

Portland State University PDXScholar

Mathematics and Statistics Faculty Publications and Presentations Fariborz Maseeh Department of Mathematics and Statistics

1-1-2000

Estimation of a Monotone Mean Residual Life

Subhash C. Kochar

Portland State University, kochar@pdx.edu

Hari Mukerjee

Francisco J. Samaniego

Let us know how access to this document benefits you.

Follow this and additional works at: http://pdxscholar.library.pdx.edu/mth_fac

 Part of the [Mathematics Commons](#)

Citation Details

Subhash, K. Mukerjee, H. and Samaniego, F. Estimation of a monotone mean residual life. *The Annals of Statistics* , Vol. 28, No. 3 (Jun., 2000), pp. 905-921.

This Article is brought to you for free and open access. It has been accepted for inclusion in Mathematics and Statistics Faculty Publications and Presentations by an authorized administrator of PDXScholar. For more information, please contact pdxscholar@pdx.edu.

ESTIMATION OF A MONOTONE MEAN RESIDUAL LIFE

BY SUBHASH C. KOCHAR, HARI MUKERJEE AND FRANCISCO J. SAMANIEGO

*Indian Statistical Institute, Wichita State University
and University of California, Davis*

In survival analysis and in the analysis of life tables an important biometric function of interest is the life expectancy at age x , $M(x)$, defined by

$$M(x) = E[X - x | X > x],$$

where X is a lifetime. M is called the mean residual life function. In many applications it is reasonable to assume that M is decreasing (DMRL) or increasing (IMRL); we write decreasing (increasing) for nonincreasing (nondecreasing). There is some literature on empirical estimators of M and their properties. Although tests for a monotone M are discussed in the literature, we are not aware of any estimators of M under these order restrictions. In this paper we initiate a study of such estimation. Our projection type estimators are shown to be strongly uniformly consistent on compact intervals, and they are shown to be asymptotically “root- n ” equivalent in probability to the (unrestricted) empirical estimator when M is strictly monotone. Thus the monotonicity is obtained “free of charge”, at least in the asymptotic sense. We also consider the nonparametric maximum likelihood estimators. They do not exist for the IMRL case. They do exist for the DMRL case, but we have found the solutions to be too complex to be evaluated efficiently.

1. Introduction. The mean residual life (MRL) of a unit or a subject at age x is the average remaining life among those population members who have survived until time x . If the lifelength of the population is described by a random variable X with survival function (s.f.) S , defined by $S(x) = P(X > x)$, then the mean residual life function is defined by

$$(1.1) \quad M(x) = E[X - x | X > x] = I[S(x) > 0] \int_x^\infty S(u) du / S(x).$$

A distribution is characterized by its MRL by the relation [Guess and Proschan (1988)]

$$(1.2) \quad S(x) = \frac{M(0)}{M(x)} \exp \left\{ - \int_0^x \frac{1}{M(u)} du \right\} I[M(x) > 0].$$

Note that $M(0)$ is just the mean of X whose existence is assumed throughout. Like the failure rate function, the MRL describes a conditional concept of aging; however, the MRL is more intuitive, especially in the health sciences. The review article by Guess and Proschan (1988) gives a nice summary of the theory of MRL and an extensive bibliography.

Received October 1998; revised March 2000.

AMS 1991 subject classifications. Primary 62P10, 62G05, 62E10.

Key words and phrases. Mean residual life, order restricted inference, asymptotic theory.

Let $0 \equiv X_0 \leq X_1 \leq X_2 \leq \dots \leq X_n$ be the order statistics from a random sample from S with support $[0, T]$ for some finite T , or $[0, \infty)$. Let S_n denote the empirical s.f. Yang (1978) considered an empirical estimate of M using

$$(1.3) \quad \widehat{M}_n(x) = I(x < X_n) \int_x^\infty S_n(u) du / S_n(x).$$

Assuming S to be (absolutely) continuous, she showed that \widehat{M}_n is strongly uniformly consistent on $[0, b]$ for any $b < T$ and that $\sqrt{n}(\widehat{M}_n - M)$ converges weakly to a Gaussian process. Hall and Wellner (1979) and Csörgő and Zitikis (1996) have strengthened some of Yang's results.

In many applications it is reasonable to assume that the life system is monotonically degenerating or improving with age. This concept has been modeled several ways, of which increasing (IFR) and decreasing (DFR) failure rates are probably the most studied; we write increasing (decreasing) for nondecreasing (nonincreasing) throughout. The somewhat weaker version of decreasing (increasing) MRL, DMRL (IMRL), which is implied by IFR (DFR), is perhaps more clear conceptually and is easier to explain to the user. New better (worse) than used in expectation, NBUE (NWUE), is an even weaker concept. NBUE and NWUE correspond to $M(x) \leq M(0)$ and $M(x) \geq M(0) \forall x$, respectively. There are other modified concepts like DIMRL (IDMRL) where the MRL is initially decreasing (increasing) and then increasing (decreasing). Gertsbach and Kordonskiy (1969) have shown that the lognormal distribution is DIMRL. However, there are no well known parametric families that are DMRL (IMRL) (for some values of a parameter) but not IFR (DFR). Hollander and Proschan (1975), Guess, Hollander and Proschan (1986), Aly (1990), Hawkins, Kochar and Loader (1992) and Lim and Park (1998) have considered tests for MRL's under various monotonicity restrictions. However, estimation of a MRL under order restrictions does not appear to have been considered in the literature. In this paper we initiate a study of such estimation procedures. Although several modifications may be considered, our basic estimators are projection type estimators that proved to have nice properties in several restricted estimation problems, improving on the nonparametric maximum likelihood estimators (NPMLE's) [see, e.g., Rojo and Samaniego (1991, 1993), Mukerjee (1996), Rojo and Ma (1996) and Rojo (1995)]. In Section 2 we describe our estimators and prove their strong uniform consistency. In Section 3 we illustrate our procedure for some data from Bjerkedal (1960). In Section 4 we consider the asymptotic distributions of our estimators. It is shown that

$$\sup_{x \leq b} \sqrt{n} |\widehat{M}_n(x) - M_n^*(x)| \xrightarrow{p} 0$$

as $n \rightarrow \infty$ for any $b < T$ if M is strictly monotone on $[0, b]$; here \widehat{M}_n is the empirical estimator and M_n^* is our monotone estimator, both based on a sample of size n . Thus all of the asymptotic distributional results for \widehat{M}_n hold for M_n^* , and we get the monotonicity "free of charge". The results and the method of proofs are similar to those for the estimation of an IFR distribution function (d.f.) in Wang (1986), which in turn are similar to those in the estimation of a

concave or a convex d.f. in Kiefer and Wolfowitz (1976). It may be possible to extend our results from $[0, b]$ to $[0, \infty)$ using weighted empiricals [Hall and Wellner (1979) and Csörgö and Zitikis (1996)]. However, we have not considered those extensions here. In Section 5 we present some modest simulation results for the DMRL case, which is of primary interest to us. Since we do not know of any suitable parametric family of d.f.'s that are DMRL but not IFR, we have considered the uniform and an IFR Weibull distribution. The MSE of M_n^* is uniformly smaller than that of \widehat{M}_n , but not by much, as is to be expected from their asymptotic "root- n " equivalence in probability as proven in Section 4. We have also considered the exponential distribution where M is a constant. Since M is not strictly decreasing the results of Section 4 do not hold, but it is interesting to note that the MSE of M_n^* is still uniformly smaller than that of \widehat{M}_n although the bias is considerably larger, especially in the right tail, where M_n^* achieves its greatest gains in terms of MSE! In Section 6 we consider the NPMLÉ's. The NPMLÉ does not exist for the IMRL case. It does exist for the DMRL case, but the computation appears to be intractable. In Section 7 we present some concluding remarks and directions of future research.

2. Estimators and consistency. One property of a MRL, M , is that $M(x) + x$ is increasing for all x so that $M' \geq -1$ whenever it exists, where $M'(x) = dM(x)/dx$ [see, e.g., Guess and Proschan (1988)]. Thus, in the DMRL case M must be continuous, and from the definition (1.1) of a MRL, the corresponding s.f., S , cannot have a jump except possibly at T , the right end point of its support. This need not be true in the IMRL case. Using the inversion formula (1.2) it can be seen that S is flat in any interval where $M' = -1$ and that S corresponds to a segment of a shifted exponential s.f. where $M' = 0$ with a mean equal to the local value of M . We note that the empirical estimator \widehat{M}_n in (1.3) consists of line segments with slope equal to -1 with a jump up at each order statistic (which gives rise to a rather ragged estimator). Our estimators simply utilize the fact that M is a DMRL (IMRL) iff $M(x) = \inf_{y \leq x} M(y)$ ($M(x) = \sup_{y \leq x} M(y)$). These estimators are given by

$$(2.1) \quad M_n^*(x) = I(x < X_n) \inf_{y \leq x} \widehat{M}_n(y), \quad \text{DMRL}$$

and

$$(2.2) \quad M_n^{**}(x) = I(x < X_n) \sup_{y \leq x} \widehat{M}_n(y), \quad \text{IMRL.}$$

Note that M_n^* (M_n^{**}) is the largest (smallest) decreasing (increasing) function that lies below (above) the empirical \widehat{M}_n (see Figure 1). M_n^* is a continuous function formed of line segments that are flat (corresponding to an exponential s.f.) or with a slope of -1 (corresponding to a flat s.f.). M_n^{**} consists of an increasing step function. Other *ad hoc* estimators could be defined using the same principle, for example, $\widetilde{M}_n(x) = \widehat{M}_n(0) \wedge \sup_{y \geq x} \widehat{M}_n(y)$, or some convex

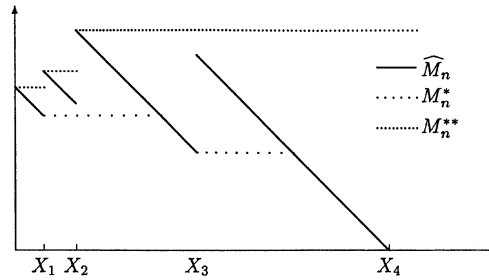


FIG. 1. Illustration of \widehat{M}_n , M_n^* and M_n^{**} .

combination of this and M_n^* for the DMRL case. Since the number of observations remaining in the right tail is small, this produces large variabilities which is borne out by our simulations. Computations of the estimators are quite simple. For example, for M_n^* , we first find

$$Y_0 = 0 < Y_1 = X_1 < Y_2 < Y_3 < \dots < Y_k = X_n,$$

where $\widehat{M}_n(Y_j^-) > \widehat{M}_n(Y_{j+1}^-)$ with $\widehat{M}_n(Y_j^-) \leq \widehat{M}_n(X_i^-)$ for all $Y_j < X_i < Y_{j+1}$. Then, $M_n^*(0) = \widehat{M}_n(0)$, and, for $Y_j < x \leq Y_{j+1}$, $0 \leq j \leq k - 1$,

$$M_n^*(x) = \min\{\widehat{M}_n(Y_j^-), \widehat{M}_n(Y_{j+1}^-) + Y_{j+1} - x\}.$$

To illustrate, in Figure 1, $k = 3$, $Y_0 = 0$, $Y_1 = X_1$, $Y_2 = X_3$ and $Y_3 = X_4$. It can be seen that as x moves to the left from Y_2 to Y_1 , $M_n^*(x)$ increases from $\widehat{M}_n(Y_2^-)$ by $Y_2 - x$ until it attains its maximum value of $\widehat{M}_n(Y_1^-)$ in the interval.

2.1. *Consistency.* Yang (1978) has shown that \widehat{M}_n is strongly uniformly consistent on $[0, b]$ for any $b < T$, where T is the right endpoint of the support of S , using only the consistency of S_n . The same holds true for M_n^* and M_n^{**} from this result and the triangle inequality of the sup-norm [see Lemmas 1 and 2, Rojo and Samaniego (1993).] under the monotonicity assumptions

$$(2.3) \quad \left| \inf_{y \leq x} \widehat{M}_n(y) - \inf_{y \leq x} M(y) \right| \leq \sup_{y \leq x} |\widehat{M}_n(y) - M(y)|,$$

which proves consistency for M_n^* . The proof for M_n^{**} is similar.

3. An example. Bjerkedal (1960) reports on two studies of survival times (in days) of guinea pigs infected with different dosages of tubercle bacilli. Although at lower dosages the distributions seem to be DMRL, at higher dosages they appear to be DIMRL (or, perhaps, DIDMRL) due to developed resistance to infection. However, this is not always easy to detect from the empirical MRL. Since intuitively we would expect the distribution to be DMRL, the

TABLE 1
Estimates of mean residual life

X_i	$\widehat{M}_n(X_i^-)$	$M_n^*(X_i)$	X_i	$\widehat{M}_n(X_i^-)$	$M_n^*(X_i)$	X_i	$\widehat{M}_n(X_i^-)$	$M_n^*(X_i)$
0	175.3	175.3	120	101.1	94.7	202	94.6	90.0
22	153.3	153.3	121	102.3	94.7	213	89.9	89.9
33	144.4	144.4	122	106.3	94.7	215	90.0	89.9
44	135.5	135.5	124	109.3	94.7	216	93.6	89.9
56	125.4	125.4	130	105.9	94.7	222	92.5	89.9
59	124.3	124.3	134	109.3	94.7	230	89.6	89.6
72	113.1	113.1	136	105.5	94.7	231	93.9	89.6
74	112.9	112.9	139	104.7	94.7	240	90.8	89.6
77	111.6	111.6	144	102.5	94.7	245	91.8	89.6
92	98.4	98.4	146	103.2	94.7	251	92.4	89.6
93	98.9	98.4	153	99.1	94.7	253	97.5	89.6
94	99.5	98.4	159	96.0	94.7	254	104.6	89.6
100	95.1	95.1	160	95.0	94.7	278	101.5	89.6
102	96.4	95.1	163	97.8	94.7	293	97.8	89.6
105	95.0	95.0	168	99.1	94.7	327	76.0	76.0
107	94.7	94.7	171	99.4	94.7	342	71.9	71.9
108	97.1	94.7	172	101.8	94.7	347	78.8	71.9
109	101.7	94.7	176	101.4	94.7	361	80.6	71.9
112	100.7	94.7	183	98.2	94.7	402	59.8	59.8
113	101.7	94.7	195	90.0	90.0	432	49.7	49.7
115	101.8	94.7	196	92.6	90.0	458	48.5	48.5
116	102.9	94.7	197	95.4	90.0	555	0.0	0.0

constrained estimator could be used as a data analytic tool for detecting deviations from it. Thus we consider regimen 4.3 in Study M under the assumption that the MRL is decreasing, both for the purposes of illustration and to demonstrate this data analytic value. The observations are discrete with 2 deaths on days 100, 107, 122, 163 and 254, and 3 deaths on day 108. There were 72 subjects with no censoring. The results obtained for the empirical and the constrained estimates of the MRL are given in Table 1. The estimate $\widehat{M}_n(X_i) = (n - i + i)/(n - i)\widehat{M}_n(X_i^-)$, $n = 72$, $1 \leq i \leq 71$.

The data seem to indicate that initially the distribution is DMRL, but some resistance might have developed among those who survived the first 100 days or so, indicating a slightly increasing MRL, and then displaying DMRL behavior after 150 days or so. It should be noted that our DMRL and IMRL estimators could be employed on some given intervals only, if the order restriction applied only to these intervals, and the consistency results will still hold if the assumptions are correct.

4. Asymptotic results. Let $b < T$ be fixed, and let $\|\cdot\|_0^b$ denote the sup-norm on $[0, b]$. To simplify the notation we write $\|\cdot\|$ for $\|\cdot\|_0^b$ throughout this section. We prove the following theorem of asymptotic equivalence of \widehat{M}_n and M_n^* . An exactly parallel result holds for the equivalence of \widehat{M}_n and M_n^{**} when M is increasing; however, the assumption of a density, that always exists for the DMRL case, needs to be added for this IMRL case.

THEOREM 4.1. *Assume that*

$$(4.1) \quad M'(x) \text{ exists and } M'(x) \leq -c_1 \forall x \in [0, b] \text{ for some } c_1 > 0,$$

$$(4.2) \quad M''(x) \text{ exists and } \|M''\| \leq c_2 < \infty$$

and

$$(4.3) \quad EX^r < \infty \text{ for some } r > 2.$$

Then

$$\sqrt{n} \|M_n^* - \widehat{M}_n\| \xrightarrow{p} 0.$$

Assumption (4.2) is similar to the assumption of uniform convexity of the cumulative hazard function in Wang (1986).

The idea of the proof is to first construct a continuous piecewise linear version of \widehat{M}_n on $[0, b]$, then show that it is eventually decreasing *a.s.*, and that both M_n^* and \widehat{M}_n are close to it on $[0, b]$ in an appropriate sense. For each n let k_n be an integer, $k_n \uparrow \infty$, and let $\Delta_n = b/k_n$. If $k_n = [n^\delta]$ is chosen, then it will be shown that we need $1/4 < \delta < 1/2$. Let

$$\alpha_j^n = jb/k_n = j\Delta_n, \quad j = 0, 1, \dots, k_n.$$

Define the linear interpolation of any function h on $[0, b]$ by

$$L_n h(\alpha_j^n) = h(\alpha_j^n), \quad j = 0, 1, \dots, k_n$$

and

$$L_n h(x) = h(\alpha_j^n) + [h(\alpha_{j+1}^n) - h(\alpha_j^n)](x - \alpha_j^n)/\Delta_n \\ \text{for } \alpha_j^n < x < \alpha_{j+1}^n, \quad j = 0, 1, \dots, k_n - 1.$$

We now prove the following propositions leading to the proof of Theorem 4.1.

PROPOSITION 4.1. *Let $A_n = \{L_n \widehat{M}_n \text{ is strictly decreasing on } [0, b]\}$. Then, under assumption (4.3), if $\log \log n = o(n\Delta_n^2)$ then*

$$(4.4) \quad P[\underline{\lim}_n A_n] = 1.$$

PROOF. Hall and Wellner (1979) show that under the assumption $EX^r < \infty$ for some $r > 2$,

$$\sup_{x \leq b_n} \sqrt{n} |\widehat{M}_n(x) - M(x)| S_n(x) / (\log \log n)^{1/2} = O_{\text{a.s.}}(1),$$

where b_n is any increasing sequence with $S_n(b_n) \rightarrow 0$ and $nS(b_n)/(\log \log n)^{1/2} \rightarrow 0$. Since $S(b) > 0$ and $S_n \xrightarrow{\text{unif}} S$ *a.s.*, we have

$$\sqrt{n} \|\widehat{M}_n - M\| / (\log \log n)^{1/2} = O_{\text{a.s.}}(1)$$

Since $L_n \widehat{M}_n(a_j^n) = \widehat{M}_n(a_j^n) \forall j$ and n ,

$$\sqrt{n} |\widehat{M}_n(a_j^n) - M(a_j^n)| / (\log \log n)^{1/2} = O_{\text{a.s.}}(1)$$

uniformly in $j \leq k_n$ from above, and

$$M(a_j^n) - M(a_{j+1}^n) \geq c_1 \Delta_n \forall j \leq k_n - 1$$

by assumption (4.8), we have

$$\widehat{M}_n(a_j^n) - \widehat{M}_n(a_{j+1}^n) \geq c_1 \Delta_n + O_{\text{a.s.}}((\log \log n/n)^{1/2}) \geq c_3 \Delta_n \text{ a.s.}$$

for all large n and $j \leq k_n - 1$ for some $c_3 > 0$ if $(\log \log n/n)^{1/2} = o(\Delta_n)$. This completes the proof of (4.4) using the piecewise linearity of $L_n \widehat{M}_n$. \square

PROPOSITION 4.2. *Let $B_n = \{\|M_n^* - L_n \widehat{M}_n\| \leq \|\widehat{M}_n - L_n \widehat{M}_n\|\}$. Then, under the assumptions of Proposition 4.1,*

$$(4.5) \quad P[\lim_n B_n] = 1.$$

PROOF. Since $L_n \widehat{M}_n$ is decreasing on $[0, b]$ a.s. for all n sufficiently large under the assumptions of Proposition 4.1, the result is immediate from the triangular inequality (2.3) of the sup-norm. \square

Now Proposition 4.2 implies that

$$(4.6) \quad \|M_n^* - \widehat{M}_n\| \leq \|M_n^* - L_n \widehat{M}_n\| + \|\widehat{M}_n - L_n \widehat{M}_n\| \leq 2\|\widehat{M}_n - L_n \widehat{M}_n\| \text{ a.s.}$$

for all large n under the assumptions of Proposition 4.1. In order to bound the last expression, we first establish an analytic bound for $\|M - L_n M\|$.

PROPOSITION 4.3. *Under assumption (4.2),*

$$(4.7) \quad \|M - L_n M\| \leq c_2 \Delta_n^2 \quad \forall n.$$

PROOF. For any fixed n and $0 \leq j \leq k_n - 1$ define

$$g(x) = M(x + a_j^n) - L_n M(x + a_j^n) \quad \text{for } x \in [0, \Delta_n],$$

and note that $g''(x) = M''(x + a_j^n)$ for $x \in (0, \Delta_n)$. Now $g(0) = g(\Delta_n) = 0$. By Taylor expansion and using our assumptions,

$$g(x) = g'(0^+)x + g''(\xi_x)x^2/2$$

for some $0 < \xi_x < x$. Using $g(\Delta_n) = 0$, we get

$$\begin{aligned} & g'(0^+)\Delta_n + g''(\xi_{\Delta_n})\Delta_n^2/2 \\ &= 0 \implies g'(0^+) \\ &= -g''(\xi_{\Delta_n})\Delta_n/2 \implies |g(x)| \\ &= |-g''(\xi_{\Delta_n})\Delta_n x/2 + g''(\xi_x)x^2/2| \leq c_2 \Delta_n^2 \quad \forall x, \end{aligned}$$

which proves (4.7). \square

PROPOSITION 4.4. Under the assumptions of Theorem 4.1, if $n^{1/4} = o(k_n)$, then

$$(4.8) \quad \sqrt{n} \|\widehat{M}_n - L_n \widehat{M}_n\| \xrightarrow{p} 0.$$

PROOF. Let $Z_n = \sqrt{n}(\widehat{M}_n - M)$ on $[0, \infty)$ and $Z(x) = (\sigma(0)/S(x))B(U(x))$, where

$$\sigma^2(x) = \text{Var}(X - x | X > x) = I[x < T] \left[\int_{(x, \infty)} (t - x)^2 f(t) dt / S(x) - M^2(x) \right],$$

$U(x) = S(x)\sigma^2(x)/\sigma^2(0)$, and B is a standard Brownian motion. Hall and Wellner (1979) have shown that, under assumption (4.3):

- (i) U is a s.f.;
- (ii) Z is a mean-zero gaussian process on $[0, T)$ with covariance function $\text{Cov}[Z(x), Z(y)] = \sigma^2(y)/S(x)$ for $0 \leq x \leq y < T$, when $EX^2 < \infty$; and
- (iii) $Z_n \implies Z$ weakly on $[0, T)$ in $D([0, T))$ in the Skorohod J_1 topology. Since S is continuous on $[0, T)$, Z has a.s. continuous paths.

We now proceed as in the proofs of Lemma 6 and Proposition 2 in Wang (1986). From the tightness conditions on the Skorohod topology we have that for every $\epsilon > 0$, $\exists \delta > 0$ and $n_0 \in \mathcal{N}$ such that

$$(4.9) \quad P[\sup\{|Z_n(t) - Z_n(s)| : |t - s| < \delta, t, s \in [0, b]\} > \epsilon] < \epsilon \quad \forall n \geq n_0.$$

Let $\epsilon > 0$ be arbitrary. Choose δ to satisfy (4.9). Since $k_n \uparrow \infty$, $\exists n_1 \in \mathcal{N}$ such that $\Delta_n < \delta \forall n \geq n_1$. Assume that $n \geq n_0 \vee n_1$.

Define the piecewise shift transformation V_n by $V_n M(t) = M(t) + [\widehat{M}_n(a_j^n) - M(a_j^n)]$ for $a_j^n \leq t < a_{j+1}^n$, $j \leq k_n - 1$, and $V_n M(b) = \widehat{M}_n(b)$ (see Figure 2). Note that $V_n M(a_j^n) = \widehat{M}_n(a_j^n)$, $j \leq k_n$. Consider

$$\sqrt{n}[\widehat{M}_n(t) - V_n M(t)] = \sqrt{n}\{[\widehat{M}_n(t) - M(t)] - [\widehat{M}_n(a_j^n) - M(a_j^n)]\} = Z_n(t) - Z_n(a_j^n)$$

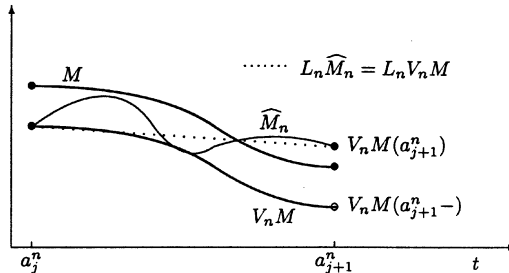


FIG. 2. The operators L_n and V_n .

for $\alpha_j^n \leq t < \alpha_{j+1}^n$. Then

$$\begin{aligned}
 \sqrt{n} \|\widehat{M}_n - V_n M\| &= \max_{0 \leq j \leq k_n - 1} \sup_{\alpha_j^n \leq t < \alpha_{j+1}^n} |Z_n(t) - Z_n(\alpha_j^n)| \\
 (4.10) \qquad &\leq \sup_{|t-s| < \delta} |Z_n(t) - Z_n(s)| \implies \sqrt{n} \|\widehat{M}_n - V_n M\| \xrightarrow{p} 0
 \end{aligned}$$

by (4.9). Note that $L_n \widehat{M}_n(\alpha_j^n) = \widehat{M}_n(\alpha_j^n) = V_n M(\alpha_j^n) \forall j \leq k_n$. Since L_n is piecewise linear and $V_n M(\alpha_j^n) = \widehat{M}_n(\alpha_j^n) \forall j$, $L_n \widehat{M}_n(t) = L_n V_n M(t) \forall t \leq b$. Hence,

$$\begin{aligned}
 \widehat{M}_n(t) - L_n \widehat{M}_n(t) &= [\widehat{M}_n(t) - V_n M(t)] + [V_n M(t) - L_n V_n M(t)] \\
 &\quad + [L_n V_n M(t) - L_n \widehat{M}_n(t)] \\
 &= [\widehat{M}_n(t) - V_n M(t)] + [V_n M(t) - L_n V_n M(t)] \\
 &= [\widehat{M}_n(t) - V_n M(t)] + [M(t) - L_n M(t)],
 \end{aligned}$$

the last equality following from the fact that $V_n M$ is piecewise shifted M . If $n^{1/4} = o(k_n)$ then $\sqrt{n} \|M - L_n M\| \rightarrow 0$ by (4.7). The proof of (4.8) then follows from (4.7), (4.10) and the above string of equalities. \square

PROOF OF THEOREM 4. The proof of the theorem is now immediate from (4.6) and (4.8) if we choose k_n such that $n^{1/4} = o(k_n)$ and $k_n = o((n/\log \log n)^{1/2})$. \square

The key to the proof of the theorem is the observation that the linearized version of \widehat{M}_n , $L_n \widehat{M}_n$, is eventually decreasing *a.s.* under our assumptions, so that if \widehat{M}_n is sufficiently close to it, so will be our restricted estimator, M_n^* . The choice of k_n in the proof is important. If $k_n = \lceil n^\delta \rceil$ is chosen, then we required $1/4 < \delta < 1/2$. Our restriction is simply monotonicity of M . It is interesting to note that for concavity or convexity restrictions, on d.f.'s, as in Kiefer and Wolfowitz (1976), or on cumulative hazard functions, as in Wang (1986), a similar result by linearization of the empirical estimates required a k_n of the order of $n^{1/3}$.

Using Theorem 4.1 we could use the same asymptotic inferences about M using M_n^* or \widehat{M}_n . In particular, Hall and Wellner (1979) have derived a conservative asymptotic confidence band for M . We state their result as a theorem below. Let B be a standard Brownian motion, let $\hat{\sigma}_n(0)$ denote the sample standard deviation of the entire sample, and for any $\beta \in (0, 1)$ let $a = a(\beta)$ be such that $P(\|B\|_0^1 \leq a) = \beta$. Let $d_n(\cdot) = \hat{\sigma}_n(0)/\sqrt{n} \widehat{S}_n(\cdot)$.

THEOREM 4.2 [Hall and Wellner (1979)]. *Under assumption (4.3),*

$$(4.11) \qquad \lim_{n \rightarrow \infty} P[|\widehat{M}_n(x) - M(x)| \leq a d_n(x) \forall x \geq 0] \geq \beta,$$

with equality for continuous s.f.'s.

TABLE 2
Approximate values of $P(a)$ for some values of a

a	2.807	2.241	1.960	1.534	1.149	0.871
$P(a)$.99	.95	.90	.75	.50	.25

Using Theorem 4.1, we could replace \widehat{M}_n by M_n^* in Theorem 4.2. The probability $P(a) \equiv P(\|B\|_0^1 \leq a)$ has an infinite series expansion in the standard normal c.d.f. [Billingsley (1968)]. Hall and Wellner (1979) show that for $a > 1.4$, the approximation $P(a) = 4\Phi(a) - 3$ gives a 3-place accuracy. They also provide a short table of values (Table 2).

5. Simulations. In this section we present some simulation results comparing the empirical estimator, \widehat{M} , with our restricted estimator, M^* , for the DMRL case. Lacking any convenient parametric family of distributions that are DMRL but not IFR, we have carried out the simulations for two IFR distributions,— the $U(0, 1)$ distribution and a *Weibull* distribution with s.f. $S(x) = \exp(-x^2)$. Note that these two distributions are strictly DMRL. We have also considered the *Exp*(1) distribution that is at the boundary of the DMRL and IMRL distributions. The results, based on 5,000 replications for a sample of size 30, are tabulated below, comparing the bias and the MSE of \widehat{M}_n and M_n^* at five quantiles. It may be noted that the $\text{Bias}^2(\widehat{M}_n)$ is negligible compared to $\text{MSE}(\widehat{M}_n)$ throughout.

It may be seen that M_n^* has a uniformly more negative bias than \widehat{M}_n , as is to be expected, but it has a uniformly smaller MSE in all cases, especially in the right tails. This appears to be true also in the exponential case even though the bias of M_n^* is very large there while that of \widehat{M}_n is rather small. A plausible explanation is that \widehat{M}_n has an asymptotic variance of $\sigma^2(x)/S(x)$ at x , which becomes large at the right tail, while M_n^* , although it has a large negative bias, avoids the contribution to the MSE when \widehat{M}_n has large positive values.

6. Nonparametric maximum likelihood estimation. Following Kiefer and Wolfowitz (1956), if a s.f. S belongs to a class \mathcal{S} , we define a sequence

TABLE 3
Comparison of \widehat{M}_n and M_n^* at quantile ξ_Q of $U(0, 1)$ distribution

Q	M(ξ_Q)	Bias(\widehat{M}_n)	Bias(M_n^*)	Mse(\widehat{M}_n)	Mse(\widehat{M}_n)/Mse(M_n^*)
.1	.45	-2.17E-04	-.00789	.00259	1.01318
.2	.40	5.17E-05	-.00794	.00228	1.00566
.5	.25	-.00044	-.00807	.00143	1.02207
.8	.10	-.00077	-.00841	.00068	1.04218
.9	.05	-.00266	-.00857	.00046	1.12036

TABLE 4

Comparison of $\widehat{M}_n M$ and M_n^* at quantile ξ_Q of the survival function $S(x) = \exp\{-x^2\}$

Q	M(ξ_Q)	Bias(\widehat{M}_n)	Bias(M_n^*)	Mse(\widehat{M}_n)	Mse(\widehat{M}_n)/Mse(M_n^*)
.1	.6363	-.00100	-.00821	.00670	1.01433
.2	.5584	-.00134	-.01253	.00649	1.02021
.5	.4237	-.00133	-.02441	.00744	1.09670
.8	.3226	.00033	-.05213	.01597	1.47369
.9	.2825	-.00956	-.08144	.02949	1.68432

$\widetilde{S}_n \in \mathcal{S}$ to be a NPMLE of S if, for every $U \in \mathcal{S}$,

$$(6.1) \quad \prod_{i=1}^n \frac{dP_n}{d(P + P_n)}(X_i) \geq \prod_{i=1}^n \frac{dP}{d(P + P_n)}(X_i),$$

where P and P_n are the probability measures corresponding to U and \widetilde{S}_n , respectively.

We first consider the DMRL case. Assume w.l.o.g. that $0 \equiv X_0 < X_1 < \dots < X_N = X_{N+1} = \dots = X_n$ are the ordered observations, where N is the number of distinct observations and $m_N \equiv n - N + 1$ is the number of repeated largest observations that may be more than one for the DMRL case. Since every DMRL s.f. is absolutely continuous with a possible jump at the endpoint of its support, it should be clear from (6.1) that P_n must be absolutely continuous w.r.t. the measure $(\lambda + \delta_0)$, where λ is the Lebesgue measure on the line and δ_0 is the point mass at X_N . The Radon-Nikodym derivative $dP/d(\lambda + \delta_0)$ is given by

$$g(x) = \begin{cases} f(x), & \text{if } x < X_N, \\ S(X_N^-), & \text{if } x = X_N, \\ 0, & \text{if } x > X_N. \end{cases}$$

Thus the NPMLE, if it exists, will be obtained by maximizing the likelihood function

$$L(S) = \left\{ \prod_{i=1}^{N-1} g(X_i) \right\} [g(X_N)]^{m_N},$$

TABLE 5

Comparison of \widehat{M}_n and M_n^* at quantile ξ_Q of the survival function $S(x) = e^{-x}$

Q	M(ξ_Q)	Bias(\widehat{M}_n)	Bias(M_n^*)	Mse(\widehat{M}_n)	Mse(\widehat{M}_n)/Mse(M_n^*)
.1	1	-.00237	-.04416	.03662	1.06165
.2	1	-.00306	-.06762	.04058	1.07985
.5	1	-.00337	-.13332	.06498	1.17958
.8	1	.00235	-.25789	.20653	1.64466
.9	1	-.03235	-.38283	.45915	1.91315

over all $S \in \mathcal{S}$. It seems reasonable to restrict our candidate s.f.'s to have the same mean, $M(0)$, for comparison purposes (Also, it can be shown that the NPMLE does not exist if the mean is allowed to vary.) Thus we will define \mathcal{S} to be the DMRL s.f.'s with mean $= \widetilde{M}_n(0)$.

Using the inversion formula (1.2), the density f is given by

$$(6.2) \quad f(x) = \frac{M(0)}{[M(x)]^2} [M'(x) + 1] \exp \left\{ - \int_0^x \frac{1}{\widehat{M}(u)} du \right\} = \frac{S(x)}{M(x)} [M'(x) + 1],$$

where $M'(x)$ exists. Note that $f(x) = 0$ if $M'(x) = -1$ and, if $M'(x) = 0$ in an interval, then $f(x) = S(x)/M(x)$ and $S(x)$ correspond to a shifted exponential distribution with mean $M(x)$ in that interval. Now there are some inherent ambiguities in defining a NPMLE when it has a density that is discontinuous at some order statistics. This is the case, for example, in the IFR case [Marshall and Proschan (1965)] where the NPMLE of the s.f. corresponds to a density that consists of segments of exponential densities that jump up at some order statistics, and it is customary to choose the right-hand limits of the density because they are larger. Since the one-sided derivatives of M exist everywhere, we adopt the convention of using the maximum of these when M' does not exist.

Suppose that the NPMLE, \widetilde{S}_n , exists with the MRL \widetilde{M}_n and density \widetilde{f}_n . The following proposition will show that \widetilde{M}_n must have the same shape as our estimator, M_n^* .

PROPOSITION 6.1. *Suppose that \widetilde{M}_n exists, and we are given only the values $\widetilde{M}_n(X_k)$ for $1 \leq k \leq N$. Then \widetilde{M}_n is completely determined on $[0, X_N]$.*

PROOF. Let $1 \leq k \leq N - 1$ be fixed. From the inversion formula (1.2) it can be seen that

$$(6.3) \quad \widetilde{S}_n(X_k) = \widetilde{S}_n(X_{k-1}) \frac{\widetilde{M}_n(X_{k-1})}{\widetilde{M}_n(X_k)} \exp \left\{ - \int_{X_{k-1}}^{X_k} \frac{1}{\widetilde{M}_n(u)} du \right\}$$

Suppose that \widetilde{M}_n has been determined on $[0, X_{k-1}]$. Then $\widetilde{S}_n(X_{k-1})$ is fixed from the inversion formula (1.2). If $\widetilde{M}_n(X_{k-1}) = \widetilde{M}_n(X_k)$, then $\widetilde{M}_n(u) = \widetilde{M}_n(X_{k-1})$ on $[X_{k-1}, X_k]$. If $\widetilde{M}_n(X_{k-1}) = a + b$ and $\widetilde{M}_n(X_k) = a$ for some $a, b > 0$, then, from (6.3), $\widetilde{S}_n(X_k)$ is maximized by choosing the largest possible \widetilde{M}_n on $[X_{k-1}, X_k]$ subject to the constraint that $\widetilde{M}_n(X_{k-1}) = a + b$, $\widetilde{M}_n(X_k) = a$, and \widetilde{M}_n is nonincreasing on $[X_{k-1}, X_k]$. From Figure 3 it is clear that the maximizing \widetilde{M}_n has the graph shown by the solid line, that is, $\widetilde{M}_n(u) = \widetilde{M}_n(X_{k-1})$ on $[X_{k-1}, X_k - b]$ and $\widetilde{M}_n(u) = \widetilde{M}_n(X_k) + X_k - u$ on $[X_k - b, X_k]$; the dotted line shows a competitor. The same argument applies to the interval $[X_{N-1}, X_N]$ using $\widetilde{S}_n(X_N)$ instead of $\widetilde{S}_n(X_{N-1})$ above. From (6.2) and (6.3), and using an induction argument on k , it is clear that, given the values of \widetilde{M}_n on

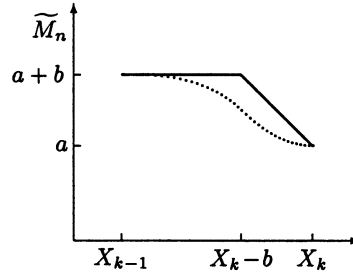


FIG. 3.

$\{X_k\}$ alone the likelihood is maximized by successively maximizing $\{\tilde{S}_n(X_k)\}$ as done above. \square

REMARK. The proof shows that $\tilde{M}_n(X_k^-) = 0$ or -1 , while $\tilde{M}_n(X_k^+) = 0$. Thus we must choose the right-hand derivative for $\tilde{f}_n(X_k)$.

We now consider the case of $n = 2$ with two distinct observations, $X_1 < X_2$. Here we have to maximize $f(X_1^+)S(X_2^-)$. Since the problem is scale invariant, we assume that the sample mean $\widehat{M}_n(0) = 1$ w.l.o.g. Then

$$X_2 = 2 - X_1, \widehat{M}_n(X_1^-) = 1 - X_1, \widehat{M}_n(X_1) = 2(1 - X_1), \text{ and } \widehat{M}_n(X_2) = 0.$$

Suppose that \tilde{M}_n exists. If $\tilde{M}_n(X_1^-) = t$, then, since $\tilde{M}_n(0) = 1$ and $\tilde{M}_n(X_2) = 0$, \tilde{M}_n is determined completely by Proposition 6.1. Note that $\widehat{M}_n(X_1^-) \leq t \leq \widehat{M}_n(X_1) \wedge 1$ by the restriction $\tilde{M}_n \geq -1$ and \tilde{M}_n is nonincreasing. For a $X_1 \leq t < 1$, corresponding to $X_1 > 1/2$, an explicit expression of \tilde{M}_n is given by [see Figure 4]

$$\tilde{M}_n(x) = \begin{cases} 1, & \text{if } 0 \leq x < X_1 + t - 1, \\ X_1 + t - x, & \text{if } X_1 + t - 1 \leq x < X_1, \\ t, & \text{if } X_1 \leq x < X_2 - t, \\ X_2 - x, & \text{if } X_2 - t \leq x < X_2, \\ 0, & \text{if } x \geq X_2. \end{cases}$$

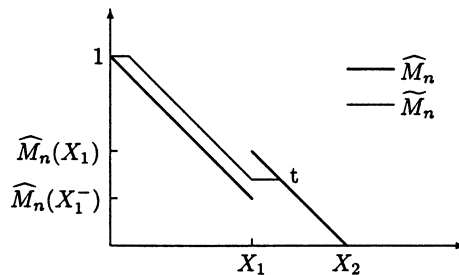


FIG. 4.

For $t = 1$, or $X_1 \leq 1/2$, $\widetilde{M}_n(x) = 1$ for $0 \leq x \leq X_2 - 1$ and $\widetilde{M}_n = \widehat{M}_n$ on $[X_2 - 1, X_2]$.

Now

$$\tilde{f}_n(X_1^+) = \frac{\tilde{S}_n(X_1)}{\widetilde{M}_n(X_1)} = \frac{\tilde{S}_n(X_1 + t - 1)}{t} = \frac{\exp\{-(X_1 + t - 1)\}}{t}$$

and

$$\begin{aligned} \tilde{S}_n(X_2^-) &= \tilde{S}_n(X_2 - t) = \tilde{S}_n(X_1 + t - 1) \exp\left\{\frac{-(X_2 - t) - X_1}{t}\right\} \\ &= \exp\left\{2 - X_1 - t - \frac{X_2 - X_1}{t}\right\}. \end{aligned}$$

The likelihood function then becomes a function of the single parameter t , and we have to maximize

$$\begin{aligned} L(t) &= \frac{1}{t} \exp\left\{3 - 2X_1 - 2t - \frac{X_2 - X_1}{t}\right\} \\ &= \frac{1}{t} \exp\left\{3 - 2X_1 - 2t - \frac{2(1 - X_1)}{t}\right\}, \end{aligned}$$

subject to $1 - X_1 \leq t \leq X_2 - X_1 = 2(1 - X_1)$, since $\widetilde{M}'_n \geq -1$. Setting $L'(t) = 0$ yields the equation

$$-\frac{1}{t^2} + \frac{1}{t} \left[-2 + \frac{2(1 - X_1)}{t^2}\right] = 0,$$

whose positive solution is given by

$$t_0 = \frac{-1 + \sqrt{1 + 16(1 - X_1)}}{4}.$$

Since $-1 + \sqrt{1 + 2a} < a$ and $-1 + \sqrt{1 + a^2} < a$ for all $a > 0$, we note that $t_0 < 2(1 - X_1) \wedge 1$ always. It can also be seen that

$$t_0 \geq 1 - X_1 \Leftrightarrow 1 - X_1 \leq 1/2 \Leftrightarrow 2(1 - X_1) = X_2 - X_1 \leq 1 = \widehat{M}_n(0) \Leftrightarrow X_1 \geq X_2/3.$$

For $X_1 \leq X_2/3$, $\widetilde{M}_n = M_n^*$. However, $\widetilde{M}_n \geq M_n^*$ in general, and it may be possible to improve on the negative bias of M_n^* using the NPMLE.

For $n \geq 3$, the NPMLE could be obtained by setting the values of $\widetilde{M}_n(X_i^-) = t_i$, $i = 1, 2, \dots, N - 1$, and then maximizing $L(t_1, \dots, t_{N-1})$ w.r.t. the $N - 1$ parameters subject to constraints that restrict each t_i to lie in a compact interval. It is clear that a solution exists, and it is also likely that it is asymptotically root- n equivalent to \widehat{M}_n in probability, the same as M_n^* . Unfortunately, even for the case $n = 3$, we get two coupled quadratic equations whose solution requires solving a quartic equation for which we have found no efficient solution.

Looking at the form of the density in (6.2), it can be seen that the NPMLE does not exist for the IMRL case since $\widetilde{M}'_n(X_i)$ could be made arbitrarily large over an arbitrarily small interval.

7. Concluding remarks. In this paper we have introduced some estimators for a MRL when it is restricted to be decreasing or increasing. They have been shown to be uniformly consistent on compact intervals bounded away from the endpoint of the support. Under the assumption of strict monotonicity, the estimators are asymptotically root- n equivalent in probability to the empirical estimator of Yang (1978) for which many asymptotic properties are known. Thus we get a monotone estimator that enjoys the same asymptotic properties as the unrestricted estimator. Simulations seem to indicate that, under the assumptions, the restricted estimator is uniformly superior to the empirical in terms of MSE, although it has a higher negative bias. This result is similar to those obtained for stochastic ordering [Rojo and Ma (1996)] and uniform stochastic ordering [Rojo and Samaniego (1993), Mukerjee (1996)] with similar projection type estimators. It has also been shown that the NPMLE does not exist for the IMRL case. It does exist for the DMRL case, and an explicit solution has been found for a sample of size 2. The solutions for larger sample sizes require solutions of coupled quadratic equations in many variables, and we have been unable to solve them effectively even for a sample of size 3. It does appear, however, that the NPMLE reduces some of the bias of our estimator. We conjecture that the NPMLE also is asymptotically root- n equivalent in probability to the empirical. It will be useful to verify this conjecture, and to come up with an effective computational scheme for evaluating the NPMLE.

One way to extend our results will be to consider the censored case. Yang (1977) has also considered an empirical estimate in this case using an estimator of the s.f. used by Aalen (1976) and Breslow and Crowley (1974) that has been shown to be asymptotically equivalent to the Kaplan-Meier (1958) estimator. She shows weak convergence of this estimator to a Gaussian process, but under the assumption that X is bounded. Our result for the uncensored case could be directly extended to the censored case, replacing the empirical s.f. by the Kaplan-Meier or the Aalen-Breslow-Crowley estimator. However, the assumption of boundedness on X appears to be unduly restrictive, and it will be worthwhile to remove this restriction.

We have not considered any testing problems in this paper. Hollander and Proschan (1975) have considered tests for exponentiality against NBUE, DMRL and IFR alternatives using integrals of empirical versions of functionals of the s.f.'s that vanish if and only if the distribution is exponential. Their test for the NBUE alternative turns out to be equivalent to a standard test for the IFR alternative even though we have $\text{IFR} \implies \text{DMRL} \implies \text{NBUE}$. It might be better to consider restricted estimators for each of the alternatives and base the test on the distribution of an exponential under the same order restriction. Work is presently under progress along these directions.

Acknowledgments. The authors are grateful to the referees for their comments which were very helpful.

REFERENCES

- ALY, E. A. A. (1990). Test for monotonicity properties of the mean residual life functions. *Scand. J. Statist.* **17** 189–200.
- AALEN, O. (1976). Nonparametric inference in connection with multiple decrement models. *Scand. J. Statist.* **3** 15–27.
- BILLINGSLEY, P. (1968). *Convergence of Probability Measures*. Wiley, New York.
- BJERKEDAL, T. (1960). Acquisition of resistance in guinea pigs infected with different doses of virulent tubercle bacilli. *Amer. J. Hygiene* **72** 130–48.
- BRESLOW, N. and CROWLEY, J. (1974). A large sample study of the life table and product limit estimates under random censorship. *Ann. Statist.* **2** 437–453.
- CSÖRGŐ, M. and ZITIKIS, R. (1996). Mean residual life processes. *Ann. Statist.* **24** 1717–1739.
- GERTSBAKH, I. B. and KORDONSKIY, K. B. (1969). *Models of Failure*. Springer, New York.
- GUESS, F., HOLLANDER, M. and PROSCHAN, F. (1986). Testing exponentiality versus a trend change in mean residual life. *Ann. Statist.* **14** 1388–1398.
- GUESS, F. and PROSCHAN, F. (1988). Mean residual life: theory and applications. In *Handbook of Statistics* (P. R. Krishnaiah and C. R. Rao, eds.) **7** 215–224. North-Holland, New York.
- HALL, W. J. and WELLNER, J. A. (1979). Estimation of mean residual life. Unpublished manuscript.
- HAWKINS, D. L., KOCHAR, S. and LOADER, C. (1992). Testing exponentiality against IDMRL distributions with unknown change point. *Ann. Statist.* **20** 280–290.
- HOLLANDER, M. and PROSCHAN, F. (1975). Testing for mean residual life. *Biometrika* **62** 585–593.
- KAPLAN, E. L. and MEIER, P. (1958). Nonparametric estimation from incomplete observations. *J. Amer. Statist. Assoc.* **53** 457–481.
- KIEFER, J. and WOLFOWITZ, J. (1956). Consistency of the maximum likelihood estimator in the presence of infinitely many incidental parameters. *Ann. Math. Statist.* **27** 887–906.
- KIEFER, J. and WOLFOWITZ, J. (1976). Asymptotically minimax estimation of concave and convex distribution functions. *Z. Wahrsch. Verw. Gebiete* **34** 73–85.
- LIM, J.-H. and PARK, D. H. (1998). A family of tests for trend change in mean residual life. *Comm. Statist. Theory Methods* **27** 1163–1179.
- MARSHALL, A. W. and PROSCHAN, F. (1965). Maximum likelihood estimation for distributions with monotone failure rate. *Ann. Math. Statist.* **36** 69–77.
- MUKERJEE, H. (1996). Estimation of survival functions under uniform stochastic ordering. *J. Amer. Statist. Assoc.* **91** 1684–1689.
- ROJO, J. (1995). On the weak convergence of certain estimators of stochastically ordered survival functions. *J. Nonparamet. Statist.* **4** 349–363.
- ROJO, J. and MA, Z. (1996). On the estimation of stochastically ordered survival functions. *J. Statist. Comput. Simulation* **55** 1–21.
- 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.
- WANG, J.-L. (1986). Asymptotically minimax estimators for distributions with increasing failure rate. *Ann. Statist.* **14** 1113–1131.
- YANG, G. (1977). Life expectancy under random censorship. *Stochastic Process. Appl.* **6** 33–39.
- YANG, G. L. (1978) Estimation of a biometric function. *Ann. Statist.* **6** 112–116.

DEPARTMENT OF STATISTICS
INDIAN STATISTICAL INSTITUTE
CALCUTTA
INDIA
EMAIL: kochar@isid.isid.ac.in

DEPARTMENT OF STATISTICS
WICHITA STATE UNIVERSITY
WICHITA, KANSAS
E-MAIL: mukerjee@twsuvm.uc.twsu.edu

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