to $(1-x)$. Barlow and Scheuer (1971) constructed an estimator $\bar{F}_n^*$ using isotonic regression ideas, and a proof of the strong consistency of $\bar{F}_n^*$ was given in BS (1971, p. 155) and Barlow *et al.* (1972, p. 259). However, some of the conditions assumed in their proof do not hold for all IFRA distributions. More precisely, a condition needed to prove the strong consistency of $\bar{F}_n^*$ is that

$$\lim_{n\to\infty} \max_{1\leqslant i\leqslant n-1} |H_n(X_{(i)}) - H(X_{(i)}| = 0 \quad \text{a.s.,} \tag{1.2}$$

where H and H_n are the true and empirical failure rate average functions defined by $H(t) = \log(1 - F(t))/t$ and $H_n(t) = \log(1 - F_n(t))/t$, respectively, and where $X_{(1)}, ..., X_{(n)}$ represent the order statistics. However, as shown by Boyles (1981), for the case $F(t) = 1 - e^{-t}$,

$$\Pr\{|H_n(X_{(1)}) - H(X_{(1)})| > \varepsilon\} \geqslant (1 - 1/n)^{n/(1-\varepsilon)} \to e^{-((1-\varepsilon)^{-1})}$$

and hence (1.2) does not hold. The strong consistency of the isotonic regression estimator continues to be an open question.

Wang (1987b) gives a nice treatment of the optimal asymptotic character of the estimator G_n, whose hazard function C_n is the GSM of the sample hazard function $[0, \lambda]$. In particular, it is shown there that G_n is $n^{1/2}$-equivalent to the sample distribution function F_n in the sense that $\sup_{x\leqslant\lambda} n^{1/2} |G_n - F_n| \xrightarrow{P} 0$, for any λ with $F(\lambda) < 1$, provided F is, essentially, strictly IFRA. When C_n^* is the GSM of the sample hazard function on $(0, \infty)$, under more stringent conditions (see Theorem 2 of Wang (1987b)), it was also shown that $\sup_{x\leqslant\lambda} n^{1/2} |F_n - G_n^*| \xrightarrow{P} 0$, where G_n^* is the distribution function with hazard rate function C_n^* and $F(\lambda) < 1$. This implies that G_n and G_n^* are weakly uniformly consistent on compact sets. The conditions of Theorem 1 of Wang (1987b) are not, for example, satisfied when F has a Weibull distribution with shape parameter $1 \leqslant \alpha < 2$, and the conditions of Theorem 2 of Wang (1987b) do not hold for the Weibul distribution with shape parameter $\alpha \geqslant 1$. The estimation of an IFRA survival function based on censored data has been presented in Wang (1987a), where estimators are constructed which converge weakly and uniformly on compact sets.

Thus, the problem of the existence of strong uniform consistent estimators for an IFRA distribution, which belong to the IFRA class, remains open. The purpose of the present paper is to find estimators of F which are members of the IFRA class and have good properties under a minimal set of conditions. In particular, it is shown that the estimator

$$\hat{\bar{F}}_n(t) = \begin{cases} \sup_{t\leqslant s\leqslant A} (\bar{F}_n(s))^{t/s}, & t < A, \\ \inf_{A\leqslant s\leqslant t} (\bar{F}_n(s))^{t/s}, & t \geqslant A, \end{cases} \tag{1.3}$$

where $\bar{F}_n = 1 - F_n$ and A is such that $0 < \bar{F}(A) < 1$, is IFRA and inherits the consistency properties of $\bar{F}_n$, and its rates of convergence. More specifically, it is shown that

$$\sup_x |\hat{\bar{F}}_n(x) - \bar{F}(x)| \leqslant k \sup_x |\bar{F}_n(x) - \bar{F}(x)| \qquad \text{for some} \quad k \geqslant 2,$$

where $k = 2$ if and only if A is the median of F.

The organization of the paper is as follows. Section 2 defines two IFRA transforms which map an arbitrary function f with support $I \subset [0, \infty)$ into IFRA functions f_1 and f_2. The fixed points of these transforms are the IFRA functions. It is shown that f_1 corresponds to the greatest IFRA minorant (GIFRAM) of f. Similarly, f_2 defines the smallest IFRA majorant (SIFRAM) of f. Section 3 considers the main result which constructs an IFRA estimator by splicing the SIFRAM and the GIFRAM of the restrictions of $\bar{F}_n$ to the intervals $[0, A]$ and $[A, \infty)$, respectively. The estimator $\hat{\bar{F}}_n$ inherits the consistency properties of $\bar{F}_n$ and its convergence rates. In the case of no censoring, letting $\bar{F}_n$ be the empirical survival function, $\hat{\bar{F}}_n$ enjoys strong uniform consistency with optimal rates $O(n^{-1/2}(\log \log n)^{1/2})$. In the case of censored data, $\bar{F}_n$ may be chosen to be the Kaplan–Meier estimator and then $\hat{\bar{F}}_n$ inherts all the consistency properties of $\bar{F}_n$. Csörgő and Horváth (1983) provide a nice treatment of the strong uniform consistency results for the product-limit etimator.

2. The IFRA Transforms

Let f be a nonnegative real-valued function with domain $I \subset [0, \infty)$, where $I = [a, b)$. By a slight abuse of the terminology, f is said to be IFRA if $-\log f$ is star-shaped. Clearly, when f is a survival function, the terminology agrees with the usual definition of an IFRA distribution function.

Let f_1 and f_2 be the greatest IFRA minorant (GIFRAM) and the smallest IFRA majorant (SIFRAM) of f, respectively. That is,

$$\begin{aligned} f_1(x) &= \sup\{g(x)\colon -\ln g \text{ is star-shaped on } I \text{ and } g \leqslant f\} \\ f_2(x) &= \inf\{g(x)\colon -\ln g \text{ is star-shaped on } I \text{ and } g \geqslant f\}. \end{aligned} \tag{2.1}$$

It is not obvious that f_1 and f_2 are well-defined, and it is not clear that functions g exist such that $g \leqslant f$ $(g \leqslant f)$ and $-\ln g$ is star-shaped. Consider the following functions:

$$h_1(t) = \inf_{a \leqslant s \leqslant t} (f(s))^{t/s}, \qquad a \leqslant t < b \tag{2.2}$$

and

$$h_2(t) = \sup_{t \leqslant s \leqslant b} (f(s))^{t/s}, \qquad a \leqslant t < b. \tag{2.3}$$

Then it is easy to see that h_1 and h_2 are IFRA with $h_1 \leqslant f$ and $h_2 \geqslant f$ so that the sets of functions over which the supremum and the infinum are taken in (2.1) are not empty. To motivate (2.2), note that if f is IFRA, then $(f(t))^{1/t}$ is nonincreasing so that $(f(t))^{1/t} = \inf_{s \leqslant t}(f(s))^{1/s}$. If f is not IFRA, f violates the previous identity and (2.2) represents the adjustment needed to map f into an IFRA function. Similar comments apply to (2.3). Moreover, it turns out that $h_1 = f_1$ and $h_2 = f_2$ as the next theorem shows.

THEOREM 2.1. *Let f be nonnegative with domain $I \subset [0, \infty)$, and let h_1 and h_2 be defined as in* (2.2) *and* (2.3). *Then,*

(i) *$f = h_1 = h_2$ if and only if f is IFRA.*

(ii) *$h_1(h_2)$ is the GIFRAM (SIFRAM) of f.*

Proof. To show (i), note that if f is IFRA, then $(f(t))^{1/t}$ is nonincreasing so that $\inf_{a \leqslant s \leqslant t}(f(s))^{t/s} = \sup_{t \leqslant s \leqslant b}(f(s))^{t/s} = f(t)$. Since h_1, h_2 are IFRA, the converse follows immediately. To prove (ii), let g be any IFRA function with $g \leqslant f$. Then by (i),

$$g(t) = \inf_{a \leqslant s \leqslant t} (g(s))^{t/s} \leqslant \inf_{a \leqslant s \leqslant t} (f(s))^{t/s} = h_1(t).$$

Hence h_1 is the GIFRAM of f. A similar argument shows that h_2 is the SIFRAM of f. ∎

It is worthwhile noting that while (2.2) defines a survival function when f is a survival function, the transform (2.3) may map a survival function $\bar{F}$ to a function $\bar{F}^*$, with $\bar{F}^* = 1$ for all x. This occurs for distributions with slowly or regularly varying tails. To see this, let $\bar{F}(t) = h(t)\, t^{-\alpha}$ where h is slowly varying at infinity and $\alpha > 0$. It follows that

$$-\ln h_2(t) = t \inf_{s \geqslant t} \left\{ \frac{-\ln h(s)}{s} + \alpha \frac{\ln s}{s} \right\} = 0,$$

since $-\ln h(s)/s \to 0$ and $\ln s/s \to 0$. When $\bar{F}$ has bounded support, (2.3) defines a true survival function. Thus, whether or not (2.3) defines a survival function depends on the tail heaviness of $\bar{F}$. One important special case for which h_2 maps f into a survival function is that of distributions with exponential tails. Rojo (1988, 1991) discusses a general concept of tail-heaviness and treats the case of exponentially tailed distributions in particular.

3. Estimation of an IFRA Distribution Function

The SIFRAM and GIFRAM of $\bar{F}_n$ are unsatisfactory as etimators of $\bar{F}$. Neither achieves uniform consistency because each has poor performance near one of the tails of F (i.e., either near zero or near ∞). However, it is possible to exploit the best properties of each estimator as the following construction will demonstrate.

Let $\bar{F}_n$ be a survival function which is estimating the IFRA distribution F. Let A be such that $0 < \bar{F}(A) < 1$. Define

$$\hat{\bar{F}}_n(t) = \begin{cases} \sup_{t \leqslant s \leqslant A} (\bar{F}_n(s))^{t/s}, & t < A, \\ \inf_{A \leqslant s \leqslant t} (\bar{F}_n(s))^{t/s}, & t \geqslant A. \end{cases} \tag{3.1}$$

Note that the estimator (3.1) is constructed by splicing, at the point A, the SIFRAM and the GIFRAM of the restrictions of $\bar{F}_n$ to the intervals $[0, A]$ and $[A, \infty)$, respectively. It follows from (3.1) that $\hat{\bar{F}}_n(A) = \bar{F}_n(A)$, while $\hat{\bar{F}}_n(t) \geqslant \bar{F}_n(t)$ for $t < A$, and $\hat{\bar{F}}_n(t) \leqslant \bar{F}_n(t)$ for $t \geqslant A$. Thus, $\hat{\bar{F}}_n(t) \to 0$ as $t \to \infty$. Also, it follows easily from (3.1) that $\hat{\bar{F}}_n$ is nondecreasing and right-continuous. That $\hat{\bar{F}}_n$ is IFRA follows easily by noting that

$$\frac{-\ln \hat{\bar{F}}_n(t)}{t} = \begin{cases} \inf\limits_{t \leqslant s \leqslant A} \dfrac{-\ln \bar{F}_n(s)}{s}, & t < A, \\ \sup\limits_{A \leqslant s \leqslant t} \dfrac{-\ln \bar{F}_n(s)}{s}, & t \geqslant A, \end{cases}$$

is nondecreasing in t.

The main result will now be proven. The proof hinges on properties of convex functions and the following technical result.

Lemma 3.1. *Let h, g be bounded functions on the interval $[a,b]$. Then,*

$$\left| \inf_{a \leqslant y \leqslant b} h(y) - \inf_{a \leqslant y \leqslant b} g(y) \right| \leqslant \sup_{a \leqslant y \leqslant b} |h(y) - g(y)|, \tag{3.2}$$

and hence, also, $|\sup_{a \leqslant y \leqslant b} h(y) - \sup_{a \leqslant y \leqslant b} g(y)| \leqslant \sup_{a \leqslant y \leqslant b} |h(y) - g(y)|$.

Proof. Let $\{y_n\}$ be a sequence in $[a, b]$ for which $g(y_n) \leqslant \inf_{a \leqslant y \leqslant b} g(y) + 1/n$. Assume without loss of generality that $\inf_{a \leqslant y \leqslant b} h(y) - \inf_{a \leqslant y \leqslant b} g(y) = A > 0$. Then $h(y_n) \geqslant \inf_{a \leqslant y \leqslant b} h(y) = \inf_{a \leqslant y \leqslant b} g(y) + A \geqslant g(y_n) + A - 1/n$; that is, $h(y_n) - g(y_n) \geqslant A - 1/n$. It follows that $\sup_{a \leqslant y \leqslant b} |h(y) - g(y)| \geqslant A$, completing the proof. ∎

We now turn our attention to the main result of the paper.

THEOREM 3.2. *Let F be IFRA and F_n be an estimator of F. Let $\hat{\bar{F}}_n$ be defined as in* (3.1). *Then,*

$$\sup_t |\hat{\bar{F}}_n(t) - \bar{F}(t)| \leqslant \max\left\{\frac{1}{\bar{F}(A)}, \frac{1}{F(A)}\right\} \sup_t |\bar{F}_n(t) - \bar{F}(t)|.$$

Proof. Using Lemma 3.1 and the fact that F is IFRA, it is easy to see that

$$\sup_{t<A} |\hat{\bar{F}}_n(t) - \bar{F}(t)| = \sup_{t\leqslant A} | \sup_{t\leqslant s\leqslant A} (\bar{F}_n(s))^{t/s} - \sup_{t\leqslant s\leqslant A} (\bar{F}(s))^{t/s}|$$
$$\leqslant \sup_{t\leqslant A} \sup_{t\leqslant s\leqslant A} |(\bar{F}_n(s))^{t/s} - (\bar{F}(s))^{t/s}|.$$

Now, since $\bar{F}(A) \leqslant \bar{F}(s)$ for $s \in [t, A]$, it follows from the concavity of $x^{t/s}$, when $t \leqslant s$, that

$$\frac{(\bar{F}_n(s))^{t/s} - (\bar{F}(s))^{t/s}}{\bar{F}_n(s) - \bar{F}(s)} \leqslant \frac{(\bar{F}(A))^{t/s}}{\bar{F}(A)} = (\bar{F}(A))^{t/s-1}.$$

Therefore,

$$\sup_{t\leqslant A} \sup_{t\leqslant s\leqslant A} |(\bar{F}_n(s))^{t/s} - (\bar{F}(s))^{t/s}| \leqslant \sup_{t\leqslant A} \sup_{t\leqslant s\leqslant A} (\bar{F}(A))^{t/s-1} |\bar{F}_n(s) - \bar{F}(s)|$$
$$\leqslant \frac{1}{\bar{F}(A)} \sup_{s\leqslant A} |\bar{F}_n(s) - \bar{F}(s)|.$$

On the other hand,

$$\sup_{t\geqslant A} |\hat{\bar{F}}_n(t) - \bar{F}(t)| = \sup_{t\geqslant A} | \inf_{A\leqslant s\leqslant t} (\bar{F}_n(s))^{t/s} - \inf_{A\leqslant s\leqslant t} (\bar{F}(s))^{t/s}|$$
$$\leqslant \sup_{t\geqslant A} \sup_{A\leqslant s\leqslant t} |\bar{F}_n(s)^{t/s} - (\bar{F}(s))^{t/s}|.$$

Now, for $s \in [A, t]$, $\bar{F}(t) \leqslant \bar{F}(s) \leqslant \bar{F}(A)$. It then follows from the convexity of $x^{t/s}$, when $t \geqslant s$, that $[(\bar{F}_n(s))^{t/s} - (\bar{F}(s))^{t/s}]/[\bar{F}_n(s) - \bar{F}(s)] \leqslant [1 - (\bar{F}(A))^{t/s}]/[1 - \bar{F}(A)]$. Therefore,

$$\sup_{t\geqslant A} |\hat{\bar{F}}_n(t) - \bar{F}(t)| \leqslant \sup_{t\geqslant A} \sup_{A\leqslant s\leqslant t} \left\{\frac{1 - (\bar{F}(A))^{t/s}}{1 - \bar{F}(A)}\right\} |\bar{F}_n(s) - \bar{F}(s)|$$
$$\leqslant \frac{1}{1 - \bar{F}(A)} \sup_{t\geqslant A} |\bar{F}_n(s) - \bar{F}(s)|,$$

from which the result follows. ∎

The following corollary is immediate.

COROLLARY 3.3. *Let $\bar{F}_n$ denote the empirical survival function. Then $\hat{\bar{F}}_n$ defined by* (3.1) *is a strong uniform consistent estimator of F with optimal rate of convergence $O(n^{-1/2}(\log\log n)^{1/2})$.*

A closed form of $\hat{\bar{F}}_n$ useful for computational purposes is obtained from (3.1). Let k be the unique integer such that $X_{(k)} \leqslant A < X_{(k+1)}$. *Then,*

$$\hat{\bar{F}}_n(t) = \begin{cases} \max\left\{ \max_{\{i:\, t < X_i \leqslant A\}} \left(1 - \frac{(i-1)}{n}\right)^{t/X_i}, \left(1 - \frac{k}{n}\right)^{t/A} \right\}, & t < A, \\ \min\left\{ \min_{\{i:\, A < X_i \leqslant t\}} \left(1 - \frac{i}{n}\right)^{t/X_i}, \left(1 - \frac{k}{n}\right)^{t/A} \right\}, & t \geqslant A, \end{cases} \tag{3.3}$$

We now turn our attention to the censored case. Let $X_1, ..., X_n$ be a random sample from the IFRA distribution F. In the same set-up as in Csörgő and Horváth (1983), another random sample $Y_1, ..., Y_n$ with (left-continuous) distribution function H censors on the right the distribution F. As a result, the observations available consist of the pairs (X_j, δ_j), $1 \leqslant j \leqslant n$, where $Z_j = \min(X_j, Y_j)$ and δ_j is the indicator of the event $\{Z_j = X_j\}$. Let $\bar{F}_n$ be the Kaplan–Meier (1958) estimator of $\bar{F}$ defined by

$$\bar{F}_n(t) = \begin{cases} \prod_{1 \leqslant j \leqslant n:\, X_j < t} \left(\frac{n - N_{j:n} - 1}{n - N_{j:n}}\right)^{\delta_j}, & t \leqslant X_{(n)}, \\ 0, & t > X_{(n)}, \end{cases}$$

where $X_{(n)}$ denotes the nth order statistic, and $N_{j:n} = \sum_{k=1}^{n} I_{\{X_k < X_j\}}$.

Note that the distribution G of Z_i is given by $\bar{G}(t) = \bar{F}(t)\,\bar{H}(t)$ for each t. Csörgő and Horváth (1983) provide a nice treatment of the closeness of $\bar{F}_n$ to $\bar{F}$ as well as the rates of convergence of $\bar{F}_n$ to $\bar{F}$. For a distribution function F^*, let $T_{F^*} = \inf\{t\colon F^*(t) = 1\}$. Also define $d_n(\bar{F}, \hat{\bar{F}}_n) = \sup_{-\infty < t \leqslant T_F} |\bar{F}(t) - \hat{\bar{F}}_n(t)|$. The following corollary gives the strong uniform convergence properties of $\hat{\bar{F}}_n$ when $\bar{F}_n$ is the Kaplan–Meier estimator, and follows directly from Theorem 3.2 and Corollary 2(ii) in Csörgő and Horváth (1983).

COROLLARY 3.4. *Let $\bar{F}_n$ be the Kaplan–Meier estimator for the IFRA survival function $\bar{F}$. If $T_F \leqslant T_H \leqslant \infty$, then $\hat{\bar{F}}_n$ defined by* (3.1) *has the property that $d_n(\bar{F}, \hat{\bar{F}}_n)$ converges to zero with probability one.*

4. SIMULATIONS

Simulations were performed to compare the MSE behavior of the estimator (3.1) with the MSE of the empirical survival function and those estimators proposed by Wang (1987b) and Barlow *et al.* (1972). Five types of distributions were considered: the standard exponential; the Gamma distribution with unit scale and shape parameter equal to two; and the Weibull distributions with unit scale and shape parameters, $\alpha = 1.1$, 1.25, 1.5. These were selected since it is believed that the tail-heaviness of the underlying distribution has an effect on the behavior of Wang's estimator. That is, the lighter the tail of the underlying distribution, the lesser the problem Wang's estimator has in estimating the tail of the distribution.

The isotonic regression estimator was computed, as given on p. 258–259 of Barlow *et al.* (1972) with equal weights of $1/n$, and Wang's estimator was computed by using the following expression:

$$\hat{\bar{F}}(s) = \left\{ \max_{k \leqslant j \leqslant n} \left(1 - \frac{(j-1)}{n}\right)^{1/X_{(j)}} \right\}^{s}, \qquad X_{(k-1)} \leqslant s < X_{(k)}.$$

Although the strong uniform consistency of the estimators defined by (3.1) does not depend on the $0 < A < \infty$ selected, the mean-squared error properties of the estimators are clearly affected by the selection of A. Since the estimator $\hat{\bar{F}}_n$ agrees with $\bar{G}_n$ as defined by Wang (1987b) on $[0, A)$, it is clear that as A increases, the maximum MSE of $\hat{\bar{F}}_n$ on $[0, A)$ will increase since the estimator $\bar{G}_n$ does not perform well in the right tail of the distribution. Similarly, as A decreases, the maximum MSE of $\hat{\bar{F}}_n$ on (A, ∞) will increase. Simulations—the results of which are not presented

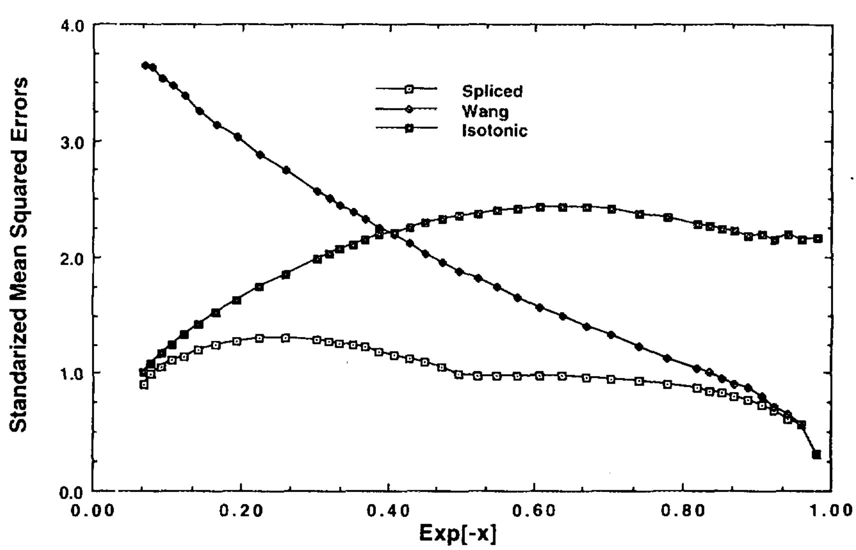

FIG. 1. Sample size is 30 from the exponential, 10,000 replicates.

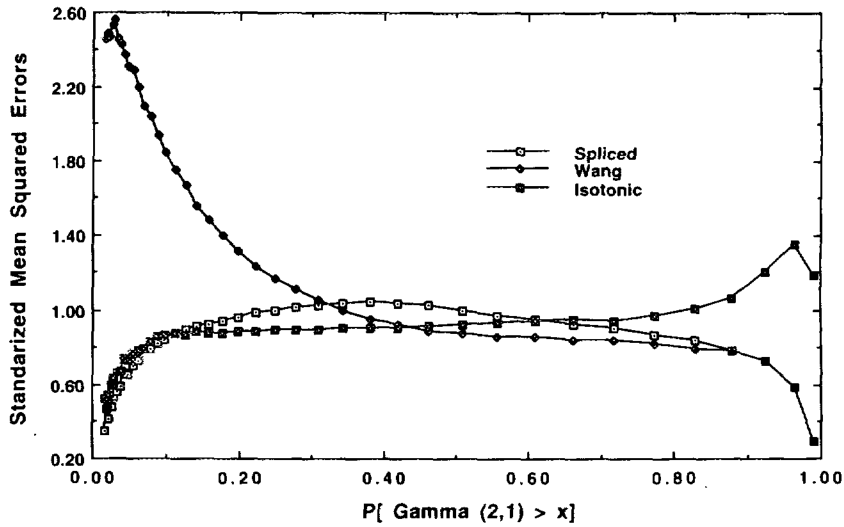

FIG. 2. Sample size is 30 from gamma (2, 1), 10,000 replicates.

here—were performed choosing A as the median of F. These simulations showed that the spliced estimator consistently beats Wang's estimator in the right tail, while also being superior to the isotonic regression estimator on the left tail. Clearly, in applications the value of the median of F is rarely known and hence a criterion for selecting A must be proposed. Since the median of F performs well as a choice for A, one possible criterion for

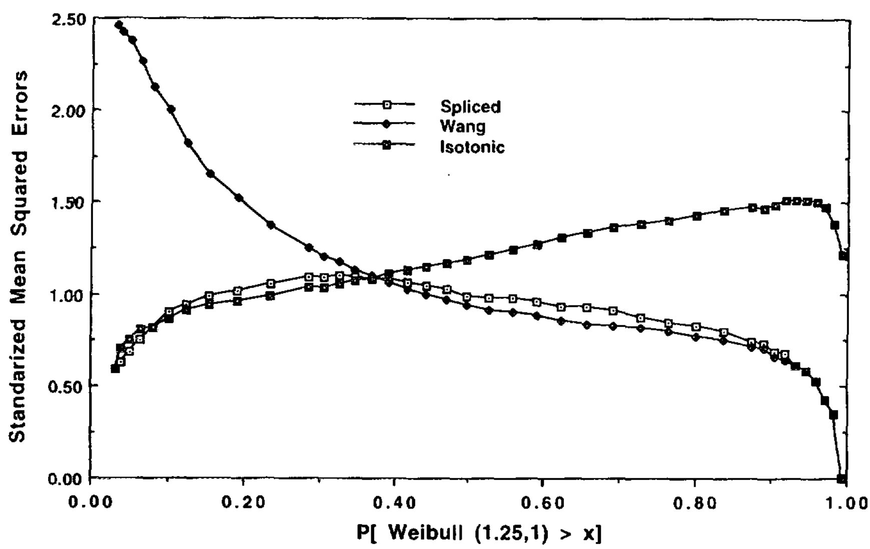

FIG. 3. Sample size is 30 from the Weibull (1.25,1), 10,000 replicates.

selecting A is to let A be the sample median. Simulations showed that the resulting estimator, although still better than Wang's estimator and the isotonic regression estimator in the right and left tails, respectively, when compared to the empirical survival function, showed a substantial increase in MSE.

An alternative criterion for selecting A is now proposed based on the intuitive notion that an IFRA estimator close to $\hat{\bar{F}}_n$ will behave, in MSE, similar to $\bar{F}_n$. Select A to minimize $\sup_t |\hat{\bar{F}}_n(t) - \bar{F}_n(t)|$. An easy argument shows that A must then be selected to be one of the order statistics. As the value of k such that $A = X_{(k)}$ minimizes $\sup_t |\hat{\bar{F}}_n(t) - \bar{F}_n(t)|$ may not be unique, we select k to be the smallest integer, $k \geqslant 2$, minimizing the expression

$$\sup_t |\hat{\bar{F}}_n(t) - \bar{F}_n(t)|$$

$$= \max \left\{ \max_{2 \leqslant i \leqslant k} \left\{ \max_{i-1 \leqslant j \leqslant k} (\bar{F}_n(X_{(j-1)}))^{X_{(i-1)}/X_{(j)}} - (\bar{F}_n(X_{(i-1)})) \right\}, \right.$$

$$\left. \max_{k+1 \leqslant i \leqslant n-1} \left\{ (\bar{F}_n(X_{(i)})) - \min_{k \leqslant j \leqslant i+1} (\bar{F}_n(X_{(j)}))^{X_{(i+1)}/X_{(j)}} \right\} \right\}.$$

The results of the simulations, presented in Figs. 1–3, were obtained with this particular selection of A. Sample sizes of 15 and 30 were generated with 10,000 replications. In all cases considered, similar results were obtained and Figs. 1–3 are a representative sample of those results.

In the figures, the mean square errors of the estimator defined by (3.1), and those of Wang and Barlow *et al.*, were standarized by dividing by $F(x)\,\bar{F}(x)/n$, the mean square error of the empirical survival function. One notes from Figs. 1–3 that, as expected, the spliced estimator consistently beats Wang's estimator in the right tail, although the difference in MSE decreases with a decrease in the heaviness of the tail of the underlying distribution. Also, the spliced estimator behaves better, in terms of MSE, than the isotonic regression estimator when estimating the left tail.

5. Asymptotic Confidence Intervals and Bands

The development of confidence intervals for $\bar{F}(t_0)$ and confidence bands for $\bar{F}$ based on $\hat{\bar{F}}_n$ appears to be a challenging technical problem in its full generality. In this section, asymptotic confidence intervals for $\bar{F}(t_0)$ will be presented when $t_0 \leqslant A$ and $\bar{F}$ is assumed to satisfy the conditions of Theorem 1 in Wang (1987b). Under similar conditions, confidence bands for $\bar{F}$ on compact intervals are also possible. These results are stated in the following theorem.

THEOREM 5.1. *Let $\bar{F}$ satisfy the assumptions of Theorem* 1 *of Wang* (1987b), *and let $\hat{\bar{F}}_n$ be defined by* (3.1). *Let $\|f\|_a^b = \sup_{a \leqslant x \leqslant b} |f(x)|$. Then*

(i) $\sqrt{n}\,(\hat{\bar{F}}_n(t_0) - \bar{F}(t_0)) \xrightarrow{\mathscr{L}} N(0, \bar{F}(t_0)\,F(t_0))$ *for* $t_0 \leqslant A$.

(ii) $\sqrt{n}\,\|\hat{\bar{F}}_n - \bar{F}\|_{t_1}^{t_2} \xrightarrow{\mathscr{L}} \|\mathscr{U}\|_{F(t_1)}^{F(t_2)}$, *where* $t_1 < t_2 \leqslant A$, *and* $\mathscr{U}$ *is a Brownian bridge.*

Proof. Note that $\hat{\bar{F}}_n(t) = 1 - G_n(t)$ for $t \leqslant A$, where $G_n(t)$ is as defined in Wang (1987b). It follows from Corollary 1 of Wang (1987b) that $\|\hat{\bar{F}}_n - \bar{F}_n\|_a^b \xrightarrow{P} 0$ for $a < b \leqslant A$ and hence, $\sqrt{n}\,(\hat{\bar{F}}_n(t) - \bar{F}_n(t)) \xrightarrow{P} 0$, for $t \leqslant A$. Since, $\sqrt{n}\,(\hat{\bar{F}}_n(t) - \bar{F}(t)) = \sqrt{n}\,(\hat{\bar{F}}_n(t) - \bar{F}_n(t)) + \sqrt{n}\,(\bar{F}_n(t) - \bar{F}(t))$, (i) follows immediately. A similar argument yields (ii), after noting that $\|\cdot\|_a^b$ is a $\|\cdot\|_{-\infty}^{\infty}$-continuous function. ▌

The reader is referred to Manija (1949) for the details of the asymptotic distribution of $\|\sqrt{n}\,(\bar{F}_n - \bar{F})\|_a^b$.

6. AN EXAMPLE

To illustrate the effect of transforming the empirical survival function into an IFRA survival function, the data given by Pike (1966) have been used. The date represent the times in days, after the start of the experiment,

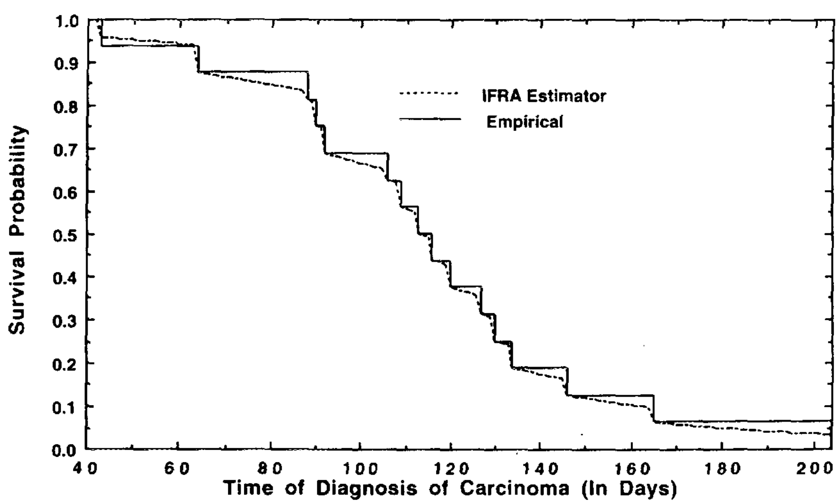

FIG. 4. The IFRA estimator and the empirical survival function for the carcinoma data.

at which carcinoma was diagnosed for a group of rats exposed to the carcinogen DMBA and are given as follows: 43, 64, 88, 90, 92, 106, 109, 113, 116, 120, 127, 130, 134, 146, 165, and 204. Pike (1966) fitted the data by a Weibull distribution with shape parameter 3. Figure 4 illustrates the estimator (3.1) with A selected to minimize $\sup_t |\hat{\bar{F}}_n(t) - \bar{F}_n(t)|$ which in this case gives $A = 64$; the empirical distribution function is also shown for comparison purposes. As expected, $\hat{\bar{F}}_n(t) \geqslant \bar{F}_n(t)$ for $t < A$ while $\hat{\bar{F}}_n(A) = \bar{F}_n(A)$ and $\hat{\bar{F}}_n(t) \leqslant \bar{F}_n(t)$ for $t > A$.

References

[1] A-Hameed, M. S., and Proschan, F. (1973). Nonstationary shock models. *Stochastic Process. Appl.* **1** 383–404.

[2] A-Hameed, M. S., and Proschan, F. (1975). Shock models with underlying birth process. *J. Appl. Probab.* **12** 18–28.

[3] Barlow, R. E., Bartholomew, D. J., Bremmer, J. N., and Brunk, H. D. (1972). *Statistical Inference Under Order Restrictions.* Wiley, New York.

[4] Barlow, R. E., and Scheuer, E. M. (1971). Estimation from accelerated life tests. *Technometrics* **13** 145–169.

[5] Birnbaum, Z. W., Esary, J. D., and Marshall, A. W. (1966). Stochastic characterization of wearout for components and systems. *Ann. Math. Statist.* **37** 816–825.

[6] Boyles, R. A.. (1981). *Statistical Inference for Two Models in Reliability*, Ph.D. dissertion. University of California, Davis.

[7] Boyles, R. A., Marshall, A. W., and Proschan, F. (1985). Inconsistency of maximum likelihood estimator of distributions having increasing faillure rate average. *Ann. Statist.* **13** 413–418.

[8] Boyles, R. A., and Samaniego, F. J. (1984). Estimating a survival curve when new is better than used. *Oper. Res.* **32** 732–740.

[9] Csörgő, S., and Horváth, L. (1983). The rate of strong uniform consistency for the product-limit estimator. *Z. Wahrsch. verw. Gebiete* **62** 411–426.

[10] Dvoretzky, A., Kiefer, J., and Wolfowitz, J. (1956). Asymptotic minimax character of the sample distribution function and the classical multinomial estimator. *Ann. Math. Statist.* **27** 642–669.

[11] Esary, J. D., Marshall, A. W., and Proschan, F. (1973). Shock models and wear processes. *Ann. Probab.* **1**, 627–629.

[12] Esary, J. D., and Marshall, A. W. (1974). Families of components and systems exposed to a compound poisson damage process. In *Reliability and Biometry Statistical Analaysis of Lifelength* (F. Proschan and R. J. Serfling, Eds.), pp. 31–46, SIAM, Philadelphia.

[13] Kiefer, J., and Wolfowitz J. (1976). Asymptotically minimax esimation of concave and convex distribution functions. *Z. Wahrsch. verw. Gebiete* **34**, 73–85.

[14] Manija, G. M. (1949). Generalization of the criterion of A. Kolmogorov for an estimate of the distribution law from the empirical data. *Dokl. Akad. Nauk USSR* **69** (No. 4) 495–497. [in Russian]

[15] Marshall, A. W., and Proschan, F. (1965). Maximum likelihood estimation for distributions with monotone failure rate. *Ann. Math. Statist.* **36** 69–77.

[16] Pike, M. C. (1966). A method of analysis of a certain class of experiments in carcinogenesis. *Biometrics* **22** 142–161.

[17] ROJO, J. (1988). *On the Concept of Tail-Heaviness*, Tech. Rep. 175, Oct. University of California, Berkeley.

[18] ROJO, J. (1992). A pure-tail ordering based on the ratio of the quantile functions. *Ann. Statist.* **20** 2100–2110.

[19] ROJO, J., AND SAMANIEGO, F. (1990). On estimating a survival curve subject to a uniform stochastically ordering constraint *J. Amer. Statist. Assoc.* **88** 566–572.

[20] WANG, J. L. (1987a). Estimating IFRA and NBU survival curves based on censored data. *Scand. J. Statist.* **14** 199–210.

[21] WANG, J. L. (1987b). Estimators of a distribution function with increasing failure rate average. *J. Statist. Plann. Inference* **16** 415–427.

[22] WANG, J. L. (1988). Optimal esitmation of star shaped distribution functions. *Statist. Decisions* **6** 21–32.