Copy for Wellner

ESTIMATION OF MEAN RESIDUAL LIFE

W. J. Hall and Jon A. Wellner
University of Rochester

January 1979

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 real line, and the variance of 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.

AMS 1970 subject classifications. Primary 62P05, 62N05, 62G15; Secondary 62G05, 62E20.

Key words and phrases: Mean residual life function, life expectancy, consistency, limiting Gaussian process, confidence bands.

1. Introduction and summary. Let $X_1,\ldots,X_n$ be a random sample from a continuous df F on $\mathbb{R}^+ = [0,\infty)$ with finite mean $\mu = E(X)$, variance $\sigma^2 \leq \infty$, and density $f(x) > 0$. Let $\bar{F} = 1 - F$ denote the survival function, let $\mathbb{F}_n$ and $\bar{\mathbb{F}}_n$ denote the empirical distribution function and empirical survival function respectively, and let

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

denote the mean residual life function or life expectancy function at age x. We use a subscript F or $\bar{F}$ 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

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

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 1 by using the techniques of Wellner (1977,1978); then her weak convergence result is extended in Theorem 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 size $n(x)$ from the conditional distribution $P(\cdot|X>x)$. Noting that $\bar{\mathbb{F}}_n(x) = n(x)/n \rightarrow \bar{F}(x)$

a.s., we would then have $n\text{Var}[\hat{e}_n(x)] = n\,\sigma^2(x)/n(x) \to \sigma^2(x)/\overline{F}(x)$. Proposition 1 and Theorem 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 1 and 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 3); this enables construction of nonparametric simultaneous confidence bands for the function e_F (Corollary 3). 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 $k = n\overline{F}_n(x)$ 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 4 below). Chiang (1968, page 214) gives the analogous marginal result for grouped data in more detail, but again without proofs; note the column $S_{\hat{e}_i}$ 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 for 43 patients suffering from chronic granulocytic 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 descreasing 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 df 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 zero 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 $H(\downarrow)$ 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 $h \in H(\downarrow)$ 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 $h(F)$ is a survival function on $\mathbb{R}^+$ and that the numerator in Condition 1a is simply $e_{h(F)}$; hence Condition 1a may be rephrased as: there exists $h \in H(\downarrow)$ such that $M_1 \equiv ||e_{h(F)}/e_F|| < \infty$.

Condition 1b. There exists $h \in H(\downarrow)$ 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 1b (see Section 3).

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

$$(2.1) \qquad \rho_{h(F)e/\overline{F}}(\hat{e}_n^*, e^*) \equiv \sup\{\frac{|\hat{e}_n(x) - e(x)|\overline{F}(x)}{h(F(x))e(x)}: \; x \le b_n\} \to_{a.s.} 0 \text{ as } n \to \infty.$$

If Condition 1b holds, then

$$(2.2) \qquad \rho_{1/\overline{F}}(\hat{e}_n^*, e^*) \equiv \sup\{|\hat{e}_n(x) - e(x)|\overline{F}(x): \; x \le b_n\} \to_{a.s.} 0 \text{ as } n \to \infty.$$

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 $x < X_{nn}$

$$\hat{e}_n(x) - e(x) = (\overline{F}(x)/\overline{\mathbb{F}}_n(x))\{-(\overline{F}(x))^{-1}\int_x^\infty (\mathbb{F}_n - F)dI + (e(x)/\overline{F}(x))(\mathbb{F}_n(x) - F(x))\}.$$

Hence

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

using Condition 1a, Theorem 1 of Wellner (1977) to show $\rho_{h(F)}(\mathbb{F}_n, F) \to 0$ 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,

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

To extend Yang's weak convergence results, we will use the special uniform empirical processes U_n of the Appendix of Shorack (1972) which converge to a special Brownian bridge process U in the strong sense that

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

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 $U_n = n^{\frac{1}{2}}(\Gamma_n - I)$ on $[0,1]$ where Γ_n is the empirical df of special uniform $(0,1)$ rv's $\xi_1, \ldots, \xi_n$.

Define the <u>mean residual life process</u> on $\mathbb{R}^+$ by

$$\begin{aligned} n^{\frac{1}{2}}(\hat{e}_n(x) - e(x)) &= (\overline{\mathbb{F}}_n(x))^{-1}\{-\textstyle\int_x^\infty n^{\frac{1}{2}}(\mathbb{F}_n - F)dI + e(x)n^{\frac{1}{2}}(\mathbb{F}_n(x) - F(x))\} \\ &=_d (\overline{\Gamma}_n(F(x)))^{-1}\{-\textstyle\int_x^\infty U_n(F)dI + e(x)U_n(F(x))\} \\ &\equiv Z_n(x) \qquad\qquad 0 \le x < F^{-1}(\xi_{nn}) \end{aligned}$$

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

$$Z(x) = (\overline{F}(x))^{-1}\{-\textstyle\int_x^\infty U(F)dI + e(x)U(F(x))\} \qquad 0 \le x < \infty. \tag{2.3}$$

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

<u>Proposition 1</u>. Suppose that $\sigma^2 = \text{Var}(X) < \infty$. For $0 \le x \le y < \infty$,

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

where

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

is the residual variance function; also

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

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

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

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

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

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

As in this proposition, the process Z is often more easily studied through the process $Z' = Z\overline{F}$; such a study continues in Section 5. Study of the variance of $Z(x)$, 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: $M_2 \equiv ||e_{q(F)}/e_F|| < \infty$ where $e_{q(F)}$ denotes the mean residual function for the survival function $q(F)$.

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

$$(2.6) \qquad \rho_{q(F)e/\overline{F}}(Z_n^*, Z^*) \equiv \sup\{\frac{|Z_n(x) - Z(x)|\overline{F}(x)}{q(F(x))e(x)} : x \le b_n\} \to_p 0 \text{ as } n \to \infty.$$

If Condition 2b holds, then

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

Proof. First write

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

where

$$Z_n^1(x) \equiv (\overline{F}(x))^{-1}\{-\int_x^\infty U_n(F)\,dI + e(x)U_n(F(x))\}, \qquad 0 \le x < \infty.$$

Then note that, using Condition 2a,

$$\rho_{q(F)e/\overline{F}}(Z_n^1, 0) \le \sup_x \frac{|\int_x^\infty U_n(F)\,dI|}{q(F(x))e(x)} + \rho_q(U_n, 0)$$

$$\le \rho_q(U_n, 0)\{M_2 + 1\} = 0_p(1);$$

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

$$\rho_{q(F)e/\overline{F}}(Z_n^1, Z) \le \sup_x \frac{|\int_x^\infty (U_n(F) - U(F))\,dI|}{q(F(x))e(x)} + \rho_q(U_n, U)$$

$$\le \rho_q(U_n, U)\{M_2 + 1\} \to_p 0.$$

Hence

$$\rho_{q(F)e/\overline{F}}(Z_n^*,Z^*) \le \left|\left|\frac{\overline{I}}{\overline{\Gamma}_n} - 1\right|\right|_0^{1-a_n} \rho_{q(F)e/\overline{F}}(Z_n^1,0) + \rho_{q(F)e/\overline{F}}(Z_n^1,Z)$$

$$= o_p(1)O_p(1) + o_p(1) = o_p(1).$$

Similarly, using Condition 2b,

$$\rho_{1/\overline{F}}(Z_n^1,0) \le \sup_x\left|\int_x^\infty U_n(F)dI\right| + \sup_x e(x)\left|U_n(F(x))\right|$$

$$\le \rho_q(U_n,0)\{\int_0^\infty q(F)dI + ||eq(F)||\} = O_p(1),$$

$$\rho_{1/\overline{F}}(Z_n^1,Z) \le \sup_x\left|\int_x^\infty (U_n(F) - U(F))dI\right| + \sup_x e(x)\left|U_n(F(x)) - U(F(x))\right|$$

$$\le \rho_q(U_n,U)\{\int_0^\infty q(F)dI + ||eq(F)||\} \to_p 0,$$

and hence

$$\rho_{1/\overline{F}}(Z_n^*,Z^*) \le \left|\left|\frac{\overline{I}}{\overline{\Gamma}_n} - 1\right|\right|_0^{1-a_n} \rho_{1/\overline{F}}(Z_n^1,0) + \rho_{1/\overline{F}}(Z_n^1,Z)$$

$$= o_p(1)O_p(1) + o_p(1) = o_p(1). \quad \square$$

3. <u>Alternative sufficient conditions; $\mathrm{Var}[Z(x)]$ as $x \to \infty$</u>. 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 $\mathrm{Var}[Z(x)]$, the asymptotic variance of $n^{\frac{1}{2}}(\hat{e}_n(x) - e(x))$, 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 $r^{-1} < \gamma < \frac{1}{2}$.

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

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

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

Corollary 1. Condition 4a implies

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

Proof. Assume 4a. Choose γ between r^{-1} and $\frac{1}{2}$; define a df 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

$$0 \le \lim \frac{\eta(x)}{e_G(x)} = \lim \frac{\eta(x)\overline{G}(x)}{\int_x^\infty \overline{G}dI}$$

$$= \lim(\eta(x)g(x) - \overline{G}(x)\eta'(x))/\overline{G}(x)$$

$$= \gamma - \lim \eta'(x) = \gamma - c \text{ by 4a.}$$

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

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

That (3.1) holds will now follow 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 "double exponential" extreme value distribution and the exponential residual life distribution, and (3.1) is their Corollary to Theorem 8. In the latter case $(0 < c \le r^{-1})$, 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. $\square$.

Similarly, sufficient conditions for Conditions 1a and 1b can be given: simply replace "2" in Condition 3 and "½" 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 $\mathrm{Var}[Z] \sim e^2/\overline{F} \sim (\overline{F}/f)^2/\overline{F} \sim 1/(-f')$. We treat three examples, not all 'typical', in more detail.

Example 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 2 with $q(t) = (1-t)^{\frac{1}{2}-\delta}$, $0 < \delta < \frac{1}{2}$. Conditions 1a and 1b hold with $h(t) = (1-t)^{1-\delta}$, $0 < \delta < 1$. Hence Theorems 1 and 2 hold where now

$$Z(x) = \frac{U(F(x))}{1-F(x)} - \frac{1}{1-F(x)} \int_{F(x)}^{1} \frac{U}{1-I} dI =_d \theta B(e^{x/\theta}), \qquad 0 \leq x < \infty.$$

and B is standard Brownian motion on $[0,\infty)$. (The process $B_1(t) = U(1-t) - \int_{1-t}^{1} (U/(1-I))dI$, $0 \leq t \leq 1$, is Brownian motion on $[0,1]$; and with $B_2(x) \equiv xB_1(1/x)$ for $1 \leq x \leq \infty$, $Z(x) = \theta B_2(1/\overline{F}(x)) = \theta B_2(e^{x/\theta})$.) Thus, in agreement with (2.4),

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

An immediate consequence is that $||Z_n^* \overline{F}|| \to_d ||Z\overline{F}|| =_d \theta \sup_{0 \leq t \leq 1} |B_1(t)|$; 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 2. (Weibull). Let $\overline{F}(x) = \exp(-x^\theta)$, $x \geq 0$, 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 2. Thus Theorems 1 and 2 hold. Also, $e(x) \sim \theta^{-1}x^{1-\theta}$ as $x \to \infty$, and hence $\mathrm{Var}[Z(x)] \sim \theta^{-2}x^{2(1-\theta)}\exp(x^\theta)$ as $x \to \infty$.

Example 3. (Pareto). Let $\overline{F}(x) = (1+cx)^{-1/c}$, $x \geq 0$, with $0 < c < \frac{1}{2}$. Then $e(x) = (1-c)^{-1}(1+cx)$, and Conditions 4a and 4b hold for $r < c^{-1}$ and $\gamma \geq c$ (and c of 4a is c). Thus Proposition 2 holds with $r > 2$ and $c > 0$ and $\mathrm{Var}[Z(x)] \sim c^{2+(1/c)}(1-c)^{-2}(1-2c)^{-1}x^{2+(1/c)}$ as $x \to \infty$. Conditions 1 and 2 hold with h and q as in Example 1, and Theorems 1 and 2 hold.

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

Not surprisingly, the limiting process 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 $Z' \equiv Z\overline{F}$ on $\mathbb{R}^+$ which appeared in (2.5) of Proposition 1. Its sample analog $Z_n' \equiv Z_n \overline{\mathbb{F}}_n$ is easily seen to be a cumulative sum (times $n^{-\frac{1}{2}}$) 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 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 Z'.

The zero-mean Gaussian process Z' has covariance function $\mathrm{Cov}[Z'(x),Z'(y)] = \mathrm{Var}[Z'(x\vee y)]$ (see (2.5)); hence, when viewed in reverse time, it has independent increments (and hence Z' is a reverse martingale). Specifically, with $Z''(s) \equiv Z'(-\log s)$, Z'' is a zero-mean Gaussian process on $[0,1]$ with independent increments and $\mathrm{Var}[Z''(s)] = \mathrm{Var}[Z'(-\log s)] \equiv \tau^2(s)$. 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 $\tau^2(0) = \lim_{\varepsilon\downarrow 0}\overline{F}(-\log\varepsilon)\sigma^2(-\log\varepsilon) = \lim_{x\to\infty}\overline{F}(x)\sigma^2(x) = 0$ 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 W on $[0,1]$ by

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

Theorem 3. W is standard Brownian motion on $[0,1]$.

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

Corollary 2. If (2.7) holds, then

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

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

Proof. By Theorem 0 of Wellner (1978) $||\frac{\overline{\mathbb{F}}_n}{\overline{F}} - 1||_0^{b_n} \to_p 0$ 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 $\mathbb{b}_n = \mathbb{F}_n^{-1}(1-a_n)$, the (n-m)th order statistic with $m = [na_n]$, leads to asymptotic confidence bands for $e = e_F$:

Corollary 3. Let $0 < a < \infty$. If (2.7) holds, $S_n^2 \to_p \sigma^2$, and $na_n/\log\log n \uparrow \infty$, then, as $n \to \infty$

(5.3) $$P(\hat{e}_n(x) - n^{-\frac{1}{2}}S_n a/\overline{\mathbb{F}}_n(x) \le e(x) \le \hat{e}_n(x) + n^{-\frac{1}{2}}S_n a/\overline{\mathbb{F}}_n(x) \quad \text{for all } 0 \le x \le \mathbb{b}_n) \to P(a)$$

where

$$P(a) \equiv P(||W||_0^1 < a) = \sum_{k=-\infty}^{\infty}(-1)^k[\Phi((2k+1)a)-\Phi((2k-1)a)]$$

$$= 1 - 4[\overline{\Phi}(a) - \overline{\Phi}(3a) + \Phi(5a) - \ldots]$$

and Φ denotes the standard normal df.

Proof. It follows immediately from Corollary 2 and $S_n \to_p \sigma > 0$ that $||Z_n \overline{\mathbb{F}}_n||_0^{b_n}/S_n \to_d ||Z\overline{F}||/\sigma = ||W||_0^1$. Finally b_n may be replaced by $\mathbb{b}_n$ without harm: letting $c_n = 2\log\log/na_n \to 0$ and using Theorem 4S of Wellner (1978), for $\tau > 1$ and all $n \ge N(\omega,\tau)$, $\mathbb{b}_n \equiv \mathbb{F}_n^{-1}(1-a_n) =_d F^{-1}(\Gamma_n^{-1}(1-a_n)) \le F^{-1}([1+\tau c_n^{\frac{1}{2}}](1-a_n))$ w.p.1. This proves (5.3); the expression for $P(a)$ is well-known (e.g. see Billingsley (1968), page 79). ▯

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

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

Thus, choosing a so that $P(a) = \beta$, (5.3) provides a two-sided simultaneous confidence band for the function e with confidence coefficient asymptotically β. In applications we suggest taking $a_n = n^{-\frac{1}{2}}$ so that $\mathbb{b}_n$ is the $(n-m)$th order statistic with $m = [n^{\frac{1}{2}}]$; 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 and Proposition 1 directly. We first require a consistent estimate of the asymptotic variance of $n^{\frac{1}{2}}(\hat{e}_n(x) - e(x))$, namely $\sigma^2(x)/\overline{F}(x)$.

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

<u>Proof</u>. 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 $\int_x^\infty y\,\overline{\mathbb{F}}_n(y)\,dy \to_{a.s.} \int_x^\infty y\overline{F}(y)\,dy$. Let $h(t) = (1-t)^{\gamma+\frac{1}{2}}$ and $q(t) = (1-t)^{\gamma}$ with $r^{-1} < \gamma < \frac{1}{2}$ 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 2. Then,

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

by Theorem 1 of Wellner (1977) since

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

By Theorem 2, Propositions 1 and 3, and Slutsky's theorem we have:

Proposition 4. Under the conditions of Proposition 3,

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

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 $[\overline{F}(y)\sigma^2(y)/\overline{F}(x)\sigma^2(x)]^{\frac{1}{2}}$ estimated by $[\overline{\mathbb{F}}_n(y)S_n^2(y)/\overline{\mathbb{F}}_n(x)S_n^2(x)]^{\frac{1}{2}}$, an asympotic 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 on 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 orgin. More specifically, a Maclaurin expansion yields

$$e(x) = \mu + (\mu f_o - 1)x + \tfrac{1}{2}[(2\mu f_o - 1)f_o + f_o']x^2 + o(x^2)$$

where $f_o = f(0)$, $f_o' = f'(0)$, 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_o is 0 and $f_o' > 0$, or f_o is near 0 (and $f_o' \geq 0$).

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 4). 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 or 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.)

Confidence bands on the difference between two mean residual life functions, and for the case of censored data, will be presented in subsequent papers.

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

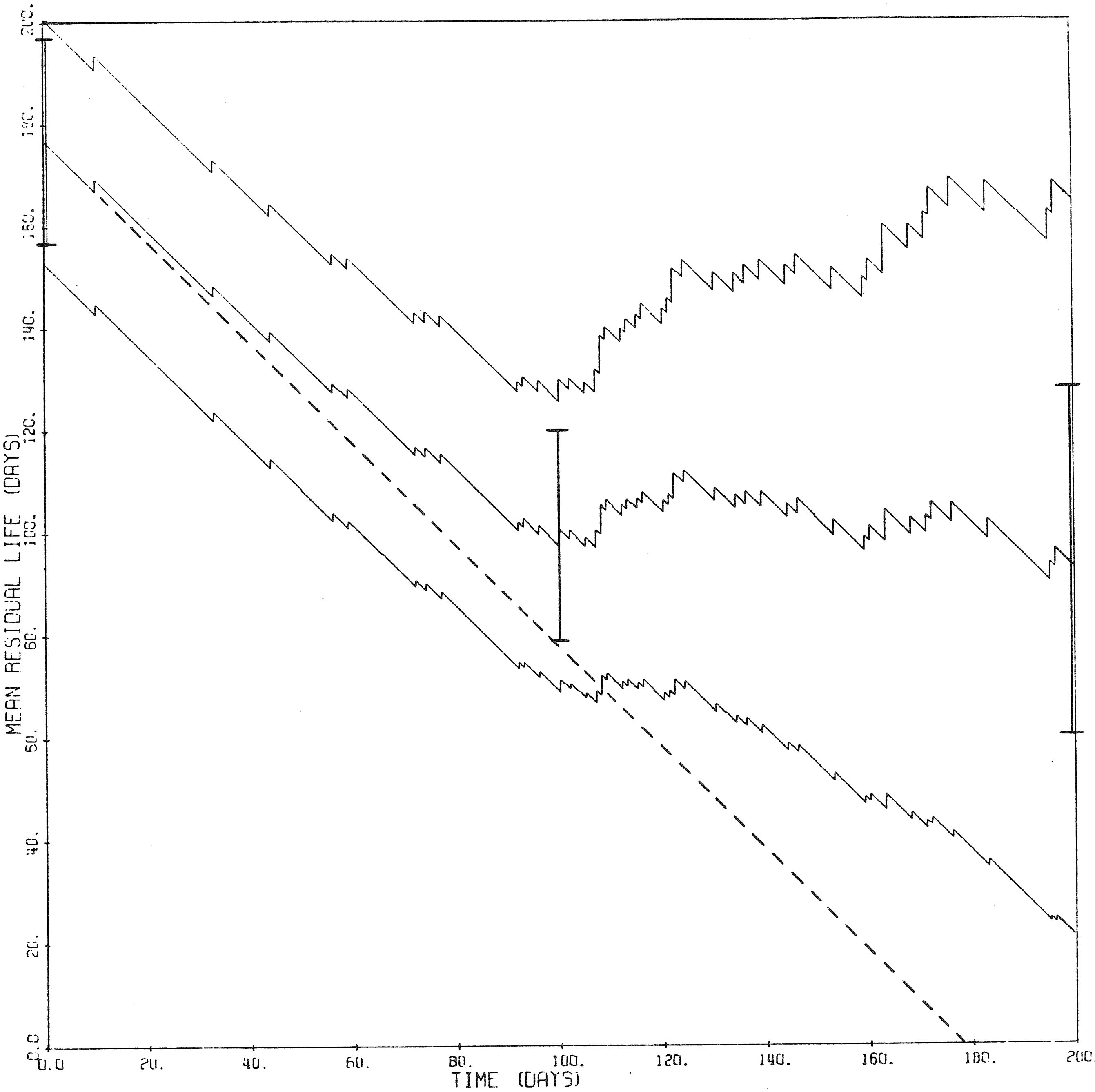
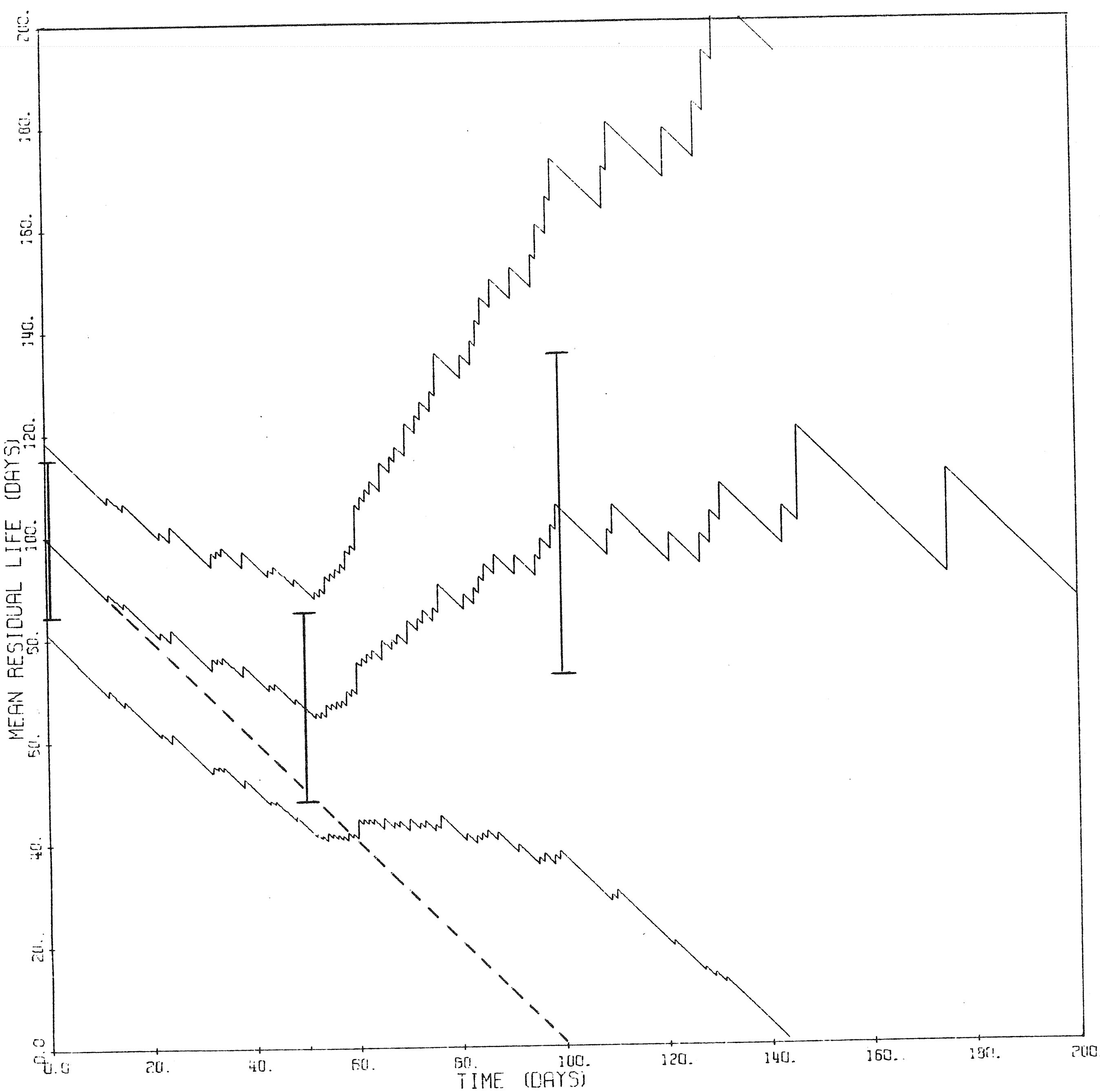


Figure 2. 90% confidence bands for mean residual life; Regimen 6.6



REFERENCES

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

Barlow, Richard E. and Campo, Raphael (1975). Total time on test processes and applications to failure data analysis. In Reliability and Fault Tree Analysis (ed. Barlow, Fussel, and Singpurwalla). SIAM, Philadelphia.

Billingsley, Patrick (1968). Convergence of Probability Measures. Wiley, New York.

Bjerkedal, Tor (1960). Acquisition of resistance in quinea pigs infected with different doses of virulent tubercle bacilli. Amer. Jour. Hygiene 72 130-148.

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

Chiang, Chin Long (1960). A stochastic study of the life table and its applications: I. Probability distributions of the biometric functions. Biometrics 16 618-635.

Chiang, Chin Long (1968). Introduction to Stochastic Processes in Biostatistics. Wiley, New York.

Gross, Alan J. and Clark, Virginia A. (1975). Survival distributions: Reliability Applications in the Biomedical Sciences. Wiley, New York.

Hollander, Myles and Proschan, Frank (1975). Tests for mean residual life. Biometrika 62 585-593.

Shorack, Galen R. (1972). Functions of order statistics. Ann. Math. Statist. 43 412-427.

Wellner, Jon A. (1977). A Glivenko-Cantelli theorem and strong laws of large numbers for functions of order statistics. Ann. Statist. 5 473-480.

Wellner, Jon A. (1978). Limit theorems for the ratio of the empirical distribution function to the true distribution function. Z. Wahrscheinlichkeitstheorie und Verw. Gebiete 45 74-88.

Wilson, Edwin B. (1938). The standard deviation of sampling for life expectancy. J. Amer. Statist. Assoc. 33 705-708.

Yang, Grace L. (1978). Estimation of a biometric function. Ann. Statist. 6 112-116.

Department of Statistics
University of Rochester
Rochester, N.Y. 14627