MEAN RESIDUAL LIFE PROCESSES

By Miklós Csörgő¹ and Ričardas Zitikis²

Carleton University

Yang and Hall and Wellner initiated investigations of the asymptotic uniform behaviour of mean residual life (MRL) processes. They obtained results holding true over fixed and expanding compact subintervals of $[0,\infty)$.

In this exposition we study MRL processes over the whole positive half-line $[0,\infty)$. We describe classes of weight functions which enable us to establish the (a) strong uniform-over- $[0,\infty)$ consistency and (b) weak uniform-over- $[0,\infty)$ approximation of MRL processes. We give examples which show the necessity of employing weight functions in order to have (a) and (b), and prove the optimality of the weight function classes which we make use of. Extending our results concerning (b), we discuss constructions of asymptotic confidence bands for unknown MRL functions. The width of the obtained confidence bands is regulated by weight functions depending on the available information on the underlying distribution function.

1. Introduction and main results. The mean residual life (MRL for short) function or remaining life expectancy function at age x is defined to be the expected remaining life given survival to age x. It is a concept of obvious interest and, indeed, one of the most important notions in actuarial, reliability and survivorship studies. For details on the variety of situations where estimating MRL is of importance, we may, for example, refer to the proceedings edited by Proschan and Serfling (1974), Miller, Efron, Brown and Moses (MEBM) (1980), Crowley and Johnson (1982) and Krishnaiah and Rao (1988).

Let X be a nonnegative random variable defined on a (fixed) probability space $\{\Omega, \mathscr{A}, \mathbf{P}\}$. Its distribution function F (usually called life distribution function) is assumed to be continuous, and $\mathbf{E}X < \infty$. As just mentioned above, the MRL function M_F at (age) $x \geq 0$ is defined by

$$M_F(x) := \mathbf{E}(X - x \mid X > x), \qquad x \ge 0.$$

Naturally, if F were known and completely specified, then M_F would be a computable function. This, of course, is rarely the case. A natural way, therefore, to get to know the MRL function M_F is via constructing an empirical estimate for it and, based on the empirical estimate and whatever

Received June 1994; revised May 1995.

¹Research supported by an NSERC Canada grant at Carleton University, Ottawa.

²Work done while the author was a Canada International Fellow at Carleton University, Ottawa, on leave from the Institute of Mathematics and Informatics, Vilnius.

AMS 1991 subject classifications. Primary 62N05, 62G15; secondary 62F17, 62E20.

Key words and phrases. Mean residual life, life expectancy, strong consistency, weak approximations, confidence bands, weighted empirical processes, Brownian motion, Brownian bridge.

regularity conditions one may be willing to assume for F, like the existence of moments and/or regular tail variation, for example, one would then wish to construct asymptotic confidence bands for the unknown M_F , using all the data that may be available. In this regard, we succeed in estimating the MRL function over the whole positive half-line. This really is also the main statistical message and implication of our paper. In this section we present and discuss our results along these lines and in a historical context as well. The proofs are given in Section 2.

Let X_1, \ldots, X_n be independent copies of the random variable X; the corresponding right-continuously defined empirical distribution function is denoted by F_n . If X^* is a random variable having the distribution function F_n when X_1, \ldots, X_n are fixed, then M_n is defined by the formula

$$M_n(x) := \mathbf{E}^*(X^* - x \mid X^* > x), \qquad x \ge 0$$

where \mathbf{E}^* denotes the conditional expectation when X_1,\ldots,X_n are fixed. In several papers, attempts to prove uniform-over- $[0,\infty)$

(a) strong consistency

and

(b) weak approximation

for the MRL process $M_n - M_F$ were made. In particular, under somewhat stronger than necessary conditions, Yang (1978) proved (a) and (b) for the process $M_n - M_F$ uniformly over the interval $[0, x_0]$ for any fixed $x_0 < x_F$, where x_F is the, possibly infinite, end of the support of F, defined by

$$x_F := \inf\{x : F(x) = 1\}.$$

Later on, Hall and Wellner (1979) proved the just mentioned Yang (1978) results under optimal moment conditions. Also, Hall and Wellner (1979) showed that (a) and (b) hold uniformly over the interval $[0, x_n]$, where the nonnegative numbers x_n possibly depending on F, are strictly less than x_F and approach x_F as $n \to \infty$. For a discussion and new derivation of these and some other closely related results on the MRL process we refer to Hall and Wellner (1981), Burke, S. Csörgő and Horváth (BCsH) (1981, 1988), M. Csörgő, S. Csörgő and Horváth (CsCsH) (1986), Shorack and Wellner (1986) and Bickel, Klaassen, Ritov and Wellner (BKRW) (1993).

In this work we show that, in general, it is impossible to have (a) and (b) uniformly over $[0,\infty)$ without employing weight functions. In particular, we refer to (1.7) and Corollary 1.4, as well as to Corollary 1.8 and Proposition 1.1 below, concerning the exponential distribution. This, in turn, suggests the following two natural approaches to the problem of studying the MRL process on $[0,\infty)$:

(A) To describe sequences $\{x_n, n \in \mathbb{N}\}$ of nonnegative numbers, or random variables $x_n, x_n \leq x_F$, such that (a) and (b) would hold true uniformly over the intervals $[0, x_n]$.

(B) To describe classes of weight functions q which would allow (a) and (b) to hold uniformly over $[0,\infty)$ when the MRL process M_n-M_F is weighted by q.

We have found the second approach (B) more natural for immediate attention. Indeed, it turns out that in some cases (like the exponential one) we have to push the MRL process M_n-M_F down [in other words, to adjust it by a function q such that $q(F(x)) \to 0$ as $x \to x_F$] in order to have (a) and (b) valid, while in some other cases [like, for example, the (0,1)-uniform one] we have to adjust M_n-M_F by a function q such that $q(F(x)) \to \infty$ as $x \to x_F$, in order to catch the true asymptotic behaviour of the process M_n-M_F . Another compelling reason for choosing the second approach (B) is discussed in more detail below when constructing confidence bands for the (unknown) MRL function M_F . We shall see (cf. Theorem 1.3 and the related discussions around there) that, using weight functions q, we are able to "regulate" the width of confidence bands according to the nature of information available to us on the distribution function F.

Thus, our goal here is to establish (B). However, we recognize of course that a resolution of (A), just like that of (B), is also of interest for the sake of the manifold applications of the MRL process. Hall and Wellner (1979) made important contributions in the first direction (A). This also poses the intriguing question of unifying the two approaches (A) and (B) via constructing a weighted approximation theory so that it would, for example, contain the results of this paper as well as those of Hall and Wellner (1979). Unfortunately, posing the question this way also poses some mathematical difficulties that we have not yet succeeded in overcoming.

We assume that the weight functions $q: [0,1] \to [0,\infty]$ we deal with are measurable and, for every $\delta > 0$,

$$(1.1) \qquad \sup\{q(t): t \in [0, 1 - \delta]\} < \infty.$$

We note in passing that assumption (1.1) can be relaxed via making some integrability assumptions on q over the intervals $[0, 1 - \delta]$, $\delta > 0$. However, at this stage, we do not see any immediate statistical or mathematical value in such generalizations.

We now state our main result on strong consistency. Denote

$$\Delta_n(q,F) := \sup_{x \in [0,\infty)} q(F(x)) |M_n(x) - M_F(x)|.$$

We assume that the function

$$(1.2) t \mapsto \frac{q(t)}{1-t} \log_2 \frac{1}{1-t}$$

is nondecreasing in a neighborhood of 1. We use the notation $\log_2 x := \log(e \vee \log(e \vee x))$ throughout the paper.

Theorem 1.1. If (1.2) obtains, $\mathbf{E}X < \infty$ and

$$(1.3) \quad \Delta(q,F) \coloneqq \int_0^{x_F} \left\{ \frac{1}{1 - F(x)} \int_x^{x_F} q(F(y)) \log_2 \frac{1}{1 - F(y)} \, dy \right\} dF(x) < \infty,$$

$$then \ \Delta_n(q,F) \to_{a.s.} \ 0 \ as \ n \to \infty.$$

We do not know whether the factor $\log_2\{1/(1-F(x))\}$ in $\Delta(q,F)$ could be omitted and the strong consistency result of Theorem 1.1 retained. However, by slightly modifying the proof of Theorem 1.1, it is easy to show that if the function $t\mapsto q(t)/(1-t)$ is nondecreasing in a neighborhood of 1 and the condition

(1.3')
$$\int_0^{x_F} \left\langle \frac{1}{1 - F(x)} \int_x^{x_F} q(F(y)) \, dy \right\rangle dF(x) < \infty$$

[that is, condition (1.3) without the factor $\log_2 \cdots$] is satisfied, then $\Delta_n(q,F) \to_P 0$ as $n \to \infty$. We omit further comments on this result, since they are quite in parallel with those given below in the case of the strong consistency of Theorem 1.1. It will suffice to note only that, due to the above remark, if we omit the factor $\log_2\{1/(1-t)\}$ in (1.5), (1.6) and (1.8) below, then the statements of Corollaries 1.2–1.4 hold true "in probability" instead of "almost surely" as stated now.

Concerning the monotonicity assumption on q given by (1.2), we note that, in order to have Theorem 1.1 hold true, it suffices to have a function h which majorizes q and satisfies conditions (1.2) and (1.3). We add also that such, seemingly trivial, generalization of Theorem 1.1 has proved to be useful on occasions (cf., e.g., the proofs of Corollaries 1.1 and 1.5).

COROLLARY 1.1 [Yang (1978); Hall and Wellner (1979)]. For every fixed $x_0 < x_F$, as $n \to \infty$,

(1.4)
$$\sup_{x \in [0, x_0]} |M_n(x) - M_F(x)| \to_{a.s.} 0.$$

The next corollary to Theorem 1.1 shows what kind of weight functions should be used to adjust the MRL process in order to assure strong consistency uniformly over the whole positive half-line $[0,\infty)$ when no information, other than $\mathbf{E}X<\infty$, on F is available.

COROLLARY 1.2. If q is the function

(1.5)
$$t \mapsto (1-t)/\log_2\{1/(1-t)\},$$

then $\Delta_n(q, F) \to_{a.s.} 0$ as $n \to \infty$.

For further elucidation on Theorem 1.1, we formulate the following two corollaries concerning the (0,1)-uniform and exponential distributions, respectively.

COROLLARY 1.3. If F is the (0,1)-uniform distribution function and q is such that

(1.6)
$$\int_0^1 q(t) \log_2\{1/(1-t)\} dt < \infty,$$

then $\Delta_n(q, F) \to_{a.s.} 0$ as $n \to \infty$.

Corollary 1.3 allows us the use of weight functions $t\mapsto (1-t)^{-1+\alpha}$, $\alpha>0$, in the (0,1)-uniform case when proving strong consistency results for the MRL processes M_n-M_F . We note that when $\alpha<1$, then this weight function goes to $+\infty$ as $t\to 1$. In comparison with corresponding results by Yang (1978) and Hall and Wellner (1979) (cf. Corollary 1.1 above), in this situation we would have to require q(t)=0 in a neighborhood of 1.

On the other hand, in the exponential case we must require

$$(1.7) q(t) \to 0, \text{ as } t \to 1,$$

as shown by the following argument: Let F be the $\text{Exp}(\theta)$ -distribution function $\theta \in (0, \infty)$, that is, $F(x) = 1 - e^{-x/\theta}$, $x \ge 0$. Then $M_F(x) = \theta$ and, therefore,

$$\Delta_{n}(q,F) \geq \sup_{x \in [X_{n:n}, x_{F})} q(F(x)) M_{F}(x) = \sup_{x \in [X_{n:n}, x_{F})} q(F(x)) \theta.$$

Consequently, if $\Delta_n(q,F) \to_{\text{a.s.}} 0$ as $n \to \infty$, then we must have (1.7). Here, and throughout, $X_{1:n} \le \cdots \le X_{n:n}$ denote the order statistics of X_1, \ldots, X_n . In general, we have the following result for exponential distributions.

COROLLARY 1.4. If F is the $\text{Exp}(\theta)$ -distribution function $\theta \in (0, \infty)$ and q is such that

(1.8)
$$\int_0^1 \frac{q(t)}{1-t} \log_2 \frac{1}{1-t} dt < \infty,$$

then $\Delta_n(q, F) \rightarrow_{a.s.} 0$ as $n \rightarrow \infty$.

The second goal of this exposition is to prove weak approximation results for the MRL process M_n-M_F . Namely, we shall describe a class of weight functions q, such that on a (possibly richer) probability space $\{\tilde{\Omega}, \tilde{\mathscr{A}}, \tilde{\mathbf{P}}\}$ there is a sequence $\{G_n, n \in \mathbf{N}\}$ of Gaussian random processes such that the quantity

$$\nabla_{n}(q,F) := \sup_{x \in [0,\infty)} q(F(x)) \Big| \sqrt{n} \left\{ M_{n}(x) - M_{F}(x) \right\} - G_{n}(x) \Big|$$

converges in probability to 0 as $n \to \infty$. For notational simplicity, we will continue using the notation $\{\Omega, \mathscr{A}, \mathbf{P}\}$ instead of $\{\tilde{\Omega}, \tilde{\mathscr{A}}, \tilde{\mathbf{P}}\}$. Furthermore, the probability space $\{\Omega, \mathscr{A}, \mathbf{P}\}$ which the two sequences of random processes $\sqrt{n}\{M_n - M_F\}$ and G_n live on is that of Theorem 1.1 of M. Csörgő, S. Csörgő,

Horváth and Mason (CsCsHM) (1986). The approximating Gaussian processes G_n are defined as follows:

$$G_n(x) := \frac{1}{1 - F(x)} B_n(F(x)) M_F(x) - \frac{1}{1 - F(x)} \int_x^\infty B_n(F(y)) dy$$

for all $x \in [0, x_F)$, and $G_n(x) := 0$ otherwise, where $\{B_n, n \in \mathbb{N}\}$ denotes the sequence of Brownian bridges that are constructed in CsCsHM [(1986), Theorem 1.1, page 32].

We assume that the function

$$(1.9) t \mapsto q(t)/(1-t)$$

is nondecreasing in a neighborhood of 1 and that $\mathbf{E}X^2 < \infty$.

Theorem 1.2. If (1.9) holds, $\mathbf{E}X^2 < \infty$ and

$$(1.10) \qquad \nabla(q,F) := \int_0^{x_F} \left\{ \frac{1}{1 - F(x)} \int_x^{x_F} q(F(y)) \, dy \right\}^2 dF(x) < \infty,$$

then $\nabla_n(q,F) \to_P 0$ as $n \to \infty$.

Just as in the case of Theorem 1.1, assumption (1.10) in Theorem 1.2 can be replaced by $\nabla(h, F) < \infty$, where h is a function such that $q \le h$, and (1.9) holds with the h instead of q.

Weak approximations for the (normalized) MRL processes $\sqrt{n} (M_n - M_F)$ by G_n were studied by Yang (1978), Hall and Wellner (1979), BCsH (1981), CsCsH [(1986), Chapter 4] and Shorack and Wellner [(1986), Section 5 in Chapter 23]. For some related discussions of MRL processes we refer to BKRW (1993). Corollaries 1.5, 1.6 and the statements (1.13), (1.14) below relate our Theorem 1.2 to these works.

COROLLARY 1.5 [Yang (1978); Hall and Wellner (1979)]. For every fixed $x_0 < x_F$, as $n \to \infty$,

(1.11)
$$\sup_{x \in [0, x_0]} \left| \sqrt{n} \left\{ M_n(x) - M_F(x) \right\} - G_n(x) \right| \to_P 0.$$

The next corollary to Theorem 1.2 answers the question of what kind of weight functions should be used in the weak approximation of MRL when the only available information on F is $\mathbf{E}X^2 < \infty$.

COROLLARY 1.6. If q is the function

$$(1.12) t \mapsto 1 - t,$$

then $\nabla_n(q,F) \to_P 0$ as $n \to \infty$.

This corollary implies the first part of Theorem 4 of Hall and Wellner [(1979), page 9] which, under somewhat stronger assumptions than $\mathbf{E}X^2 < \infty$, states that

states that
$$(1.13) \quad \sup_{x \in [0, x_n]} (1 - F(x)) | \sqrt{n} \{ M_n(x) - M_F(x) \} - G_n(x) | \to_P 0$$

as $n \to \infty$, where $\{x_n \colon n \in \mathbf{N}\}$ is a sequence of numbers such that $1 - F(x_n) \to 0$ and $n\{1 - F(x_n)\} \to \infty$, $n \to \infty$. Furthermore, Hall and Wellner [(1979), Theorem 3 on page 9, under somewhat stronger assumptions than $\mathbf{E}X^2 < \infty$] and CsCsH [(1986), Theorem 4.1(iv) on page 39, under only the assumption $\mathbf{E}X^2 < \infty$] proved the result

EX
$$<\infty$$
 proved the result
$$\sup_{x\in[0,\infty)} \left|\{1-F_n(x)\}\sqrt{n}\left\{M_n(x)-M_F(x)\right\}\right| \\ -\{1-F(x)\}G_n(x)\big|\to_P 0, \qquad n\to\infty,$$

which is slightly different from those given in Corollary 1.6 and (1.13). Clearly, (1.14) is implied by Corollary 1.6. We shall see later on that results like (1.14) are very useful when investigating the problem of constructing asymptotic confidence bands for the (unknown) MRL function M_F .

We now state some further corollaries to Theorem 1.2 that are of interest.

COROLLARY 1.7. If F is the (0,1)-uniform distribution function and

$$\int_0^1 q^2(t) dt < \infty,$$

then $\nabla_n(q,F) \to_P 0$ as $n \to \infty$.

COROLLARY 1.8. If F is the $\text{Exp}(\theta)$ -distribution function $\theta \in (0, \infty)$ and

$$\int_0^1 \left\langle \frac{q(t)}{1-t} \right\rangle^2 dt < \infty,$$

then $\nabla_n(q,F) \to_P 0$ as $n \to \infty$.

In particular, Corollary 1.8 says that $\nabla_n(q,F) \to_P 0$, $n \to \infty$, for any weight function $q(t) = (1-t)^{1/2+\alpha}$, $\alpha > 0$. On the other hand, the next proposition states that $\nabla_n(q,F) \to_P 0$, $n \to \infty$, fails to hold in case $\alpha = 0$.

PROPOSITION 1.1. If F is the Exp(1)-distribution function, and $q(t) = \sqrt{1-t}$, then there exists a constant c > 0 such that

(1.17)
$$\Lambda := \liminf_{n \to \infty} \mathbf{P} \{ \nabla_n(q, F) \ge \frac{1}{2} \} \ge c > 0.$$

Based on Theorem 1.2, we now discuss several constructions of asymptotic confidence bands for the MRL function M_F . Fix any confidence level $\alpha \in (0,1)$ and let z_{α} be the smallest number z such that

(1.18)
$$P(z) := \mathbf{P} \Big\{ \sup_{x \in [0, x_F)} q(F(x)) |G(x)| \le z \Big\} = 1 - \alpha,$$

where $G(\cdot)$ is a Gaussian process such that $\{G(x), 0 \le x < \infty\} =_{\mathscr{D}} \{G_n(x), 0 \le x < \infty\}$ for each $n \ge 1$.

We note in passing that the function P is continuous provided that q is not identically equal to 0 on $[0,x_F)$ [cf. Tsirel'son (1975)]. Some further details and remarks on this and other related topics can be found in Bickel and Freedman (1981), M. Csörgő, S. Csörgő and Mason (CsCsM) (1984) and S. Csörgő and Mason (1989).

In our next theorem (cf. Theorem 1.3) we will also assume that

$$(1.19) q is continuous on $[0,1),$$$

and that for every $c \in (0, 1)$,

(1.20)
$$\sup_{t \in (0,1)} h(ct)/h(t) < \infty \quad \text{where } h(t) \coloneqq q(1-t).$$

We note that assumption (1.20) is a technical one needed in the proof of Theorem 1.3 and, roughly speaking, it says that if the function q diverges to $+\infty$ as $t \to 1$, then it should diverge with a polynomial rate.

THEOREM 1.3. Under the assumptions of Theorem 1.2, and those of (1.19) and (1.20) we have

(1.21)
$$\mathbf{P}\left\{M_{F}(x) \in \left[M_{n}(x) - \frac{z_{\alpha}}{\sqrt{n}\,q(F_{n}(x))}, \right.\right.$$

$$\left.M_{n}(x) + \frac{z_{\alpha}}{\sqrt{n}\,q(F_{n}(x))}\right], x \in [0, X_{n:n})\right\}$$

$$= 1 - \alpha + o(1), \qquad n \to \infty.$$

In order to use Theorem 1.3 in practical situations, we have to discuss two problems. Namely, in applications we need to choose q and to estimate z_{α} . As to the weight function q, it has to satisfy the conditions $\nabla(q,F)<\infty$, and thus its choice depends on the unknown F. If $x_p<\infty$, then, roughly speaking, one can choose q as if the distribution function F were (0,1)-uniform (cf. Corollary 1.7). In the case of $x_p=\infty$, the problem is more complicated. For example, if F is the $\operatorname{Exp}(\theta)$ -distribution, then any q satisfying (1.16) can be used in Theorem 1.3. Other parametric families [cf., e.g., Miller (1981) for an account of them] could be investigated in a similar way. When we do not know the parametric family from which F comes, but we know, or can guess, the number of moments F has, then one may choose q according to the following proposition.

Proposition 1.2. Let $\nu > 1/2$ be any number and let

(1.22)
$$\mathbf{E} X^{2/(2\nu-1)} < \infty.$$

Then for any weight function q such that

(1.23)
$$q(t) \le c(1-t)^{\nu} \text{ for all } t \in [0,1],$$

the condition $\nabla(q, F) < \infty$ is satisfied.

In the case of $\nu=1$, Proposition 1.2 results in the same conclusion as that of Corollary 1.6 (cf. also the remark right after Theorem 1.2 above). The case $\nu>1$ is of no interest, since $\mathbf{E}X^2<\infty$ is assumed for the validity of Theorem 1.2. Thus, without loss of generality, the value of $\nu>1/2$ in Proposition 1.2 can be restricted to $\nu\in(1/2,1]$. We recall also that, due to Proposition 1.1 and Theorem 1.2, the value $\nu=1/2$ cannot be accommodated by Proposition 1.2. Naturally, if $\mathbf{E}X^r<\infty$, $r\geq 2$, then with $q(t)=(1-t)^\nu$ and $\nu\in(1/2,1]$ such that $r=2/(2\nu-1)$, we have Theorem 1.3.

Concerning the problem of z_{α} in (1.21), we now discuss some methods of construction for its estimation. The first approach is taken from Hall and Wellner [(1979), Section 7] [cf. also Shorack and Wellner (1986), page 779]. Their idea is based on the fact that the distribution of the process $x \mapsto G(x)$ is equal to that of the process

$$x\mapsto \frac{\sigma(x)}{1-F(x)}W\bigg(\frac{\{1-F(x)\}\sigma^2(x)}{\sigma^2(0)}\bigg),$$

where $\sigma^2(x) \coloneqq \operatorname{Var}(X-x \mid X>x)$ and W is a standard Wiener process on [0,1] [cf. Exercise 1 on pages 778–779 of Shorack and Wellner (1986) for further details on this topic]. Now, using empirical estimates for F and σ , and the fact that the distribution of $\sup\{|W(t)|: t \in (0,1)\}$ is known, we arrive at an (empirical) estimator of z_{α} , denoted by $z_{\alpha,n}$. For more details and properties of the estimator $z_{\alpha,n}$ we refer to Hall and Wellner (1979) and Shorack and Wellner (1986).

An alternative approach to constructing an estimator for z_{α} is provided by bootstrapping the MRL process M_n-M_F as follows. Let X_1^*,\ldots,X_n^* be independent and identically distributed random variables with the distribution function F_n , where X_1,\ldots,X_n are fixed. Furthermore, let F_n^* be the empirical distribution function corresponding to the random sample X_1^*,\ldots,X_n^* . Then we define

$$M_n^*(x) := \frac{1}{1 - F_n^*(x)} \int_x^{\infty} \{1 - F_n^*(y)\} dy$$

for all $x \in [0, X_{n : n}^*)$, and $M_n^*(x) := 0$ otherwise. In other words, we have

$$M_n^*(x) = \mathbf{E}^{**}(X^{**} - x \mid X^{**} > x), \qquad x \ge 0,$$

where the notation corresponds to that used in the definition of $M_n(x)$. If we now denote

$$P_n^*(x) \coloneqq \mathbf{P}^* \bigg\{ \sup_{x \in [0, X_{n \cdot n}^*)} q(F_n(x)) \big| M_n^*(x) - M_n(x) \big| \le z \bigg\},$$

where \mathbf{P}^* stands for the conditional probability when X_1,\ldots,X_n are fixed, then the bootstrap based estimator $z_{\alpha,\,n}^*$ for z_{α} is defined as the smallest number z such that

$$P_n^*(z) \geq 1 - \alpha$$
.

There has been a lot of material published on various aspects of the bootstrap since Efron (1979). Investigations of bootstrapping MRL and other related empirical processes began in CsCsM (1984) and continued with further developments in CsCsH (1986) and S. Csörgő and Mason (1989) [cf. also Chung (1989) and Yu (1993)]. For example, using ideas and techniques developed in CsCsH (1986) and S. Csörgő and Mason (1989), one can prove (in the spirit of Theorem 1.3) that

(1.24)
$$\mathbf{P}\left\{M_{F}(x) \in \left[M_{n}(x) - \frac{z_{\alpha, n}^{*}}{\sqrt{n} q(F_{n}(x))}, \right.\right.$$

$$\left.M_{n}(x) + \frac{z_{\alpha, n}^{*}}{\sqrt{n} q(F_{n}(x))}\right], x \in [0, X_{n:n}^{*})\right\}$$

$$\geq 1 - \alpha + o(1), \qquad n \to \infty,$$

under the conditions of Theorem 1.3.

We conclude this section by calling attention to the interesting and mathematically challenging problem of investigating mean residual life under random censorship (from the right). Concerning this subject, we refer to the recent research monographs Andersen, Borgan, Gill and Keiding (ABGK) (1993), BKRW (1993) and the references in these works. The very first steps in constructing confidence bands for the MRL function under random censorship were made by Yang (1977), Hall and Wellner (1980) and BCsH (1981). In these and subsequent papers, results are proved over (fixed or expanding) compact intervals. Thus, in the spirit of the present paper, it would be of interest to obtain "a.s." and "in P" results for MRL under random censorship which would hold over the whole half-line $[0,\infty)$. This, in turn, would first require carrying out another challenging research program on studying weighted "a.s." and "in P" approximations of empirical processes under random censorship, using various empirical distribution estimators of F, like for example those of Berman (1963) and Kaplan and Meier (1958). That is to say, one would need results along the lines of those that are already available in the case of uncensored empirical processes.

2. Proofs. We start with some general remarks. First, without loss of generality, we may and do assume throughout that the weight function q is such that for every $\delta > 0$ there is a constant $c = c(\delta) \in (0,1)$ such that

(2.1)
$$c < q(t) < 1/c \text{ for all } t \in [0, 1 - \delta].$$

Indeed, $\Delta(q,F) \leq \Delta(\bar{q},F)$, where $\bar{q}(t) \coloneqq q(t)$ for every $t>1-\delta$ and $\bar{q}(t) \coloneqq \lambda$ for every $t \in [0,1-\delta]$, and where λ denotes $\sup\{q(t)+1\colon t \in [0,1-\delta]\}$, which is finite on account of (1.1). The same arguments are applicable for studying the quantity $\nabla_n(q,F)$ as well. The second remark is the simple observation that

(2.2)
$$M_F(x) = \frac{1}{1 - F(x)} \int_x^{\infty} \{1 - F(y)\} dy$$

for all $x \in [0, x_F)$ and $M_F(x) = 0$ for all $x \ge x_F$. Also,

(2.3)
$$M_n(x) = \frac{1}{1 - F_n(x)} \int_x^{\infty} \{1 - F_n(y)\} dy$$

for all $x \in [0, X_{n:n})$ and $M_n(x) = 0$ for all $x \ge X_{n:n}$.

PROOF OF THEOREM 1.1. We have

$$\begin{split} \Delta_{n}(q,F) & \leq \sup_{x \in [0,X_{n:n})} q(F(x)) |M_{n}(x) - M_{F}(x)| \\ & + \sup_{x \in [X_{n:n},x_{F})} q(F(x)) |M_{n}(x) - M_{F}(x)| \\ & =: \Lambda_{2} + \Lambda_{1}. \end{split}$$

For proving the theorem, we show that $\Lambda_1, \Lambda_2 \to_{\text{a.s.}} 0$, $n \to \infty$. The fact $\Lambda_1 \to_{\text{a.s.}} 0$, $n \to \infty$, is immediate. Indeed, since $M_n(x) = 0$ for all $x \ge X_{n : n}$,

$$\Lambda_1 = \sup_{x \in [X_{n:n}, x_F)} q(F(x)) M_F(x).$$

Also, we have that $X_{n:n} \to_{\text{a.s.}} x_F$, $n \to \infty$ and that the function in (1.2) is nondecreasing in a neighborhood of 1, say on $[1-\delta,1)$ for some $\delta > 0$. Given such a $\delta > 0$ and n large enough, we get (using the above remarks) the following bound:

$$\Lambda_1 \le \int_{X_{p+p}}^{x_F} q(F(y)) \log_2 \frac{1}{1 - F(y)} \, dy / \log_2 \frac{1}{1 - F(X_{n+p})}.$$

This bound, along with the assumption (1.3) and the fact that $X_{n:n} \to_{\text{a.s.}} x_F$, $n \to \infty$, completes the proof of $\Lambda_1 \to_{\text{a.s.}} 0$, $n \to \infty$.

The proof of

$$(2.4) \Lambda_2 \to_{a.s.} 0, n \to \infty,$$

is a bit more complicated. We start with Corollary 7(i) on page 81 of Wellner (1978), which states that, given $\tau > 1$, for almost all elementary events ω one can find a number $N := N(\omega, \tau)$ such that

(2.5)
$$1 - F_n(x) \ge \left\{1 - F(x)\right\} / \left\{\tau \log_2 \frac{1}{1 - F(x)}\right\}$$

for all $x \in [0, X_{n : n})$ and all $n \ge N$. Now, using (2.2), (2.3) and (2.5), we easily arrive at the bound

$$(2.6) |M_n(x) - M_F(x)| \le \tau \psi_1(x) + \tau \psi_2(x), x \in [0, X_{n:n}),$$

where

$$\begin{split} \psi_1(x) &\coloneqq \frac{M_F(x)}{1 - F(x)} \log_2 \frac{1}{1 - F(x)} \big| F_n(x) - F(x) \big|, \\ \psi_2(x) &\coloneqq \frac{1}{1 - F(x)} \log_2 \frac{1}{1 - F(x)} \bigg| \int_x^{\infty} \{ F_n(y) - F(y) \} \, dy \bigg|. \end{split}$$

We show that

$$(2.7) \qquad \Lambda_3 \coloneqq \sup_{x \in [0, X_{n:n})} q(F(x)) \psi_1(x) \to_{\text{a.s.}} 0, \qquad n \to \infty,$$

(2.8)
$$\Lambda_4 := \sup_{x \in [0, X_{n:n})} q(F(x)) \psi_2(x) \to_{\text{a.s.}} 0, \qquad n \to \infty.$$

It is easy to arrive at (2.7). Let $\delta > 0$ be such that the function in (1.2) is nondecreasing on $[1 - \delta, 1)$. Given such a $\delta > 0$ and n large enough, we get

(2.9)
$$\Lambda_{3} \leq \sup_{x \in [0, F^{-1}(1-\delta))} q(F(x)) \psi_{1}(x)$$

$$+ \sup_{x \in [F^{-1}(1-\delta), X_{n:n})} q(F(x)) \psi_{1}(x)$$

$$=: \Lambda_{5} + \Lambda_{6}.$$

Applying first the assumption (1.1) and then the Glivenko–Cantelli theorem, we conclude that $\Lambda_5 \to_{\text{a.s.}} 0$ as $n \to \infty$. Furthermore, since the function in (1.2) is nondecreasing on $[1 - \delta, 1)$ and (1.3) is assumed, using the weighted Glivenko–Cantelli theorem due to Lai (1974) and Wellner (1977), we get that $\Lambda_6 \to_{\text{a.s.}} 0$, $n \to \infty$. Now (2.9) completes the proof of (2.7).

Proving now (2.8), with the same $\delta > 0$ as above and n large enough, we get

$$\begin{array}{ll} \Lambda_{4} \leq \sup_{x \in [\,0,\,F^{-1}(1-\,\delta))} q\big(F(\,x)\big)\psi_{2}(\,x) \\ &+ \sup_{x \in [\,0,\,F^{-1}(1-\,\delta),\,X_{n\,:\,n})} q\big(F(\,x)\big)\psi_{2}(\,x) \\ &=: \Lambda_{7} + \Lambda_{8}. \end{array}$$

Because of (1.1), we have

(2.11)
$$\Lambda_7 \leq c \int_0^\infty \left| F_n(y) - F(y) \right| dy,$$

and by Kolmogorov's strong law of large numbers in (the separable Banach space) $L_1(0,\infty)$ we get $\int_0^\infty |F_n(y)-F(y)|\,dy\to_{\rm a.s.} 0,\ n\to\infty$, since $\mathbf{E} X<\infty$ is assumed [cf. Corollary 7.10 on page 189 of Ledoux and Talagrand (1991)]. Hence by (2.11), $\Lambda_7\to_{\rm a.s.} 0,\ n\to\infty$. For showing that $\Lambda_8\to_{\rm a.s.} 0$ as $n\to\infty$, we denote

$$\psi_3(x) := \frac{q(F(x))}{1 - F(x)} \log_2 \frac{1}{1 - F(x)}.$$

Then, on account of the function in (1.2) assumed to be nondecreasing on $[1 - \delta, 1)$, we have

(2.12)
$$\Lambda_8 \le \int_{F^{-1}(1-\delta)}^{x_F} \psi_3(x) |F_n(x) - F(x)| dx.$$

By Kolmogorov's strong law of large numbers in $L_1(0,\infty)$ and assumption (1.3), we get that the right-hand side of (2.12) goes to 0 a.s. as $n\to\infty$. Hence, $\Lambda_8\to_{\mathrm{a.s.}} 0$ and, via (2.10), we have $\Lambda_4\to_{\mathrm{a.s.}} 0$, as $n\to\infty$. This also completes the proof of Theorem 1.1. \square

PROOF OF COROLLARY 1.1. Let the weight function q be as follows: q(F(x)) = 1 for all $x \le x_0$ and

$$q(F(x)) = \{1 - F(x)\}/\log_2(1/\{1 - F(x)\})$$

for all $x > x_0$. Theorem 1.1 implies that $\Delta_n(q, F) \to_{\text{a.s.}} 0$ as $n \to \infty$. However,

(2.13)
$$\Delta_n(q, F) \ge \sup_{x \in [0, x_0]} |M_n(x) - M_F(x)|,$$

and hence the corollary. \Box

PROOF OF COROLLARY 1.2. Note that with the function q as in (1.5), $\Delta(q,F) \leq \int_0^1 M_F(Q(t)) \, dt$. Since $\mathbf{E} X < \infty$, we have that $\int_0^1 M_F(Q(t)) \, dt < \infty$ by Hardy's inequality:

(2.14)
$$\int_0^1 \left| \frac{1}{t} \int_0^t f(s) \, ds \right|^p \, dt \le c_p \int_0^1 |f(t)|^p \, dt.$$

For details we refer to CsCsH [(1986), page 40]. \Box

PROOF OF COROLLARIES 1.3 AND 1.4. Use Theorem 1.1 and Hardy's inequality (2.14) with p=1. \square

PROOF OF THEOREM 1.2. We have

$$\begin{split} \nabla_{n}(q,F) & \leq \sup_{x \in [0,X_{n:n})} q(F(x)) \Big| \sqrt{n} \left\{ M_{n}(x) - M_{F}(x) \right\} - G_{n}(x) \Big| \\ & + \sup_{x \in [X_{n:n},\infty)} q(F(x)) \Big| \sqrt{n} \left\{ M_{n}(x) - M_{F}(x) \right\} - G_{n}(x) \Big| \\ & =: \Lambda_{9} + \Lambda_{10}. \end{split}$$

We first prove that $\Lambda_9 \to_P 0$ as $n \to \infty$. Denote

$$e_n(t) = \sqrt{n} \left\{ F_n(Q(t)) - t \right\},\,$$

where Q stands for the quantile function corresponding to F and let

$$M_n^1(x) := \frac{1}{1 - F(x)} \Big\{ e_n(F(x)) M_F(x) - \int_x^{\infty} e_n(F(y)) dy \Big\}.$$

With this notation we have

$$(2.15) \quad \sqrt{n} \left\{ M_n(x) - M_F(x) \right\} = \left\{ \frac{1 - F(x)}{1 - F_n(x)} - 1 \right\} M_n^1(x) + M_n^1(x)$$

for all $x \in [0, X_{n:n})$. The representation (2.15) implies

$$\Lambda_{9} \leq \sup_{x \in [0, X_{n:n})} \frac{q(F(x))M_{F}(x)}{1 - F(x)} |e_{n}(F(x)) - B_{n}(F(x))|$$

$$+ \sup_{x \in [0, X_{n:n})} \frac{q(F(x))}{1 - F(x)} \left| \int_{x}^{\infty} \{e_{n}(F(y)) - B_{n}(F(y))\} dy \right|$$

$$+ \sup_{x \in [0, X_{n:n})} \left| \frac{1 - F(x)}{1 - F_{n}(x)} - 1 \left| \frac{q(F(x))}{1 - F(x)} |e_{n}(F(y))| M_{F}(x) \right|$$

$$+ \sup_{x \in [0, X_{n:n})} \left| \frac{1 - F(x)}{1 - F_{n}(x)} - 1 \left| \frac{q(F(x))}{1 - F(x)} |\int_{x}^{\infty} e_{n}(F(y)) dy \right|$$

$$=: A_{1} + A_{2} + A_{3} + A_{4},$$

where the Brownian bridges $\{B_n(t): 0 \le t \le 1\}$ are the ones constructed in Theorem 1.1 of CsCsHM (1986). We now prove $A_i \to_P 0$, as $n \to \infty$, for all $i = 1, \ldots, 4$.

Step 1 $(A_1 \to_P 0, n \to \infty)$. Since the function in (1.9) is nondecreasing in a neighborhood of 1, $t \mapsto q(t) M_F(Q(t))/(1-t)$ is a Chibisov–O'Reilly function by (1.10). Hence $A_1 \to_P 0$, $n \to \infty$, by CsCsHM [(1986), Theorem 4.2.1, page 65]. This ends Step 1.

Step 2 $(A_2 \rightarrow_P 0, n \rightarrow \infty)$. We start with an elementary bound

$$A_{2} \leq \sup_{t \in (0, 1 - \delta)} \frac{q(t)}{1 - t} \left| \int_{t}^{1 - \delta} \{e_{n}(s) - B_{n}(s)\} dQ(s) \right|$$

$$+ 2 \sup_{t \in [1 - \delta, 1)} \frac{q(t)}{1 - t} \left| \int_{t}^{1} e_{n}(s) dQ(s) \right|$$

$$+ 2 \sup_{t \in [1 - \delta, 1)} \frac{q(t)}{1 - t} \left| \int_{t}^{1} B_{n}(s) dQ(s) \right|$$

$$=: A_{21} + 2A_{22} + 2A_{23}, \quad \delta \in (0, 1).$$

Because of assumption (1.1) and the fact that $\sup_{t \in (0,1)} |e_n(t) - B_n(t)| \to_P 0$ as $n \to \infty$, we conclude that $A_{21} \to_P 0$, $n \to \infty$. We now show that for every ε_1 , $\varepsilon_2 > 0$ there exists a $\delta > 0$ and $N \in \mathbf{N}$ such that

(2.18)
$$\mathbf{P}\{I > \varepsilon_2\} < \varepsilon_1, \qquad I \in \{A_{22}, A_{23}\}$$

for all n > N

Clearly, (2.18) will also complete the proof of $A_2 \to_P 0$, $n \to \infty$. Since the proof of (2.18) is the same for both $I = A_{22}$ and $I = A_{23}$, we shall carry it out

in the case of $I=A_{23}$ only. Toward this we need some elementary calculations. We have

$$\frac{q(t)}{1-t} \left| \int_{t}^{1} B_{n}(s) dQ(s) \right| \\
= \frac{q(t)}{1-t} \left| \int_{t}^{1} \frac{1-s}{q(s)} d_{s} \left\{ \int_{1-\delta}^{s} B_{n}(v) \frac{q(v)}{1-v} dQ(v) \right\} \right| \\
\leq 2(c+1) \sup_{s \in [t,1)} \left| \int_{1-\delta}^{s} B_{n}(v) \frac{q(v)}{1-v} dQ(v) \right|,$$

where, on account of (1.9), the constant c > 0 is such that

$$\frac{q(t)}{1-t}\lim_{s\to 1}\frac{1-s}{q(s)}\le c$$

for all $t \in [1 - \delta, 1)$. Using (2.19), with the same constant c > 0 we arrive at

$$(2.20) I \leq 2(c+1) \sup_{t \in [1-\delta,1)} \left| \int_{1-\delta}^{t} B_n(s) \frac{q(s)}{1-s} dQ(s) \right|.$$

Furthermore,

$$\int_{1-\delta}^{t} B_{n}(s) \frac{q(s)}{1-s} dQ(s) = \int_{1-\delta}^{t} \frac{B_{n}(s)}{1-s} d_{s} \left\{ -\int_{s}^{1} q(z) dQ(z) \right\}
= -\frac{B_{n}(s)}{1-s} \int_{s}^{1} q(z) dQ(z) \Big|_{1-\delta}^{t}
+ \int_{1-\delta}^{1} \left\{ \int_{s}^{1} q(z) dQ(z) \right\} d\frac{B_{n}(s)}{1-s}.$$

Since the function $t\mapsto \psi_4(t):=\int_t^1 q(z)\,dQ(z)/(1-t)$ is square integrable by (1.10), it is a Chibisov–O'Reilly function. Thus, $B_n(t)\psi_4(t)\to_P 0$ as $t\to 1$ [cf., e.g., CsCsHM (1986) or see M. Csörgő, Shao and Szyszkowicz (CsShSz) (1991) and M. Csörgő and Horváth (1993) on this and related problems]. Thus, on combining the latter with (2.21) and (2.20), we see that (2.18) follows from

$$(2.22) J \coloneqq \mathbf{P} \Big\{ \sup_{t \in [1-\delta,1)} |C_n(t)| > \varepsilon_2 \Big\} < \varepsilon_1,$$

where

$$C_n(t) \coloneqq \int_{1-\delta}^t \left\langle \int_s^1 q(z) \ dQ(z) \right
brace d rac{B_n(s)}{1-s}.$$

Fix $z \in [0,1]$ for a moment and let $\mathcal{F}(z)$ denote the σ -algebra generated by the random variables $\{B_n(v), v \in [0,z]\}$. It is well known that $\{(B_n(z)/(1-z), \mathcal{F}_n(z)): z \in [0,1)\}$ is a martingale for each $n \in \mathbf{N}$. Thus, for each $n \in \mathbf{N}$, $\{(C_n(z), \mathcal{F}_n(z)): z \in [0,1)\}$ is also a martingale. Therefore, and because of

(1.10), we may use the Birnbaum-Marshall inequality to estimate the quantity J in (2.22). We get that J does not exceed (up to a constant) the quantity

(2.23)
$$\int_{1-\delta}^{1} \left\{ \int_{z}^{1} q(v) \, dQ(v) \right\}^{2} \, d_{z} \left\{ \int_{0}^{z} \frac{1}{(1-u)^{2}} \, du \right\}$$

$$= \int_{1-\delta}^{1} \left\{ \frac{1}{1-z} \int_{z}^{1} q(v) \, dQ(v) \right\}^{2} \, dz,$$

which clearly goes to 0 as $\delta \to 1$, since (1.10) is assumed. This completes the proof of (2.18) and that of the statement $A_2 \to_P 0$, $n \to \infty$, as well. We note in passing that for an interpretation of $z \mapsto \int_0^z (1-u)^{-2} du$ in (2.23) one may have a look at Al-Hussaini and Elliot [(1984), top of page 612].

Step 3 $(A_3 \to_P 0, n \to \infty)$. Fix any $\delta > 0$. Since $X_{n:n} \to_{\text{a.s.}} x_F$ as $n \to \infty$, the statement $A_3 \to_P 0, n \to \infty$, reduces to the statement $1\{X_{n:n} \ge F^{-1}(1 - \delta)\}A_3 \to_P 0, n \to \infty$. Hence we assume throughout that

$$(2.24) X_{n+n} \ge F^{-1}(1-\delta)$$

and show that $A_3 \to_P 0$, $n \to \infty$, under this assumption. For notational simplicity, we let

$$\psi_5(t) \coloneqq q(t) M_F(Q(t)) / (1-t)$$

and

$$a_3(x) := \left| \frac{1 - F(x)}{1 - F_n(x)} - 1 \right| \psi_5(F(x)) \left| e_n(F(x)) \right|.$$

Then

$$(2.25) A_3 = \sup_{x \in [0, X_{n:n})} a_3(x)$$

$$\leq \sup_{x \in [0, F^{-1}(1-\delta))} a_3(x) + \sup_{x \in [F^{-1}(1-\delta), X_{n:n})} a_3(x)$$

$$=: A_{31} + A_{32}.$$

On noting that, by (1.1), $\psi_5(F(x)) \le c$ for all $x \in [0, F^{-1}(1 - \delta))$ and that $\sup_{t \in (0,1)} |e_n(t)| = \mathscr{O}_P(1), \ n \to \infty$, by the Glivenko–Cantelli theorem, we conclude $A_{31} \to_P 0, \ n \to \infty$.

To complete the proof of $A_3 \to_P 0$, $n \to \infty$, we need to show that for every ε_1 , $\varepsilon_2 > 0$ there exists $\delta > 0$ and $N \in \mathbf{N}$ such that

$$(2.26) \mathbf{P}\{A_{32} \ge \varepsilon_2\} < \varepsilon_1$$

for all $n \ge N$. The main idea of the proof of (2.26) is based on the following observation: Since ψ_5 is a Chibisov–O'Reilly function, using Lemma 4.2.1 of CsCsHM [(1986), page 66] we can find a local function ψ_5^* such that

(2.27)
$$\psi_5(t)/\psi_5^*(t) \to 0 \text{ as } t \to 1.$$

We note in passing that a function ψ_5^* that satisfies the latter relationship would be called an Erdös-Feller-Kolmogorov-Petrovski (EFKP) upper-class

function in CsCsHM (1986). Calling it simply a local function here is in agreement with CsShSz [(1991), page 239] [cf. also M. Csörgő and Horváth (1993), Proposition 1.3, page 194]. As a consequence of (2.27) we have that

$$\sup_{t \in [1-\delta, 1)} \psi_5(t)/\psi_5^*(t) \to 0, \qquad \delta \to 1.$$

Combining the latter with Remark 1(ii) on page 75 of Wellner (1978), which tells us that

(2.28)
$$\sup_{x \in [0, X_{n \cdot n}]} \frac{1 - F(x)}{1 - F_n(x)} = \mathscr{O}_P(1), \qquad n \to \infty,$$

and with CsCsHM [(1986), Theorem 4.2.3], we arrive at (2.26).

Step 4 ($A_4 \rightarrow_P 0, n \rightarrow \infty$). Denote

$$a_4(x) := \left| \frac{1 - F(x)}{1 - F_n(x)} - 1 \right| \frac{q(F(x))}{1 - F(x)} \left| \int_x^{\infty} e_n(F(y)) \, dy \right|.$$

Then, for every $\delta \in (0, 1)$,

$$\begin{split} A_4 &= \sup_{x \in [0, X_{n:n})} a_4(x) \\ &\leq \sup_{x \in [0, F^{-1}(1-\delta))} a_4(x) + \sup_{x \in [F^{-1}(1-\delta), X_{n:n})} a_4(x) \\ &=: A_{41} + A_{42}. \end{split}$$

The proof of $A_{41} \to_P 0$, $n \to \infty$, goes along the lines of that of $A_{31} \to_P 0$, $n \to \infty$, in Step 3, only now use $\int_0^1 e_n(t) \, dQ(t) = \mathscr{O}_P(1)$, $n \to \infty$, instead of $\sup_{t \in (0,1)} |e_n(t)| = \mathscr{O}_P(1)$. The proof of the fact that we have $\mathbf{P}\{A_{42} > \varepsilon_2\} < \varepsilon_1$ is given in Step 2 [cf. (2.18)]. These remarks conclude the proof of $A_4 \to_P 0$, $n \to \infty$.

Steps 1–4 and (2.16) complete the proof of $\Lambda_9 \to_P 0$, $n \to \infty$. Consequently, Theorem 1.2 will be proved if we show that $\Lambda_{10} \to_P 0$ as $n \to \infty$. To start with, we note that

$$\Lambda_{10} \leq \sup_{x \in [X_{n:n}, x_{F})} q(F(x)) \sqrt{n} M_{F}(x)
+ \sup_{x \in [X_{n:n}, x_{F})} q(F(x)) |B_{n}(F(x))| M_{F}(x) / \{1 - F(x)\}
+ \sup_{x \in [X_{n:n}, x_{F})} \frac{q(F(x))}{1 - F(x)} |\int_{x}^{x_{F}} B_{n}(F(y)) dy|
=: B_{1} + B_{2} + B_{3}.$$

Fix any $\delta > 0$ such that the function in (1.9) is nondecreasing on $[1 - \delta, 1)$. Since $\mathbf{1}\{X_{n:n} < F^{-1}(1 - \delta)\} \to_{a.s.} 0$, $n \to \infty$, we assume below that

$$X_{n+n} \geq F^{-1}(1-\delta),$$

which is equivalent, of course, to $F(X_{n:n}) \ge 1 - \delta$.

Step 5 $(B_1 \to_P 0, n \to \infty)$. It is well known and easy to show that $n(1 - F(X_{n:n})) = \mathscr{O}_P(1), n \to \infty$. Hence $B_1 \to_P 0, n \to \infty$, if

$$(2.30) \sup_{x \in [X_{n:n}, x_F)} q(F(x)) M_F(x) / \sqrt{1 - F(x)} \to_P 0, \quad n \to \infty.$$

We now prove (2.30). First, note that the function in (1.9) is nondecreasing on $[1 - \delta, 1)$. Hence, $q(F(x))M_F(x) \leq \int_x^{x_F} q(F(y)) dy$. Due to (1.10), the function

$$t\mapsto \psi_6(t)\coloneqq rac{1}{1-t}\!\int_t^1\!\!q(s)\,dQ(s)$$

is a Chibisov-O'Reilly function. Hence, $\psi_6(t)\sqrt{1-t} \to 0$ as $t \to 1$ by CsCsHM [(1986), Proposition 3.1(i), page 56]. These remarks conclude the proof of (2.30), since $X_{n:n} \to_{\text{a.s.}} x_F$ as $n \to \infty$.

Step 6 $(B_2 \to 0, n \to \infty)$. Since the function $t \mapsto q(t)M_F(Q(t))/(1-t)$ is a Chibisov–O'Reilly function and $X_{n+n} \to_{\text{a.s.}} x_F$ as $n \to \infty$, we obtain $B_2 \to_P 0$, $n \to \infty$, as desired [cf. Remark 3.2 on page 56 of CsCsHM (1986)].

Step 7 $(B_3 \to_P 0, n \to \infty)$. Because of the assumption $X_{n:n} \ge F^{-1}(1-\delta)$ made above, we have

$$B_3 \le \sup_{t \in \lceil 1 - \delta, 1 \rceil} \frac{q(t)}{1 - t} \left| \int_t^1 B_n(s) \ dQ(s) \right|.$$

Now the statement (2.18) completes the proof of $B_3 \to_P 0$, $n \to \infty$.

Because of Steps 5–7 and (2.29), we have that $\Lambda_{10} \to_P 0$, $n \to \infty$. This also completes the proof of Theorem 1.2. \square

PROOF OF COROLLARY 1.5. Let q be such that q(F(x)) = 1 for all $x \in [0, x_0]$ and q(F(x)) = 1 - F(x) for all $x \ge x_0$. According to Theorem 1.2, $\nabla_n(q, F) \to_P 0$ as $n \to \infty$, which completes the proof, due to having

$$\nabla_n(q, F) \ge \sup_{x \in [0, x_0]} |M_n(x) - M_F(x)|. \qquad \Box$$

PROOF OF COROLLARY 1.6. With the weight function q defined by (1.12), we have $\nabla(q,F)=\int_0^1 M_F^2(Q(t))\,dt$. However, the latter quantity is finite because of the assumption $\mathbf{E}X^2<\infty$; this is seen by using Hardy's inequality (2.14) with p=2. For more details we refer to calculations presented on page 40 of CsCsH (1986). \square

Proof of Corollaries 1.7 and 1.8. Follows from Theorem 1.2 and Hardy's inequality (2.14) with p=2. \square

PROOF OF PROPOSITION 1.1. Since F is the Exp(1)-distribution function, we have $M_F(Q(t)) = 1$ for all $t \in (0,1)$. Also, $M_n(Q(t)) = 0$ for all $t \in [U_{n:n},1)$ where $U_{n:n} := F(X_{n:n})$. Thus, with the weight function q as in the proposition,

$$(2.31) \Delta_n(q, F) \ge I(U_{n+n}),$$

where, for every $\varepsilon \in (0, 1)$,

$$I(arepsilon)\coloneqq \sup_{t\in [arepsilon,1)} \sqrt{1-t} \left| \sqrt{n}
ight. + rac{B_n(t)}{1-t} - rac{1}{1-t} \int_t^1 rac{B_n(s)}{1-s} \, ds
ight|.$$

Because of (2.31),

(2.32)
$$\Lambda \ge \liminf_{n \to \infty} \mathbf{P} \{ I(U_{n:n}) \ge \frac{1}{2} \}.$$

It is easy to show that, for any fixed $d \in (0, 1]$,

$$\mathbf{P}\Big\{I(U_{n:n}) \geq \frac{1}{2}\Big\} \geq \mathbf{P}\Big\{I\Big(1 - \frac{d}{n}\Big) \geq \frac{1}{2}\Big\} - \mathbf{P}\Big\{U_{n:n} \geq 1 - \frac{d}{n}\Big\}$$

and that $\mathbf{P}\{U_{n:n} \geq 1 - d/n\}$ converges to $1 - e^{-d}$ as $n \to \infty$. Thus, in view of (2.32),

$$(2.33) \qquad \Lambda \geq \Lambda_{11} - (1-e^{-d}), \quad \text{where } \Lambda_{11} \coloneqq \liminf_{n \to \infty} \mathbf{P} \bigg\{ I \bigg(1 - \frac{d}{n} \bigg) \geq \frac{1}{2} \bigg\}.$$

Let us suppose that we have proved the inequality

$$(2.34) \qquad \qquad \Lambda_{11} \ge c_0 > 0$$

for a universal constant $c_0 > 0$. Then (2.33) implies the bound

$$\Lambda \geq c_0 - (1 - e^{-d}),$$

which is the claim of Proposition 1.1 if we take d > 0 small enough. Hence, in order to complete this proof, it suffices to show (2.34). Denote

$$W(t) := B(1-t) - \int_0^t B(1-s)/s \, ds,$$

where $B(\cdot)$ is a Brownian bridge. It is well known and easy to show that W is a Wiener process on [0,1]. Thus, $\Gamma := \sqrt{n/d} \, W(d/n)$ is a Gaussian random variable with mean 0 and variance 1. Now, with this notation and remarks we easily arrive at the following bounds:

$$\mathbf{P}\left\{I\left(1-\frac{d}{n}\right) \geq \frac{1}{2}\right\} = \mathbf{P}\left\{\sup_{t\in[0,\,d/n]} \left|\sqrt{nt} + \frac{W(t)}{\sqrt{t}}\right| \geq \frac{1}{2}\right\}$$
$$\geq \mathbf{P}\left\{\sup_{t\in[0,\,d/n]} \left|\frac{W(t)}{\sqrt{t}}\right| - \sqrt{d} \geq \frac{1}{2}\right\}$$
$$\geq \mathbf{P}\left\{|\Gamma| \geq \frac{1}{2} + \sqrt{d}\right\} \geq \mathbf{P}\left\{|\Gamma| \geq \frac{3}{2}\right\}$$

as claimed in (2.34) with $c_0 = \mathbf{P}\{|\Gamma| \ge \frac{3}{2}\}$. \square

PROOF OF THEOREM 1.3. The following elementary inequality holds for all random variables ξ and η , and numbers z and $\delta \geq 0$:

(2.35)
$$\begin{aligned} |\mathbf{P}\{|\xi| \le z\} - \mathbf{P}\{|\eta| \le z\}| \\ \le \mathbf{P}\{|\eta| \in (z - \delta, z + \delta]\} + \mathbf{P}\{|\xi - \eta| \ge \delta\}. \end{aligned}$$

Throughout the proof we use the following values of ξ , η and z:

$$\begin{split} \xi &\coloneqq \sqrt{n} \sup_{x \in [0, X_{n:n})} q(F_n(x)) \big| M_n(x) - M_F(x) \big|, \\ \eta &\coloneqq \sqrt{n} \sup_{x \in [0, x_F)} q(F(x)) \big| G_n(x) \big|, \\ z &\coloneqq z_{\alpha}. \end{split}$$

(The value of $\delta > 0$ is not fixed yet.) We continue with several remarks, leading to the proof of the theorem. First, we have $\mathbf{P}\{|\eta| \leq z\} = 1 - \alpha$ by definition. Second, $\mathbf{P}\{|\eta| \in (z_{\alpha} - \delta, z_{\alpha} + \delta]\}$ converges to 0 as $\delta \to 0$ by continuity. Third, assuming that we can show

then, because of the first two remarks and the inequality (2.35), the theorem is proved. Consequently, we are left with proving (2.36). Denote

$$\Lambda_{12} \coloneqq \sup_{x \in [0, X_{n:n})} |q(F_n(x)) - q(F(x))| \sqrt{n} |M_n(x) - M_F(x)|.$$

Now it is evident that (2.36) is a consequence of Theorem 1.2 and the fact that $\Lambda_{12} \to_P 0$ ($n \to \infty$), which, in turn, is a consequence of

$$\Lambda_{13}, \Lambda_{14} \rightarrow_P 0, \qquad n \rightarrow \infty,$$

where

$$\Lambda_{13} := \sup_{x \in [0, X_{n:n})} |q(F_n(x)) - q(F(x))| M_F(x) |e_n(F(x))| / \{1 - F_n(x)\},$$

$$\Lambda_{14} := \sup_{x \in [0, X_{n:n})} |q(F_n(x)) - q(F(x))| \left| \int_x^\infty e_n(F(y)) \, dy \right| / \{1 - F_n(x)\}.$$

Hence Steps 8 and 9 deal with showing this convergence in probability to zero for the two random variables Λ_{13} and Λ_{14} .

Step 8 ($\Lambda_{13} \to 0$, $n \to \infty$). Let the function inside the sup of Λ_{13} be denoted by ψ_{13} , that is to say, $\Lambda_{13} = \sup_{x \in [0, X_{n:n})} \psi_{13}(x)$. We may again assume without loss of generality that, for some $\delta \in (0,1)$ and n large,

$$X_{n+n} \geq F^{-1}(1-\delta)$$
.

Thus,

$$(2.37) \quad \Lambda_{13} \leq \sup_{x \in [0, F^{-1}(1-\delta))} \psi_{13}(x) + \sup_{x \in [F^{-1}(1-\delta), X_{n:n})} \psi_{13}(x) =: C_1 + C_2.$$

The proof of $C_1 \to_P 0$, $n \to \infty$, is based on

(2.38)
$$\sup_{x \in [0, F^{-1}(1-\delta)]} |q(F_n(x)) - q(F(x))| \to_P 0, \quad n \to \infty,$$

which follows from the Glivenko–Cantelli theorem and the continuity assumption (1.19), and on having $\sup_{t \in (0,1)} |e_n(t)| = \mathscr{O}_P(1), n \to \infty$.

To complete the proof of $\Lambda_{13} \to_P 0$, $n \to \infty$, we show that for every ε_1 , $\varepsilon_2 > 0$ there exists a $\delta > 0$ and $N \in \mathbf{N}$ such that

$$(2.39) \mathbf{P}\{C_2 \ge \varepsilon_2\} < \varepsilon_1$$

for all $n \geq N$. The main ingredients of the proof of (2.39) are assumption (1.20) and Inequality 1(3) on page 419 of Shorack and Wellner (1986), which states that for every $\varepsilon > 0$ there exists a subset $\Omega_{\varepsilon} \subset \Omega$ (\equiv the set of all elementary events) such that $\mathbf{P}(\Omega_{\varepsilon}) < \varepsilon$ and, on Ω_{ε} ,

$$(2.40) F_n(x) \le 1 - c(1 - F(x)) \text{for all } x \in [0, x_F),$$

where $c \in (0, 1)$ is a (nonrandom) constant. Now, the proof of (2.39) can be completed as follows. Since $\delta > 0$ can be taken arbitrarily small, we may assume the function in (1.9) to be nondecreasing on $[1 - \delta, 1)$. It is easy to see that for proving (2.39) it is enough to show that

(2.41)
$$\mathbf{P}(\{C_2 \geq \varepsilon_2\} \cap \Omega_{\varepsilon_1}) < \varepsilon_1.$$

However, on the set Ω_{ε_1} we have the bounds

$$(2.42) \frac{q(F_n(x))}{1 - F_n(x)} \le \frac{q(1 - c(1 - F(x)))}{c(1 - F(x))} \le C \frac{q(F(x))}{1 - F(x)}$$

for all $x \in [F^{-1}(1-\delta), X_{n:n})$, where the first and second inequalities in (2.42) are due to (2.40) and (1.20), respectively. Consequently, on the set Ω_{ε} ,

$$(2.43) C_2 \le c \sup_{x \in [F^{-1}(1-\delta), X_{n:n})} q(F(x)) M_F(x) |e_n(F(x))| / \{1 - F(x)\},$$

where constant $c \ge 0$ is nonrandom and does not depend on n. The estimate (2.43) together with (2.26) completes the proof of (2.41) and that of the statement (2.39) as well. This completes the proof of $\Lambda_{13} \to_P 0$, $n \to \infty$.

Step 9 ($\Lambda_{14} \to_P 0$, $n \to \infty$). The proof of $\Lambda_{14} \to_P 0$, $n \to \infty$, is very similar to that of $\Lambda_{13} \to_P 0$, $n \to \infty$. The main difference is the use of bound (2.18) instead of (2.26). We omit further details which, in the light of the above calculations, are elementary.

This completes the proof of $\Lambda_{12} \to_P 0$, $n \to \infty$, as well as that of Theorem 1.3. \square

Proof of Proposition 1.2. Because of (1.23), we have $\Delta(q, F) < \infty$ if

(2.44)
$$\int_0^1 \left\{ \frac{1}{1-t} \int_t^1 (1-s)^{\nu} dQ(s) \right\}^2 dt < \infty.$$

Using the integration by parts formula, we get that (2.44) follows from

(2.45)
$$\int_0^1 \left\{ \frac{1}{1-t} \int_t^1 Q(s) (1-s)^{\nu-1} ds \right\}^2 dt < \infty.$$

Using now Hardy's inequality (2.14) with p=2, we get that (2.45) is a consequence of

(2.46)
$$\int_0^\infty s^2 (1 - F(s))^{2\nu - 2} dF(x) < \infty.$$

However, $1 - F(s) \le s^{-2/(2\nu-1)} \mathbf{E} X^{2/(2\nu-1)}$. Hence assumption (1.22) completes the proof of (2.46) and that of the proposition as well. \square

Acknowledgments. Lectures by Hao Yu on empirical reliability renewed our interest in solving the problems discussed in this paper. R. Norvaiša made several helpful comments. J. A. Wellner acquainted us with the very interesting unpublished paper by Hall and Wellner (1979). The Editor, M. Woodroofe, an Associate Editor and two referees made a number of very useful comments and suggestions, and pointed out several oversights, misprints and a mistake as well. We have incorporated their suggestions in various places in our revision and hope to have succeeded in producing a much improved version of our original manuscript. We sincerely thank them all for their time, as well as for their careful attention to and constructive criticism of our paper.

REFERENCES

AL-HUSSAINI, A. and ELLIOTT, R. J. (1984). The single jump process with some statistical applications. *Theory Probab. Appl.* **29** 607–613.

Andersen, P. K., Borgan, Ø., Gill, R. D. and Keiding, N. (1993). Statistical Models Based on Counting Processes. Springer, New York.

Berman, S. M. (1963). A note on extreme values, competing risks and semi-Markov processes.

Ann. Math. Statist. 34 1104-1106.

BICKEL, P. J. and FREEDMAN, D. A. (1981). Some asymptotic theory for the bootstrap. *Ann. Statist.* **9** 1196–1217.

Bickel, P. J., Klaassen, C. A. J., Ritov, Y. and Wellner, J. A. (1993). *Efficient and Adaptive Estimation in Semiparametric Models*. Johns Hopkins Univ. Press.

Burke, M. D., Csörgő, S. and Horváth, L. (1981). Strong approximations of some biometric estimates under random censorship. Z. Wahrsch. Verw. Gebiete 56 87–112.

Burke, M. D., Csörgő, S. and Horváth, L. (1988). A correction to and improvement of "Strong approximations of some biometric estimates under random censorship." *Probab. Theory Related Fields* **79** 51–57.

Chung, C.-J. F. (1989). Confidence bands for percentile residual lifetime under random censorship model. J. Multivariate Anal. 29 94–126.

CROWLEY, J. and JOHNSON, R. A. (1982). Survival Analysis. IMS, Hayward, CA.

CSÖRGŐ, M., CSÖRGŐ, S. and HORVÁTH, L. (1986). An Asymptotic Theory for Empirical Reliability and Concentration Processes. Lecture Notes in Statist. 33. Springer, New York.

Csörgő, M., Csörgő, S., Horváth, L. and Mason, D. (1986). Weighted empirical and quantile processes. *Ann. Probab.* 14 86–118.

CSÖRGŐ, M., CSÖRGŐ, S. and MASON, D. M. (1984). Applications of a new approximation of uniform empirical and quantile processes: bootstrapping empirical functions. Technical Report 42, Laboratory for Research Statistics and Probability, Carleton Univ.

- Csörgő, M. and Horváth, L. (1993). Weighted Approximations in Probability and Statistics. Wiley, Chichester.
- Csörgő, M., Shao, Q.-M. and Szyszkowicz, B. (1991). A note on local and global functions of a Wiener process and some Rényi-type statistics. *Studia Sci. Math. Hungar.* **26** 239–259.
- CSÖRGŐ, S. and MASON, D. M. (1989). Bootstrapping empirical functions. Ann. Statist. 17 1447–1471.
- EFRON, B. (1979). Bootstrap methods: another look at the jackknife. Ann. Statist. 7 1-26.
- HALL, W. J. and WELLNER, J. A. (1979). Estimation of mean residual life. Unpublished manuscript.
- HALL, W. J. and WELLNER, J. A. (1980). Confidence bands for a survival curve from censored data. Biometrika 67 133-143.
- HALL, W. J. and WELLNER, J. A. (1981). Mean residual life. In Statistics and Related Topics (M. Csörgő et al., eds.) 169–184. North-Holland, Amsterdam.
- KAPLAN, E. L. and MEIER, P. (1958). Nonparametric estimation from incomplete observations. J. Amer. Statist. Assoc. 53 457–481.
- KRISHNAIAH, P. R. and RAO, C. R. (1988). Quality control and reliability. Handbook of Statistics 7. North-Holland, Amsterdam.
- Lai, T. L. (1974). Convergence rates in the strong law of large numbers for random variables taking values in Banach spaces. *Bull. Inst. Math. Acad. Sinica* 2 67–85.
- LEDOUX, M. and TALAGRAND, M. (1991). Probability in Banach Spaces. Springer, New York.
- MILLER, R. G., JR. (1981). Survival Analysis. Wiley, New York.
- MILLER, R. G., JR., EFRON, B., BROWN, B. W., JR. and Moses, L. E. (1980). Biostatistics Casebook. Wiley, New York.
- PROSCHAN, F. and SERFLING, R. J. (1974). Reliability and Biometry—Statistical Analysis of Lifelength. SIAM, Philadelphia, PA.
- Shorack, G. R. and Wellner, J. A. (1986). Empirical Processes with Applications to Statistics. Wiley, New York.
- TSIREL'SON, J. L. (1975). The density of the distribution of the maximum of a Gaussian process. Theory Probab. Appl. 20 847–856.
- WELLNER, J. A. (1977). A Glivenco-Cantelli theorem and strong laws of large numbers for functions of order statistics. Ann. Statist. 5 473-480.
- Wellner, J. A. (1978). Limit theorems for the ratio of the empirical distribution function to the true distribution function. Z. Wahrsch. Verw. Gebiete 45 73-88.
- 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.
- Yu, H. (1993). Weak convergence for empirical and quantile processes of associated sequences with applications to reliability and economics. Ph.D. thesis, Carleton Univ.

DEPARTMENT OF MATHEMATICS AND STATISTICS CARLETON UNIVERSITY 1125 COLONEL BY DRIVE OTTAWA CANADA K1S 5B6