IMS Collections
Volume Title
Vol. 0 (0000) 1–6

arXiv: `math.PR/0000000`

Estimation of Mean Residual Life

W. J. Hall[1] **and Jon A. Wellner**[2,*]

University of Rochester and University of Washington

Abstract: Yang (1978) considered an empirical estimate of the mean residual life function on a fixed finite interval. She proved it to be strongly uniformly consistent and (when appropriately standardized) weakly convergent to a Gaussian process. These results are extended to the whole half line, and the variance of the the limiting process is studied. Also, nonparametric simultaneous confidence bands for the mean residual life function are obtained by transforming the limiting process to Brownian motion.

Contents

1. Introduction and summary

This is an updated version of a Technical Report, Hall and Wellner (1979), that has been referenced repeatedly in the literature — e.g., Csörgo and Zitikis (1996), Berger et al. (1988), CsörgHo et al. (1986), Hu et al. (2002), Kochar et al. (2000), Qin and Zhao (2007) — although not having been published. We take this opportunity to honor the many achievements of Andrei Yakovlev and his devotion to modeling life processes throughout his career by making this work broadly available.

Let $X_1, \ldots, X_n$ be a random sample from a continuous d.f. F on $\mathbb{R}^+ = [0, \infty)$ with finite mean $\mu = E(X)$, variance $\sigma^2 \le \infty$, and density $f(x) > 0$. Let $\overline{F} =$

*Supported in part by NSF Grant DMS-1104832 and the Alexander von Humboldt Foundation
[1]Department of Biostatistics, University of Rochester, Rochester, NY e-mail: `hall@bst.rochester.edu`
[2]Statistics Box 354322
University of Washington
Seattle, WA 98195-4322 e-mail: `jaw@stat.washington.edu`
AMS 2000 subject classifications: Primary 62P05, 62N05, 62G15; secondary 62G05, 62E20
Keywords and phrases: life expectancy, consistency, limiting Gaussian process, confidence bands

$1 - F$ denote the survival function, let $\mathbb{F}_n$ and $\overline{\mathbb{F}}_n$ denote the empirical distribution function and empirical survival function respectively, and let

$$e(x) \equiv e_F(x) \equiv E(X - x | X > x) = \int_x^\infty \overline{F} dI / \overline{F}(x), \quad 0 \leq x < \infty$$

denote the *mean residual life function* or *life expectancy function* at age x. We use a subscript F or $\overline{F}$ on e interchangeably, and I denotes the identity function and Lebesgue measure on $\mathbb{R}^+$.

A natural nonparametric or life table estimate of e is the random function $\hat{e}_n$ defined by

$$\hat{e}_n(x) = \left\{ \int_x^\infty \overline{\mathbb{F}}_n dI / \overline{\mathbb{F}}_n(x) \right\} 1_{[0, X_{nn})}(x)$$

where $X_{nn} \equiv \max_{1 \leq i \leq n} X_i$; that is, the average, less x, of the observations exceeding x. Yang (1978) studied $\hat{e}_n$ on a fixed finite interval $0 \leq x \leq T < \infty$. She proved that $\hat{e}_n$ is a strongly uniformly consistent estimator of e on $[0, T]$, and that, when properly centered and normalized, it converges weakly to a certain limiting Gaussian process on $[0, T]$.

We first extend Yang's (1978) results to all of $\mathbb{R}^+$ by introducing suitable metrics. Her consistency result is extended in Theorem 2.1 by using the techniques of Wellner (1977, 1978); then her weak convergence result is extended in Theorem 2.2 using Shorack (1972) and Wellner (1978).

It is intuitively clear that the variance of $\hat{e}_n(x)$ is approximately $\sigma^2(x)/n(x)$ where $\sigma^2(x) = Var[X - x | X > x]$ is the residual variance and $n(x)$ is the number of observations exceeding x; the formula would be justified if these $n(x)$ observations were a random sample of fixed sized $n(x)$ from the conditional distribution $P(\cdot | X > x)$. Noting that $\overline{\mathbb{F}}_n(x) = n(x)/n \to \overline{F}(x)$ a.s., we would than have

$$n Var[\hat{e}_n(x)] = n\sigma^2(x)/n(x) \to \sigma^2(x)/\overline{F}(x).$$

Proposition 2.1 and Theorem 2.2 validate this (see (2.4) below): the variance of the limiting distribution of $n^{1/2}(\hat{e}_n(x) - e(x))$ is precisely $\sigma^2(x)/\overline{F}(x)$.

In Section 3 simpler sufficient conditions for Theorems 2.1 and 2.2 2 are given and the growth rate of the variance of the limiting process for large x is considered; these results are related to those of Balkema and de Haan (1974). Exponential, Weibull, and Pareto examples are considered in Section 4.

In Section 5, by transforming (and reversing) the time scale and rescaling the state space, we convert the limit process to standard Brownian motion on the unit interval (Theorem 5.1); this enables construction of nonparametric simultaneous confidence bands for the function e_F (Corollary 5.2). Application to survival data of guinea pigs subject to infection with tubercle bacilli as given by Bjerkedal (1960) appears in Section 6.

We conclude this section with a brief review of other previous work. Estimation of the function e, and especially the discretized life-table version, has been considered by Chiang; see pages 630-633 of Chiang (1960) and page 214 of Chiang (1968). (Also see Chiang (1968), page 189, for some early history of the subject.) The basis for *marginal* inference (i.e. at a specific age x) is that the estimate $\hat{e}_n(x)$ is approximately normal with estimated standard error $S_k/\sqrt{k}$, where $k = n\overline{\mathbb{F}}_n(x)$ is the observed number of observations beyond x and S_k is the sample standard deviation of those observations. A partial justification of this is in Chiang (1960), page 630, (and is made precise in Proposition 5.2 below). Chiang (1968), page 214, gives the analogous marginal result for grouped data in more detail, but again without proofs; note the solumn $S_{\hat{e}_i}$ in his Table 8, page 213, which is based on a modification and correction of a variance formula due to Wilson (1938). We know of no earlier work on simultaneous inference (confidence bands) for mean residual life.

A plot of (a continuous version of) the estimated mean residual life function of 43 patients suffering from chronic gramulocytic leukemia is given by Bryson and Siddiqui (1969). Gross and Clark (1975) briefly mention the estimation of e in a life - table setting, but do not discuss the variability of the estimates (or estimates thereof). Tests for exponentiality against decreasing mean residual life alternatives have been considered by Hollander and Proschan (1975).

2. Convergence on $\mathbb{R}^+$; covariance function of the limiting process

Let $\{a_n\}_{n\geq 1}$ be a sequence of nonnegative numbers with $a_n \to 0$ as $n \to \infty$. For any such sequence and a d.f. F as above, set $b_n = F^{-1}(1 - a_n) \to \infty$ as $n \to \infty$. Then, for any function f on $\mathbb{R}^+$, define f^* equal to f for $x \leq b_n$ and 0 for $x > b_n$: $f^*(x) = f(x)1_{[0,b_n]}(x)$. Let $\|f\|_a^b \equiv \sup_{a\leq x\leq b} |f(x)|$ and write $\|f\|$ if $a = 0$ and $b = \infty$.

Let $\mathcal{H}(\downarrow)$ denote the set of all nonnegative, decreasing functions h on $[0, 1]$ for which $\int_0^1 (1/h)dI < \infty$.

Condition 1a. There exists $h \in \mathcal{H}(\downarrow)$ such that

$$M_1 \equiv M_1(h, F) \equiv \sup_x \frac{\int_x^\infty h(F)dI/h(F(x))}{e(x)} < \infty.$$

Since $0 < h(0) < \infty$ and $e(0) = E(X) < \infty$, Condition 1a implies that $\int_0^\infty h(F)dI < \infty$. Also note that $h(F)/h(0)$ is a survival function on $\mathbb{R}^+$ and that the numerator in Condition 1a is simply $e_{h(F)/h(0)}$; hence Condition 1a may be rephrased as: there exists $h \in \mathcal{H}(\downarrow)$ such that $M_1 \equiv \|e_{h(F)/h(0)}/e_F\| < \infty$.

Condition 1b. There exists $h \in \mathcal{H}(\downarrow)$ for which $\int_0^\infty h(F)dI < \infty$ and $\|eh(F)\| < \infty$.

Bounded e_F and existence of a moment of order greater than 1 is more than sufficient for Condition 1b (see Section 3).

Theorem 2.1. *Let $a_n = \alpha n^{-1} \log\log n$ with $\alpha > 1$. If Condition 1a holds for a particular $h \in \mathcal{H}(\downarrow)$, then*

$$\begin{aligned} &\rho_{h(F)e/\overline{F}}(\hat{e}_n^*, e^*) \\ &\quad\equiv \sup\left\{\frac{|\hat{e}_n(x) - e(x)|\overline{F}(x)}{h(F(x))e(x)}\colon\ x \leq b_n\right\} \to_{a.s.} 0 \quad as \quad n \to \infty. \end{aligned} \tag{2.1}$$

If Condition 1b holds, then

$$\begin{aligned} &\rho_{1/\overline{F}}(\hat{e}_n^*, e^*) \\ &\quad\equiv \sup\{|\hat{e}_n(x) - e(x)|\overline{F}(x)\colon\ x \leq b_n\} \to_{a.s.} 0 \quad as \quad n \to \infty. \end{aligned} \tag{2.2}$$

The metric in (2.2) turns out to be a natural one (see Section 5); that in (2.1) is typically stronger.

Proof. First note that for $x < X_{nn}$

$$\hat{e}_n(x) - e(x) = \frac{\overline{F}(x)}{\overline{\mathbb{F}}_n(x)} \left\{ \frac{-\int_x^\infty (\mathbb{F}_n - F)dI}{\overline{F}(x)} + \frac{e(x)}{\overline{F}(x)} (\mathbb{F}_n(x) - F(x)) \right\}.$$

Hence

$$\begin{aligned} \rho_{h(F)e/\overline{F}}(\hat{e}_n^*, e^*) &\leq \left\| \frac{\overline{F}}{\overline{\mathbb{F}}_n} \right\|_0^{b_n} \left\{ \sup_x \frac{|\int_x^\infty (\mathbb{F}_n - F)dI|}{h(F(x))e(x)} + \sup_x \frac{|\mathbb{F}_n(x) - F(x)|}{h(F(x))} \right\} \\ &\leq \left\| \frac{\overline{F}}{\overline{\mathbb{F}}_n} \right\|_0^{b_n} \cdot \rho_{h(F)}(\mathbb{F}_n, F)(M_1 + 1) \\ &\to_{a.s.} 0 \end{aligned}$$

using Condition 1a, Theorem 1 of Wellner (1977) to show $\rho_{h(F)}(\mathbb{F}_n, F) \to_{a.s.} 0$ a.s., and Theorem 2 of Wellner (1978) to show that $\limsup_n \|\overline{F}/\overline{\mathbb{F}}_n\|_0^{b_n} < \infty$ a.s..

Similarly, using Condition 1b,

$$\begin{aligned} \rho_{1/\overline{F}}(\hat{e}_n^*, e^*) &\leq \left\| \frac{\overline{F}}{\overline{\mathbb{F}}_n} \right\|_0^{b_n} \left\{ \sup_x \left| \int_x^\infty (\mathbb{F}_n - F)dI \right| + \sup_x e(x)|\mathbb{F}_n(x) - F(x)| \right\} \\ &\leq \left\| \frac{\overline{F}}{\overline{\mathbb{F}}_n} \right\|_0^{b_n} \cdot \rho_{h(F)}(\mathbb{F}_n, F) \left(\int_0^\infty h(F)dI + \|eh(F)\| \right) \\ &\to_{a.s.} 0. \end{aligned}$$

■

To extend Yang's weak convergence results, we will use the special uniform empirical processes $\mathbb{U}_n$ of the Appendix of Shorack (1972) or Shorack and Wellner (1986) which converge to a special Brownian bridge process $\mathbb{U}$ in the strong sense that

$$\rho_q(\mathbb{U}_n, \mathbb{U}) \to_p 0 \quad \text{as} \quad n \to \infty$$

for $q \in \mathcal{Q}(\downarrow)$, the set of all continuous functions on $[0,1]$ which are monotone decreasing on $[0,1]$ and $\int_0^1 q^{-2} dI < \infty$. Thus $\mathbb{U}_n = n^{1/2}(\Gamma_n - I)$ on $[0,1]$ where Γ_n is the empirical d.f. of special uniform $(0,1)$ random variables $\xi_1, \ldots, \xi_n$.

Define the *mean residual life process* on $\mathbb{R}^+$ by

$$\begin{aligned} n^{1/2}(\hat{e}_n(x) - e(x)) &= \frac{1}{\overline{\mathbb{F}}_n(x)} \left\{ - \int_x^\infty n^{1/2}(\mathbb{F}_n - F)dI + e(x)n^{1/2}(\mathbb{F}_n(x) - F(x)) \right\} \\ &\stackrel{d}{=} \frac{1}{\overline{\Gamma}_n(F(x))} \left\{ - \int_x^\infty \mathbb{U}_n(F)dI + e(x)\mathbb{U}_n(F(x)) \right\} \\ &\equiv \mathbb{Z}_n(x), \quad 0 \leq x < F^{-1}(\xi_{nn}) \end{aligned}$$

where $\xi_{nn} = \max_{1 \leq i \leq n} \xi_i$, and $\mathbb{Z}_n(x) \equiv -n^{1/2}e(x)$ for $x \geq F^{-1}(\xi_{nn})$. Thus $\mathbb{Z}_n$ has the same law as $n^{1/2}(\hat{e}_n - e)$ and is a function of the special process $\mathbb{U}_n$. Define the corresponding limiting process $\mathbb{Z}$ by

$$\text{(2.3)} \qquad \mathbb{Z}(x) = \frac{1}{\overline{F}(x)} \left\{ - \int_x^\infty \mathbb{U}(F)dI + e(x)\mathbb{U}(F(x)) \right\}, \qquad 0 \leq x < \infty.$$

If $\sigma^2 = Var(X) < \infty$ (and hence under either Condition 2a or 2b below), $\mathbb{Z}$ is a mean zero Gaussian process on $\mathbb{R}^+$ with covariance function described as follows:

Proposition 2.1. *Suppose that $\sigma^2 = Var(X) < \infty$. For $0 \le x \le y < \infty$*

$$Cov[\mathbb{Z}(x), \mathbb{Z}(y)] = \frac{\overline{F}(y)}{\overline{F}(x)} Var[\mathbb{Z}(y)] = \frac{\sigma^2(y)}{\overline{F}(y)} \tag{2.4}$$

where

$$\sigma^2(t) \equiv Var[X - t | X > t] = \frac{\int_t^\infty (x-t)^2 F(x)}{\overline{F}(t)} - e^2(t)$$

is the residual variance function; also

$$Cov[\mathbb{Z}(x)\overline{F}(x), \mathbb{Z}(y)\overline{F}(y)] = Var[\mathbb{Z}(y)\overline{F}(y)] = \overline{F}(y)\sigma^2(y). \tag{2.5}$$

Proof. It suffices to prove (2.5). Let $\mathbb{Z}' \equiv \mathbb{Z}\overline{F}$; from (2.3) we find

$$\begin{aligned} Cov[\mathbb{Z}'(x), \mathbb{Z}'(y)] &= e(x)e(y)F(x)\overline{F}(y) - e(x)\int_y^\infty F(x)\overline{F}(z)dz \\ &\quad - e(y)\int_x^\infty (F(y \wedge z) - F(y)F(z))dz \\ &\quad + \int_x^\infty \int_y^\infty (F(z \wedge w) - F(z)F(w))dzdw. \end{aligned}$$

Expressing integrals over (x, ∞) as the sum of integrals over (x, y) and (y, ∞), and recalling the defining formula for $e(y)$, we find that the right side reduces to

$$\begin{aligned} &\int_y^\infty \int_y^\infty (F(z \wedge z) - F(z)F(w)dzdw - e^2(y)F(y)\overline{F}(y) \\ &= \int_y^\infty (t-y)^2 dF(t) - \overline{F}(y)e^2(y) \\ &= \overline{F}(y)\sigma^2(y) \end{aligned}$$

which, being free of x, is also $Var[\mathbb{Z}'(y)]$. ∎

As in this proposition, the process $\mathbb{Z}$ is often more easily studied through the process $\mathbb{Z}' = \mathbb{Z}\overline{F}$; such a study continues in Section 5. Study of the variance of $\mathbb{Z}(x)$, namely $\sigma^2(x)/\overline{F}(x)$, for large x appears in Section 3.

Condition 2a. $\sigma^2 < \infty$ and there exists $q \in \mathcal{Q}(\downarrow)$ such that

$$M_2 \equiv M_2(q, F) \equiv \sup_x \frac{\int_x^\infty q(F)dI / q(F(x))}{e(x)} < \infty.$$

Since $0 < q(0) < \infty$ and $e(0) = E(X) < \infty$, Condition 2a implies that $\int_0^\infty q(F)dI < \infty$; Condition 2a may be rephrased as: $M_2 \equiv \|e_{q(F)/q(0)}/e_F\| < \infty$ where $e_{q(F)/q(0)}$ denotes the mean residual life function for the survival function $q(F)/q(0)$.

Condition 2b. $\sigma^2 < \infty$ and there exists $q \in \mathcal{Q}(\downarrow)$ such that $\int_0^\infty q(F)dI < \infty$.

Bounded e_F and existence of a moment of order greater than 2 is more than sufficient for 2b (see Section 3).

Theorem 2.2. *(Process convergence). Let $a_n \to 0$, $na_n \to \infty$. If Condition 2a holds for a particular $q \in \mathcal{Q}(\downarrow)$, then*

$$\rho_{q(F)e/\overline{F}}(\mathbb{Z}_n^*, \mathbb{Z}^*) \equiv \sup\left\{\frac{|\mathbb{Z}_n(x) - \mathbb{Z}(x)|\overline{F}(x)}{q(F(x))e(x)}: \ x \le b_n\right\} \to_p 0 \quad as \ \ n \to \infty. \tag{2.6}$$

If Condition 2b holds, then

$$\rho_{1/\overline{F}}(\mathbb{Z}_n^*, \mathbb{Z}^*) \equiv \sup\{|\mathbb{Z}_n(x) - \mathbb{Z}(x)|\overline{F}(x): \ x \le b_n\} \to_p 0 \quad as \ \ n \to \infty. \tag{2.7}$$

Proof. First write

$$\mathbb{Z}_n(x) - \mathbb{Z}(x) = \left\{\frac{\overline{F}(x)}{\overline{\Gamma}_n(F(x))} - 1\right\}\mathbb{Z}_n^1(x) + (\mathbb{Z}_n^1(x) - \mathbb{Z}(x))$$

where

$$\mathbb{Z}_n^1(x) \equiv \frac{1}{\overline{F}(x)}\left\{-\int_x^\infty \mathbb{U}_n(F)dI + e(x)\mathbb{U}_n(F(x))\right\}, \qquad 0 \le x < \infty.$$

Then note that, using Condition 2a,

$$\begin{aligned}\rho_{q(F)e/\overline{F}}(\mathbb{Z}_n^1, 0) &\le \sup_x \frac{|\int_x^\infty \mathbb{U}_n(F)dI}{q(F(x))e(x)} + \rho_q(\mathbb{U}_n, 0)\\ &\le \rho_q(\mathbb{U}_n, 0)\{M_2 + 1)\} = O_p(1);\end{aligned}$$

that $\|\overline{I}/\overline{\Gamma}_n - 1\|_0^{1-a_n} \to_p 0$ by Theorem 0 of Wellner (1978) since $na_n \to \infty$; and, again using Condition 2a, that

$$\begin{aligned}\rho_{q(F)e/\overline{F}}(\mathbb{Z}_n^1, \mathbb{Z}) &\le \sup_x \frac{|\int_x^\infty (\mathbb{U}_n(F) - \mathbb{U}(F))dI|}{q(F(x))e(x)} + \rho_q(\mathbb{U}_n, \mathbb{U})\\ &\le \rho_q(\mathbb{U}_n, \mathbb{U})\{M_2 + 1\} \to_p 0.\end{aligned}$$

Hence

$$\begin{aligned}\rho_{q(F)e/\overline{F}}(\mathbb{Z}_n^*, \mathbb{Z}^*) &\le \left\|\frac{\overline{I}}{\overline{\Gamma}_n} - 1\right\|_0^{1-a_n} \rho_{q(F)e/\overline{F}}(\mathbb{Z}_n^1, 0) + \rho_{q(F)e/\overline{F}}(\mathbb{Z}_n^1, \mathbb{Z})\\ &= o_p(1)O_p(1) + o_p(1) = o_p(1).\end{aligned}$$

Similarly, using Condition 2b

$$\begin{aligned}\rho_{1/\overline{F}}(\mathbb{Z}_n^1, 0) &\le \sup_x \left|\int_x^\infty \mathbb{U}_n(F)dI\right| + \sup_x e(x)|\mathbb{U}_n(F(x))|\\ &\le \rho_q(\mathbb{U}_n, 0)\left\{\int_0^\infty q(F)dI + \|eq(F)\|\right\} = O_p(1),\end{aligned}$$

$$\begin{aligned}\rho_{1/\overline{F}}(\mathbb{Z}_n^1, \mathbb{Z}) &\le \sup_x \left|\int_x^\infty (\mathbb{U}_n(F) - \mathbb{U}(F))dI\right| + \sup_x e(x)|\mathbb{U}_n(F(x)) - \mathbb{U}(F(x))|\\ &\le \rho_q(\mathbb{U}_n, \mathbb{U})\left\{\int_0^\infty q(F)dI + \|eq(F)\|\right\} \to_p 0,\end{aligned}$$

and hence

$$\begin{aligned}\rho_{1/\overline{F}}(\mathbb{Z}_n^*, \mathbb{Z}^*) &\le \left\|\frac{\overline{I}}{\overline{\Gamma}_n} - 1\right\|_0^{1-a_n} \rho_{1/\overline{F}}(\mathbb{Z}_n^1, 0) + \rho_{1/\overline{F}}(\mathbb{Z}_n^1, \mathbb{Z})\\ &= o_p(1)O_p(1) + o_p(1) = o_p(1).\end{aligned}$$

∎

3. Alternative sufficient conditions; $Var[\mathbb{Z}(x)]$ as $x \to \infty$.

Our goal here is to provide easily checked conditions which will imply the somewhat cumbersome Conditions 2a and 2b; similar conditions also appear in the work of Balkema and de Haan (1974), and we use their results to extend their formula for the residual coefficient of variation for large x ((3.1) below). This provides a simple description of the behavior of $Var[\mathbb{Z}(x)]$, the asymptotic variance of $n^{1/2}(\hat{e}_n(x) - e(x))$ as $x \to \infty$.

Condition 3. $E(X^r) < \infty$ for some $r > 2$.

Condition 4a. Condition 3 and $\lim_{x\to\infty} \frac{d}{dx}(1/\lambda(x)) = c < \infty$ where $\lambda = f/\overline{F}$, the hazard function.

Condition 4b. Condition 3 and $\limsup_{x\to\infty}\{\overline{F}(x)^{1+\gamma}/f(x)\} < \infty$ for some $r^{-1} < \gamma < 1/2$.

Proposition 3.1. *If Condition 4a holds, then* $0 \le c \le r^{-1}$*, Condition 2a holds, and the squared residual coefficient of variation tends to* $1/(1-2c)$*:*

$$\lim_{x\to\infty} \frac{\sigma^2(x)}{e^2(x)} = \frac{1}{1-2c}. \tag{3.1}$$

If Condition 4b holds, then Condition 2b holds.

Corollary 3.1. *Condition 4a implies*

$$Var[\mathbb{Z}(x)] \sim \frac{e^2(x)}{\overline{F}(x)}(1-2c)^{-1} \quad \text{as} \quad x \to \infty.$$

Proof. Assume 4a. Choose γ between r^{-1} and $1/2$; define a d.f. G on $\mathbb{R}^+$ by $\overline{G} = \overline{F}^\gamma$ and note that $g/\overline{G} = \gamma f/\overline{F} = \gamma\lambda$. By Condition 3 $x^r\overline{F}(x) \to 0$ as $x \to \infty$ and hence $x^{\gamma r}\overline{G}(x) \to 0$ as $x \to \infty$. Since $\gamma r > 1$, G has a finite mean and therefore $e_G(x) = \int_x^\infty \overline{G}dI/\overline{G}(x)$ is well-defined.

Set $\eta = 1/\lambda = \overline{F}/f$, and note that $\eta(x)\overline{G}(x) \to 0$ as $x \to \infty$. (If $\limsup \eta(x) < \infty$, then it holds trivially; otherwise, $\eta(x) \to \infty$ (because of 4a) and $\lim \eta(x)\overline{G}(x) = \lim(\eta(x)/x)(x\overline{G}(x)) = \lim \eta''(x)x\overline{G}(x) = 0$ by 4a and L'Hopital. Thus by L'Hopital's rule

$$\begin{aligned} 0 \quad &\le \quad \lim \frac{\eta(x)}{e_G(x)} = \lim \frac{\eta(x)\overline{G}(x)}{\int_x^\infty \overline{G}dI} \\ &= \quad \lim \frac{\eta(x)g(x) - \overline{G}(x)\eta'(x)}{\overline{G}(x)} \\ &= \quad \gamma - \lim \eta'(x) = \gamma - c \quad \text{by 4a.} \end{aligned}$$

Thus $c \le \gamma$ for any $\gamma > r^{-1}$ and it follows that $c \le r^{-1}$. It is elementary that $c \ge 0$ since $\eta = 1/\lambda$ is nonnegative.

Choose $q(t) = (1-t)^\gamma$. Then $\gamma - c > 0$, $q \in \mathcal{Q}(\downarrow)$, and to verify 2a it now suffices to show that $\lim(\eta(x)/e_F(x)) = 1 - c < \infty$ since it then follows that

$$\lim \frac{e_G(x)}{e_F(x)} = \lim \frac{\eta(x)/e_F(x)}{\eta(x)/e_G(x)} = \frac{1-c}{\gamma - c} < \infty.$$

By continuity and $e_G(0) < \infty$, $0 < e_F(0) < \infty$, this implies Condition 2a. But $r > 2$ implies that $x\overline{F}(x) \to 0$ as $x \to \infty$ so $\eta(x)\overline{F}(x) \to 0$ and hence

$$\lim \frac{\eta(x)}{e_F(x)} = \lim \frac{\eta\overline{F}(x)}{\int_x^\infty \overline{F} dI} = \lim(1 - \eta'(x)) = 1 - c.$$

That (3.1) holds will now follows from results of Balkema and de Haan (1974), as follows: Their Corollary to Theorem 7 implies that $P(\lambda(t)(X-t) > x|X > t) \to e^{-x}$ if $c = 0$ and $\to (1 + cx)^{-1/c}$ if $c > 0$. Thus, in the former case, F is in the domain of attraction of the Pareto residual life distribution and its related extreme value distribution. Then Theorem 8(a) implies convergence of the (conditional) mean and variance of $\lambda(t)(X-t)$ to the mean and variance of the limiting Pareto distribution, namely $(1-c)^{-1}$ and $(1-c)^{-2}(1-2c)^{-1}$. But the conditional mean of $\lambda(t)(X_t)$ is simply $\lambda(t)e(t)$, so that $\lambda(t) \sim (1-c)^{-1}/e(t)$ and (3.1) now follows.

If Condition 4b holds, let $q(F) = \overline{F}^\gamma$ again. Then $\int_0^\infty q(F)dI < \infty$, and it remains to show that $\limsup\{e(x)\overline{F}(x)^\gamma\} < \infty$. This follows from 4b by L'Hopital. ∎

Similarly, sufficient conditions for Conditions 1a and 1b can be given: simply replace "2" in Condition 3 and "1/2" in Condition 4b with "1", and the same proof works. Whether (3.1) holds when r in Condition 3 is exactly 2 is not known.

4. Examples.

The typical situation, when $e(x)$ has a finite limit and Condition 3 holds, is as follows: $e \sim \overline{F}/f \sim f/(-f')$ as $x \to \infty$ (by L'Hopital), and hence 4b, 2b, and 1b hold; also $\eta' \equiv (\overline{F}/f)' = [(F/f)(-f/f')] - 1 \to 0$ (4a with $c = 0$, and hence 2a and 1a hold), $\sigma(x) \sim e(x)$ from (3.1), and $Var[\mathbb{Z}] \sim e^2/\overline{F} \sim (\overline{F}/f)^2/\overline{F} \sim 1/(-f')$. We treat three examples, not all 'typical', in more detail.

Example 4.1. *(Exponential). Let $\overline{F}(x) = \exp(-x/\theta)$, $x \geq 0$, with $0 < \theta < \infty$. Then $e(x) = \theta$ for all $x \geq 0$. Conditions 4a and 4b hold (for all r, $\gamma \geq 0$) with $c = 0$, so Conditions 2a and 2b hold by Proposition 3.1 with $q(t) = (1-t)^{1/2-\delta}$, $0 < \delta < 1/2$. Conditions 1a and 1b hold with $h(t) = (1-t)^{1-\delta}$, $0 < \delta < 1$. Hence Theorems 2.1 and 2.2 hold where now*

$$\mathbb{Z}(x) = \frac{\mathbb{U}(F(x))}{1-(x)} - \frac{1}{1-F(x)} \int_{F(x)}^1 \frac{\mathbb{U}}{1-I} dI \stackrel{d}{=} \theta\mathbb{B}(e^{x/\theta}), \quad 0 \leq x < \infty$$

and $\mathbb{B}$ is standard Brownian motion on $[0, \infty)$. (The process $\mathbb{B}_1(t) = \mathbb{U}(1-t) - \int_{1-t}^1 (\mathbb{U}/(1-I))dI$, $0 \leq t \leq 1$, is Brownian motion on $[0, 1]$; and with $\mathbb{B}_2(x) \equiv x\mathbb{B}_1(1/x)$ for $1 \leq x \leq \infty$, $\mathbb{Z}(x) = \theta\mathbb{B}_2(1/\overline{F}(x)) = \theta\mathbb{B}_2(e^{x/\theta})$.) Thus, in agreement with (2.4),

$$Cov[\mathbb{Z}(x), \mathbb{Z}(y)] = \theta^2 e^{(x\wedge y)/\theta}, \quad 0 \leq x, y < \infty.$$

An immediate consequence is that $\|\mathbb{Z}_n^*\overline{F}\| \to_d \|\overline{F}\| \stackrel{d}{=} \theta \sup_{0\leq t\leq 1} |\mathbb{B}_1(t)|$; generalization of this to other F's appears in Section 5. (Because of the "memoryless" property of exponential F, the results for this example can undoubtedly be obtained by more elementary methods.)

Example 4.2. *(Weibull). Let* $\overline{F}(x) = \exp(-x^\theta)$, $x \ge 0$, *with* $0 < \theta < \infty$. *Conditions 4a and 4b hold (for all* $r, \gamma > 0$*) with* $c = 0$, *so Conditions 1 and 2 hold with* h *and* q *as in Example 1 by Proposition 3.1. Thus Theorems 2.1 and 2.2 hold. Also,* $e(x) \sim \theta^{-1}x^{1-\theta}$ *as* $x \to \infty$, *and hence* $Var[Z(x)] \sim \theta^{-2}x^{2(1-\theta)}\exp(x^\theta)$ *as* $x \to \infty$.

Example 4.3. *(Pareto). Let* $\overline{F}(x) = (1+cx)^{-1/c}$, $x \ge 0$, *with* $0 < c < 1/2$. *Then* $e(x) = (1-c)^{-1}(1+cx)$, *and Conditions 4a and 4b hold for* $r < c^{-1}$ *and* $\gamma \ge c$ *(and* c *of 4a is* c*). Thus Proposition 3.1 holds with* $r > 2$ *and* $c > 0$ *and* $Var[\mathbb{Z}(x)] \sim c^{2+(1/c)}(1-c)^{-2}(1-2c)^{-1}x^{2+(1/c)}$ *as* $x \to \infty$. *Conditions 1 and 2 hold with* h *and* q *as in Example 1, and Theorems 2.1 and 2.2 hold.*

If instead $1/2 \le c < 1$, *then* $E(X) < \infty$ *but* $E(X^2) = \infty$, *and 4a and 4b hold with* $1 < r < 1/c \le 2$ *and* $\gamma \ge c$. *Hence Condition 1 and Theorem 2.1 hold, but Condition 2 (and hence our proof of Theorem 2.2) fails. If* $c \ge 1$, *then* $E(X) = \infty$ *and* $e(x) = \infty$ *for all* $x \ge 0$.

Not surprisingly, the limiting process $\mathbb{Z}$ has a variance which grows quite rapidly, exponentially in the exponential and Weibull cases, and as a power (> 4) of x in the Pareto case.

5. Confidence bands for e.

We first consider the process $\mathbb{Z}' \equiv \mathbb{Z}\overline{F}$ on $\mathbb{R}^+$ which appeared in (2.5) of Proposition 2.1. Its sample analog $\mathbb{Z}'_n \equiv \mathbb{Z}_n\overline{\mathbb{F}}_n$ is easily seen to be a cumulative sum (times $n^{-1/2}$) of the observations exceeding x, each centered at $x + e(x)$; as x decreases the number of terms in the sum increases. Moreover, the corresponding increments apparently act asymptotically independently so that $\mathbb{Z}'_n$, in reverse time, is behaving as a cumulative sum of zero-mean independent increments. Adjustment for the non-linear variance should lead to Brownian motion. Let us return to the limit version $\mathbb{Z}'$.

The zero-mean Gaussian process $\mathbb{Z}'$ has covariance function $Cov[\mathbb{Z}'(x), \mathbb{Z}'(y)] = Var[\mathbb{Z}'(x \vee y)]$ (see (2.5)); hence, when viewed in reverse time, it has independent increments (and hence $\mathbb{Z}'$ is a reverse martingale). Specifically, with $\mathbb{Z}''(s) \equiv \mathbb{Z}'(-\log s)$, $\mathbb{Z}''$ is a zero-mean Gaussian process on $[0,1]$ with independent increments and $Var[\mathbb{Z}''(s)] = Var[\mathbb{Z}'(-\log s)] \equiv \tau^2(s)$. Hence τ^2 is increasing in s, and, from (2.5),

$$\tau^2(s) = \overline{F}(-\log s)\sigma^2(-\log s). \tag{5.1}$$

Now $\tau^2(1) = \sigma^2(0) = \sigma^2$, and

$$\tau^2(0) = \lim_{\epsilon\downarrow 0} \overline{F}(-\log\epsilon)\sigma^2(-\log\epsilon) = \lim_{x\to\infty} \overline{F}(x)\sigma^2(x) = 0$$

since

$$0 \le \overline{F}(x)\sigma^2(x) \le \overline{F}(x)E(X^2|X > x) = \int_x^\infty y^2 dF(y) \to 0.$$

Since $f(x) > 0$ for all $x \ge 0$, τ^2 is strictly increasing.

Let g be the inverse of τ^2; then $\tau^2(g(t)) = t$, $g(0) = 0$, and $g(\sigma^2) = 1$. Define $\mathbb{W}$ on $[0,1]$ by

$$\mathbb{W}(t) \equiv \sigma^{-1}\mathbb{Z}''(g(\sigma^2 t)) = \sigma^{-1}\mathbb{Z}'(-\log g(\sigma^2 t)). \tag{5.2}$$

Theorem 5.1. $\mathbb{W}$ *is standard Brownian motion on* $[0,1]$.

Proof. $\mathbb{W}$ is Gaussian with independent increments and $Var[\mathbb{W}(t)] = t$ by direct computation. ∎

Corollary 5.1. *If (2.7) holds, then*

$$\rho(\mathbb{Z}_n^{'*}, \mathbb{Z}^{'*}) \equiv \sup_{x \le b_n} |\mathbb{Z}_n(x)\overline{\mathbb{F}}_n(x) - \mathbb{Z}(x)\overline{F}(x)| = o_p(1)$$

and hence $\|\mathbb{Z}_n\overline{\mathbb{F}}_n\|_0^{b_n} \to_d \|\mathbb{Z}\overline{F}\| = \sigma\|\mathbb{W}\|_0^1$ *as* $n \to \infty$.

Proof. By Theorem 0 of Wellner (1978) $\|\overline{\mathbb{F}}_n/\overline{F} - 1\|_0^{b_n} \to_p 0$ as $n \to \infty$, and this together with (2.7) implies the first part of the statement. The second part follows immediately from the first and (5.2). ∎

Replacement of σ^2 by a consistent estimate S_n^2 (e.g. the sample variance based on all observations), and of b_n by $\hat{b}_n = \mathbb{F}_n^{-1}(1 - a_n)$, the $(n-m)$−th order statistic with $m = [na_n]$, leads to asymptotic confidence bands for $e = e_F$:

Corollary 5.2. *Let* $0 < a < \infty$, *and set* $\hat{d}_n(x) \equiv n^{-1/2} a S_n / \overline{\mathbb{F}_n(x)}$. *If (2.7) holds,* $S_n^2 \to_p \sigma^2$, *and* $na_n / \log\log n \uparrow \infty$, *then, as* $n \to \infty$

$$P\left(\hat{e}_n(x) - \hat{d}_n(x) \le e(x) \le \hat{e}_n(x) + \hat{d}_n(x) \quad \textit{for all} \quad 0 \le x \le \hat{b}_n\right) \to Q(a) \tag{5.3}$$

where

$$\begin{aligned} Q(a) &\equiv P(\|\mathbb{W}\|_0^1 < a) = \sum_{k=-\infty}^{\infty} (-1)^k \{\Phi((2k+1)a) - \Phi((2k-1)a)\} \\ &= 1 - 4\{\overline{\Phi}(a) - \overline{\Phi}(3a) + \overline{\Phi}(5a) - \cdots\} \end{aligned}$$

and Φ *denotes the standard normal d.f.*

Proof. It follows immediately from Corollary 5.1 and $S_n \to_p \sigma > 0$ that

$$\|\mathbb{Z}_n\overline{\mathbb{F}}_n\|_0^{b_n}/S_n \to_d \|\mathbb{Z}\overline{F}\|/\sigma = \|\mathbb{W}\|_0^1.$$

Finally b_n may be replaced by $\hat{b}_n$ without harm: letting $c_n = 2\log\log/(na_n) \to 0$ and using Theorem 4S of Wellner (1978), for $\tau > 1$ and all $n \ge N(\omega, \tau)$, $\hat{b}_n \equiv \mathbb{F}_n^{-1}(1 - a_n) \stackrel{d}{=} F^{-1}(\Gamma_n^{-1}(1 - a_n)) \le F^{-1}(\{1 + \tau c_n^{1/2}\}(1 - a_n))$ w.p. 1. This proves the convergence claimed in the corollary; the expression for $Q(a)$ is well-known (e.g. see Billingsley (1968), page 79). ∎

The approximation $1 - 4\overline{\Phi}(a)$ for $Q(a)$ gives 3-place accuracy for $a > 1.4$. A short table appears below:

TABLE 1
$Q(a)$ for selected a

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

Thus, choosing a so that $Q(a) = \beta$, (5.3) provides a two-sided simultaneous confidence band for the function e with confidence coefficient asymptotically β.

In applications we suggest taking $a_n = n^{-1/2}$ so that $\hat{b}_n$ is the $(n-m)-$th order statistic with $m = [n^{1/2}]$; we also want m large enough for an adequate central limit effect, remembering that the conditional life distribution may be quite skewed. (In a similar fashion, one-sided asymptotic bands are possible, but they will be less trustworthy because of skewness.)

Instead of simultaneous bands for all real x one may seek (tighter) bands on $e(x)$ for one or two specific $x-$values. For this we can apply Theorem 2.2 and Proposition 2.1 directly. We first require a consistent estimator of the asymptotic variance of $n^{1/2}(\hat{e}_n(x) - e(x))$, namely $\sigma^2(x)/\overline{F}(x)$.

Proposition 5.1. *Let $0 \le x < \infty$ be fixed and let $S_n^2(x)$ be the sample variance of those observations exceeding x. If Condition 3 holds then $S_n^2(x)/\overline{\mathbb{F}}_n(x) \to_{a.s.} \sigma^2(x)/\overline{F}(x)$.*

Proof. Since $\overline{\mathbb{F}}_n(x) \to_{a.s.} \overline{F}(x) > 0$ and

$$S_n^2(x) = \frac{2\int_x^\infty (y-x)\overline{\mathbb{F}}_n(y)dy}{\overline{\mathbb{F}}_n(x)} - \hat{e}_n^2(x),$$

it suffices to show that $\int_x^\infty y\mathbb{F}_n(y)dy \to_{a.s.} \int_x^\infty y\overline{F}(y)dy$. Let $h(t) = (1-t)^{\gamma+1/2}$ and $q(t) = (1-t)^\gamma$ with $r^{-1} < \gamma < 1/2$ so that $h \in \mathcal{H}(\downarrow)$, $q \in \mathcal{Q}(\downarrow)$, and $\int_0^\infty q(F)dI < \infty$ by the proof of Proposition 3.1. Then,

$$\left| \int_x^\infty y\overline{\mathbb{F}}_n(y)dy - \int_x^\infty y\overline{F}(y)dy \right| \le \rho_{h(F)}(\mathbb{F}_n, F)\int_0^\infty Ih(F)dI \to_{a.s.} 0$$

by Theorem 1 of Wellner (1977) since

$$\int_0^\infty Ih(F)dI = \int_0^\infty (I^2\overline{F})^{1/2}q(F)dI < \infty.$$

■

By Theorem 2.2, Propositions 2.1 and 5.1, and Slutsky's theorem we have:

Proposition 5.2. *Under the conditions of Proposition 5.1,*

$$d_n(x) \equiv n^{1/2}(\hat{e}_n(x) - e(x))\overline{\mathbb{F}}_n^{1/2}(x)/S_n(x) \to_d N(0,1) \quad as \quad n \to \infty.$$

This makes feasible an asymptotic confidence interval for $e(x)$ (at this particular fixed x). Similarly, for $x < y$, using the joint asymptotic normality of $(d_n(x), d_n(y))$ with asymptotic correlation $\{\overline{F}(y)\sigma^2(y)/\overline{F}(x)\sigma^2(x)\}^{1/2}$ estimated by

$$\{\overline{\mathbb{F}}_n(y)S_n^2(y)/\overline{\mathbb{F}}_n(x)S_n^2(x)\}^{1/2},$$

an asymptotic confidence ellipse for $(e(x), e(y))$ may be obtained.

6. Illustration of the confidence bands.

We illustrate with two data sets presented by Bjerkedal (1960) and briefly mention one appearing in Barlow and Campo (1975).

Bjerkedal gave various doses of tubercle bacilli to groups of 72 guinea pigs and recorded their survival times. We concentrate on Regimens 4.3 and 6.6 (and briefly mention 5.5, the only other complete data set in Bjerkedal's study M); see Figures 1 and 2 below.

First consider the estimated mean residual life $\hat{e}_n$, the center jagged line in each figure. Figure 1 has been terminated at day 200; the plot would continue approximately horizontally, but application of asymptotic theory to this part of $\hat{e}_n$, based only the last 23 survival times (the last at 555 days), seems unwise. Figure 2 has likewise been terminated at 200 days, omitting only nine survival times (the last at 376 days); the graph of $\hat{e}_n$ would continue downward. The dashed diagonal line is $\overline{X}-x$; if all survival times were equal, say μ, then the residual life function would be $(\mu - x)^+$, a lower bound on $e(x)$ near the origin. More specifically, a Maclaurin expansion yields

$$e(x) = \mu + (\mu f_0 - 1)x + (1/2)\{(2\mu f_0 - 1)f_0 + f_0'\}x^2 + o(x^2)$$

where $f_0 = f(0)$, $f_0' = f'(0)$, if f' is continuous at 0, or

$$e(x) = \mu - x + \frac{\mu d}{r!}x^r + o(x^r)$$

if $f^{(s)}(0) = 0$ for $s < (r-1)$ (≥ 0) and $= d$ for $s = r-1$ (if $f^{(r-1)}$ is continuous at 0). It thus seems likely from Figures 1 and 2 that in each of these cases either $f_0 = 0$ and $f_0' > 0$ or f_0 is near 0 (and $f_0' \geq 0$).

Also, for large x, $e(x) \sim 1/\lambda(x)$, and Figure 1 suggests that the corresponding λ and e have finite positive limits, whereas the e of Figure 2 may eventually decrease (λ increase). We know of no parametric F that would exhibit behavior quite like these.

The upper and lower jagged lines in the figures provide 90% (asymptotic) confidence bands for the respective e's, based on (5.3). At least for Regimen 4.3, a constant e (exponential survival) can be rejected.

The vertical bars at $x = 0$, $x = 100$, and $x = 200$ in Figure 1, and at 0, 50, and 100 in Figure 2, are 90% (asymptotic) pointwise confidence intervals on e at the corresponding $x-$values (based on Proposition 5.2). Notice that these intervals are not much narrower than the simultaneous bands early in the survival data, but are substantially narrower later on.

A similar graph for Regimen 5.5 (not presented) is somewhat similar to that in Figure 2, with the upward turn in $\hat{e}_n$ occurring at 80 days instead of at 50, and a possible downward turn at somewhere around 250 days (the final death occurring at 598 days).

A similar graph was prepared for the failure data on 107 right rear tractor brakes presented by Barlow and Campo (1975), page 462. It suggests a quadratic decreasing e for the first 1500 to 2000 hours (with $f(0)$ at or near 0 but $f'(0)$ definitely positive), with $\overline{X} = 2024$, and with a possibly constant of slightly increasing e from 1500 or so to 6000 hours. The e for a gamma distribution with $\lambda = 2$ and $\alpha = .001$ ($e(x) = \alpha^{-1}(\alpha x + 2)/(\alpha x + 1)$ with $\alpha = .001$) fits reasonably well – i.e. is within the confidence bands, even for 25% confidence. Note that this is in excellent agreement with Figures 2.1(b) and 3.1(d) of Barlow and Campo (1975). Bryson and Siddiqui's (1969) data set was too small ($n = 43$) for these asymptotic methods, except possibly early in the data set.)

7. Further developments

The original version of this paper, Hall and Wellner (1979), ended with a one-sentence sketch of two remaining problems: "Confidence bands on the difference

FIG 1. 90% *confidence bands for mean residual life; Regimen 4.3*

Fig 2. 90% *confidence bands for mean residual life; Regimen 6.6* between two mean residual life functions, and for the case of censored data, will be presented in subsequent papers." Although we never did address these questions ourselves, others took up these further problems.

Our aim in this final section is to briefly survey some of the developments since 1979 concerning mean residual life, including related studies of median residual life and other quantiles, as well as developments for censored data, alternative inference strategies, semiparametric models involving mean or median residual life, and generalizations to higher dimensions. For a review of further work up to 1988 see Guess and Proschan (1988).

7.1. Confidence bands and inference

CsörgHo et al. (1986) gave a further detailed study of the asymptotic behavior of the mean residual life process as well as other related processes including the Lorenz curve. Berger et al. (1988) developed tests and confidence sets for comparing two mean residual life functions based on independent samples from the respective populations. These authors also gave a brief treatment based on comparison of *median residual life*, to be discussed in Subsection 7.3 below. Csörgo and Zitikis (1996) introduced weighted metrics into the study of the asymptotic behavior of the mean residual life process, thereby avoiding the intervals $[0, x_n]$ changing with n involved in our Theorems 2.1 and 2.2, and thereby provided confidence bands and intervals for e_F in the right tail. Zhao and Qin (2006) introduced empirical likelihood methods to the study of the mean residual life function. They obtained confidence intervals and confidence bands for compact sets $[0, \tau]$ with $\tau < \tau_F \equiv \inf\{x\colon F(x) = 1\}$.

7.2. Censored data

Yang (1977/78) initiated the study of estimated mean residual life under random right censorship. She used an estimator $\hat{F}_n$ which is asymptotically equivalent to the Kaplan - Meier estimator and considered, in particular, the case when X is bounded and stochastically smaller than the censoring variable C. In this case she proved that $\sqrt{n}(\hat{e}(x)-e(x))$ converges weakly (as $n \to \infty$) to a Gaussian process with mean zero. Csörgo and Zitikis (1996) give a brief review of the challenges involved in this problem; see their page 1726. Qin and Zhao (2007) extended their earlier study (Zhao and Qin (2006)) of empirical likelihood methods to this case, at least for the problem of obtaining pointwise confidence intervals. The empirical likelihood methods seem to have superior coverage probability properties in comparison to the Wald type intervals which follow from our Proposition 5.2. Chaubey and Sen (1999) introduced smooth estimates of mean residual life in the uncensored case. In Chaubey and Sen (2008) they introduce and study smooth estimators of e_F based on corresponding smooth estimators of $\overline{F} = 1 - F$ introduced by Chaubey and Sen (1998).

7.3. Median and quantile residual life functions

Because mean residual life is frequently difficult, if not impossible, to estimate in the presence of right-censoring, it is natural to consider surrogates for it which do not depend on the entire right tail of F. Natural replacements include median residual life and corresponding *residual life quantiles*. The study of median residual life was apparently initiated in Schmittlein and Morrison (1981). Characterization issues and basic properties have been investigated by Gupta and Langford (1984), Joe and Proschan (1984b), and Lillo (2005). Joe and Proschan (1984a) proposed comparisons of two populations based on their corresponding median (and other quantile) residual life functions. As noted by Joe and Proschan, "Some results differ notably from corresponding results for the mean residual life function". Jeong et al. (2008) investigated estimation of median residual life with right-censored data for one-sample and two-sample problems. They provided an interesting illustration of their methods using a long-term follow-up study (the National Surgical Adjuvant Breast and Bowel Project, NSABP) involving breast cancer patients.

7.4. Semiparametric models for mean and median residual life

Oakes and Dasu (1990) investigated a characterization related to a *proportional mean residual life* model: $e_G = \psi e_F$ with $\psi > 0$. Maguluri and Zhang (1994) studied several methods of estimation in a semiparametric regression version of the proportional mean residual life model, $e(x|z) = \exp(\theta^T z) e_0(x)$ where $e(x|z)$ denotes the conditional mean residual life function given $Z = z$. Chen et al. (2005) provide a nice review of various models and study estimation in the same semiparametric proportional mean residual life regression model considered by Maguluri and Zhang (1994), but in the presence of right censoring. Their proposed estimation method involves inverse probability of censoring weighted (IPCW) estimation methods (Horvitz and Thompson (1952); Robins and Rotnitzky (1992)). Chen and Cheng (2005) use counting process methods to develop alternative estimators for the proportional mean residual life model in the presence of right censoring. The methods of estimation considered by Maguluri and Zhang (1994), Chen et al. (2005), and Chen and Cheng (2005) are apparently inefficient. Oakes and Dasu (2003) consider information calculations and likelihood based estimation in a two-sample version of the proportional mean residual life model. Their calculations suggest that certain weighted ratio-type estimators may achieve asymptotic efficiency, but a definitive answer to the issue of efficient estimation apparently remains unresolved. Chen and Cheng (2006) proposed an alternative additive semiparametric regression model involving mean residual life. Ma and Yin (2010) considered a large family of semiparametric regression models which includes both the additive model proposed by Chen and Cheng (2006) and the proportional mean residual life model considered by earlier authors, but advocated replacing mean residual life by median residual life. Gelfand and Kottas (2003) also developed a median residual life regression model with additive structure and took a semiparametric Bayesian approach to inference.

7.5. Monotone and Ordered mean residual life functions

Kochar et al. (2000) consider estimation of e_F subject to the shape restrictions that e_F is increasing or decreasing. The main results concern ad-hoc estimators that are simple monotizations of the basic nonparametric empirical estimators $\hat{e}_n$ studied here. These authors show that the nonparametric maximum likelihood estimator does not exist in the increasing MRL case and that although the nonparametric MLE exists in the decreasing MRL case, the estimator is difficult to compute. Ebrahimi (1993) and Hu et al. (2002) study estimation of two mean residual life functions e_F and e_G in one- and two-sample settings subject to the restriction $e_F(x) \leq e_G(x)$ for all x. Hu et al. (2002) also develop large sample confidence bands and intervals to accompany their estimators.

7.6. Bivariate residual life

Jupp and Mardia (1982) defined a multivariate mean residual life function and showed that it uniquely determines the joint multivariate distribution, extending the known univariate result of Cox (1962); see Hall and Wellner (1981) for a review of univariate results of this type. See Ma (1996, 1998) for further multivariate characterization results. Kulkarni and Rattihalli (2002) introduced a bivariate mean residual life function and propose natural estimators.

References

Balkema, A. A. and de Haan, L. (1974). Residual life time at great age. *Ann. Probability* **2** 792–804.

Barlow, R. E. and Campo, R. (1975). Total time on test processes and applications to failure data analysis. In *Reliability and fault tree analysis (Conf., Univ. California, Berkeley, Calif., 1974)*. Soc. Indust. Appl. Math., Philadelphia, Pa., 451–481.

Berger, R. L., Boos, D. D. and Guess, F. M. (1988). Tests and confidence sets for comparing two mean residual life functions. *Biometrics* **44** 103–115.

Billingsley, P. (1968). *Convergence of probability measures*. John Wiley & Sons Inc., New York.

Bjerkedal, T. (1960). Acquisition of resistance in guinea pigs infected with different doses of virulent tubercle bacilli. *Amer. Jour. Hygiene* **72** 130–148.

Bryson, M. C. and Siddiqui, M. M. (1969). Some criteria for aging. *J. Amer. Statist. Assoc.* **64** 1472–1483.

Chaubey, Y. P. and Sen, A. (2008). Smooth estimation of mean residual life under random censoring. In *Beyond parametrics in interdisciplinary research: Festschrift in honor of Professor Pranab K. Sen*, vol. 1 of *Inst. Math. Stat. Collect.* Inst. Math. Statist., Beachwood, OH, 35–49.

Chaubey, Y. P. and Sen, P. K. (1998). On smooth estimation of hazard and cumulative hazard functions. In *Frontiers in probability and statistics (Calcutta, 1994/1995)*. Narosa, New Delhi, 91–99.

Chaubey, Y. P. and Sen, P. K. (1999). On smooth estimation of mean residual life. *J. Statist. Plann. Inference* **75** 223–236. The Seventh Eugene Lukacs Conference (Bowling Green, OH, 1997).

Chen, Y. Q. and Cheng, S. (2005). Semiparametric regression analysis of mean residual life with censored survival data. *Biometrika* **92** 19–29.

Chen, Y. Q. and Cheng, S. (2006). Linear life expectancy regression with censored data. *Biometrika* **93** 303–313.

Chen, Y. Q., Jewell, N. P., Lei, X. and Cheng, S. C. (2005). Semiparametric estimation of proportional mean residual life model in presence of censoring. *Biometrics* **61** 170–178.

Chiang, C. L. (1960). A stochastic study of the life table and its applications: I. probability distributions of the biometric functions. *Biometrics* **16** 618–635.

Chiang, C. L. (1968). *Introduction to Stochastic Processes in Biostatistics*. Wiley, New York.

Cox, D. R. (1962). *Renewal theory*. Methuen & Co. Ltd., London.

CsörgHo, M., Csörgo, S. and Horváth, L. (1986). *An asymptotic theory for empirical reliability and concentration processes*, vol. 33 of *Lecture Notes in Statistics*. Springer-Verlag, Berlin.

Csörgo, M. and Zitikis, R. (1996). Mean residual life processes. *Ann. Statist.* **24** 1717–1739.

Ebrahimi, N. (1993). Estimation of two ordered mean residual lifetime functions. *Biometrics* **49** 409–417.

El Barmi, H. and Mukerjee, H. (2004). Consistent estimation of distributions with type II bias with applications in competing risks problems. *Ann. Statist.* **32** 245–267.

Gelfand, A. E. and Kottas, A. (2003). Bayesian semiparametric regression for median residual life. *Scand. J. Statist.* **30** 651–665.

Gross, A. J. and Clark, V. A. (1975). *Survival distributions: reliability applications in the biomedical sciences*. Wiley, New York.

Guess, F. and Proschan, F. (1988). Mean residual life: theory and application. In *Handbook of Statistics: Quality Control and Reliability*, vol. 7. North-Holland, Amsterdam, 215–224.

Gupta, R. C. and Langford, E. S. (1984). On the determination of a distribution by its median residual life function: a functional equation. *J. Appl. Probab.* **21** 120–128.

Hall, W. J. and Wellner, J. (1981). Mean residual life. In *Statistics and related topics (Ottawa, Ont., 1980)*. North-Holland, Amsterdam, 169–184.

Hall, W. J. and Wellner, J. A. (1979). Estimation of mean residual life. Tech. rep., University of Rochester.

Hollander, M. and Proschan, F. (1975). Tests for the mean residual life. *Biometrika* **62** 585–593.

Horvitz, D. G. and Thompson, D. J. (1952). A generalization of sampling without replacement from a finite universe. *J. Amer. Statist. Assoc.* **47** 663–685.

Hu, X., Kochar, S. C., Mukerjee, H. and Samaniego, F. J. (2002). Estimation of two ordered mean residual life functions. *J. Statist. Plann. Inference* **107** 321–341. Statistical inference under inequality constraints.

Jeong, J.-H., Jung, S.-H. and Costantino, J. P. (2008). Nonparametric inference on median residual life function. *Biometrics* **64** 157–163, 323–324.

Joe, H. and Proschan, F. (1984a). Comparison of two life distributions on the basis of their percentile residual life functions. *Canad. J. Statist.* **12** 91–97.

Joe, H. and Proschan, F. (1984b). Percentile residual life functions. *Oper. Res.* **32** 668–678.

Jupp, P. E. and Mardia, K. V. (1982). A characterization of the multivariate Pareto distribution. *Ann. Statist.* **10** 1021–1024.

Kochar, S. C., Mukerjee, H. and Samaniego, F. J. (2000). Estimation of a monotone mean residual life. *Ann. Statist.* **28** 905–921.

Kulkarni, H. V. and Rattihalli, R. N. (2002). Nonparametric estimation of a bivariate mean residual life function. *J. Amer. Statist. Assoc.* **97** 907–917.

Lillo, R. E. (2005). On the median residual lifetime and its aging properties: a characterization theorem and applications. *Naval Res. Logist.* **52** 370–380.

Ma, C. (1996). Multivariate survival functions characterized by constant product of mean remaining lives and hazard rates. *Metrika* **44** 71–83.

Ma, C. (1998). Characteristic properties of multivariate survival functions in terms of residual life distributions. *Metrika* **47** 227–240.

Ma, Y. and Yin, G. (2010). Semiparametric median residual life model and inference. *Canad. J. Statist.* **38** 665–679.

MAGULURI, G. and ZHANG, C.-H. (1994). Estimation in the mean residual life regression model. *J. Roy. Statist. Soc. Ser. B* **56** 477–489.

OAKES, D. and DASU, T. (1990). A note on residual life. *Biometrika* **77** 409–410.

OAKES, D. and DASU, T. (2003). Inference for the proportional mean residual life model. In *Crossing boundaries: statistical essays in honor of Jack Hall*, vol. 43 of *IMS Lecture Notes Monogr. Ser.* Inst. Math. Statist., Beachwood, OH, 105–116.

QIN, G. and ZHAO, Y. (2007). Empirical likelihood inference for the mean residual life under random censorship. *Statist. Probab. Lett.* **77** 549–557.

ROBINS, J. M. and ROTNITZKY, A. (1992). Recovery of information and adjustment for dependent censoring using surrogate markers. In *AIDS Epidemiology, Methodological Issues*. Birkhauser, Boston, 297–331.

SCHMITTLEIN, D. C. and MORRISON, D. G. (1981). The median residual lifetime: a characterization theorem and an application. *Oper. Res.* **29** 392–399.

SHORACK, G. R. (1972). Functions of order statistics. *Ann. Math. Statist.* **43** 412–427.

SHORACK, G. R. and WELLNER, J. A. (1986). *Empirical Processes with Applications to Statistics*. Wiley Series in Probability and Mathematical Statistics: Probability and Mathematical Statistics, John Wiley & Sons Inc., New York.

SOLARI, A., SALMASO, L., EL BARMI, H. and PESARIN, F. (2008). Conditional tests in a competing risks model. *Lifetime Data Anal.* **14** 154–166.

WELLNER, J. A. (1977). A Glivenko-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.

WILSON, E. B. (1938). The standard deviation of sampling for life expectancy. *J. Amer. Statist. Assoc.* **33** 705–708.

YANG, G. (1977/78). Life expectancy under random censorship. *Stochastic Processes Appl.* **6** 33–39.

YANG, G. L. (1978). Estimation of a biometric function. *Ann. Statist.* **6** 112–116.

ZHAO, Y. and QIN, G. (2006). Inference for the mean residual life function via empirical likelihood. *Comm. Statist. Theory Methods* **35** 1025–1036.