Tests and Confidence Sets for Comparing Two Mean Residual Life Functions

by

Roger L. Berger, Dennis D. Boos, and Frank M. Guess

Institute of Statistics Mimeograph Series #1673

December, 1985

SUMMARY

New confidence procedures and tests for comparing two populations based on their mean residual life functions are presented. The methods have the flexibility of handling crossings of the mean residual life functions. Another advantage is that the consistency class of the tests is given exactly by the alternative hypothesis. Examples illustrate how to apply the methods.

Some key words: Intersection-union principle; mean residual life functions; nonparametric tests; nonparametric confidence procedures.

1. INTRODUCTION

In this paper we describe hypothesis tests and confidence procedures that are useful for comparing the mean residual life functions from two populations. The mean residual life function is defined by

$$e_F(t) = E(X - t \mid X > t)$$

$$= \int_t^\infty \bar{F}(u)du/\bar{F}(t) \qquad \text{for } \bar{F}(t) > 0 ,$$

where the random variable X has distribution function F(t) and $\bar{F}(t) = 1 - F(t)$. If X measures the lifelength for an item, then $e_F(t)$ is the expected remaining life in the item given it has survived until time t. Since $e_F(t)$ can be inverted to obtain F, e.g., see Cox (1962, p. 128), mean residual life functions can be used directly to model lifelength distributions.

In this paper, we develop a nonparametric test of

$$H_0: \quad e_F(t) \le e_G(t) \qquad \text{for some } t \in [T_1, T_2]$$

$$\text{versus} \tag{1.1}$$

$$H_a: \quad e_F(t) > e_G(t) \qquad \text{for all } t \in [T_1, T_2]$$

based on independent samples from distribution functions F and G. An important feature of the hypotheses (1.1) is that the test may be inverted to obtain confidence statements of the form "$e_F(t) > e_G(t)$ for all $t \in \hat{I}$," where $\hat{I}$ is an interval of values determined by the data.

The mean residual life function arises naturally and is of practical interest in many applications. These include survivorship studies in medical settings, rate setting for life insurance, and industrial burn-in procedures. Mean residual life is also of interest when the variable is not lifelength. If X is the medical cost of a certain type of patient, then $e_F(t)$ is the expected remaining cost given that the amount t has been paid. Guess and Proschan (1985) provide a recent survey of important mean residual life applications and theory.

In all these applications, comparisons between two mean residual life functions might be of interest. For example, if Vendor A's electronic systems have lifelengths with distribution function F and Vendor B's have distribution function G, then a confidence statement that "$e_F(t) > e_G(t)$ for all $t \in [0,3000)$" would imply that for systems of age up to 3000 hours, Vendor A's systems would be preferred to Vendor B's in terms of average remaining life.

A variety of other population measures have also been used by reliability and biomedical researchers to compare populations. For example, failure rate functions have been compared in Chikkagoudar and Shuster (1974), Kochar (1979, 1981), and Cheng (1985), and percentile residual life functions have been compared in Joe and Proschan (1984).

Our test for mean residual life differences is constructed using the intersection-union method. This method and the resulting confidence statements are quite different from those used by the above authors and could be profitably applied in their contexts as well. Two other distinctive features of our approach are the following.

1. We allow the user the flexibility of choosing an interval $[T_1, T_2]$ of values on which to compare mean residual life functions. A common choice, however, would be $T_1 = 0$ and T_2 some large number within the expected range of the data. Thus, the mean residual life functions may cross each other, but suitable choices of T_1 and T_2 allow meaningful tests to be made. We discuss further the occurrence of crossing mean residual life functions in Section 4. See also Stablein and Koutrouvelis (1985) for examples of crossing failure rates.
2. Related to the possibility of crossings is the question of the consistency class. The consistency class for our tests is given exactly by H_a of (1.1). This is in contrast to most useful results available for comparing F and G in the literature. Such results are typically based on a natural parameter or functional of (F,G) which is estimated from the samples. However, the consistency class from such procedures is usually larger than the stated alternative. For such tests one can reject the null hypothesis but never really know if one is in the alternative set or just in the consistency class.

In Section 2 we develop the test for (1.1) and give relevant asymptotic results. Section 3 investigates the related confidence procedures. Examples in Section 4 illustrate how to use our methods.

2. TESTING FOR LARGER MEAN RESIDUAL LIFE

The goal is to test (1.1) based on independent random samples $X_1,\ldots,X_m$ and $Y_1,\ldots,Y_n$ from F and G. Our approach is to use the intersection-union principle discussed by Gleser (1973) and Berger and Sinclair (1984). We define an asymptotic level α test for each simple problem

$$H_{0t}: \quad e_F(t) \leq e_G(t)$$
$$H_{at}: \quad e_F(t) > e_G(t) \,, \qquad (2.1)$$

and then reject H_0 if and only if H_{0t} is rejected at level α for each $t \in [T_1, T_2]$. We will also demonstrate pairs $(F,G) \in H_0$ such that

$$P\,(\text{reject } H_0) \to \alpha \qquad \text{as } m \text{ and } n \to \infty \,.$$

These pairs show that the intersection-union test is not only level α but also size α asymptotically.

The obvious test statistic for (2.1) is

$$Z_{mn}(t) = \left\{\hat{e}_F(t) - \hat{e}_G(t)\right\} \Big/ \left\{\frac{S_m^2(t)}{m(t)} + \frac{S_n^2(t)}{n(t)}\right\}^{1/2}, \qquad (2.2)$$

where $\hat{e}_F(t) + t$ = the average of the X_i's greater than t, $S_m^2(t)$ = the sample variance of the X_i's greater than t, and $m(t)$ = the number of X_i's greater than t. The analogous quantities for the Y_i's are $\hat{e}_G(t)$, $S_n^2(t)$, and $n(t)$. Define $S_m^2(t) = 0$ when $m(t) \leq 1$, $S_n^2(t) = 0$ when $n(t) \leq 1$, and $Z_{mn}(t) = 0$ when $\min\{m(t), n(t)\} \leq 1$. Chiang (1960, pp. 226–7) discusses a discrete version of $Z_{mn}(t)$ for life table analysis; Hall and Wellner (1979) discuss the one-sample version of $Z_{mn}(t)$ and Elandt-Johnson and Johnson (1980, Section 8.3) discuss related statistics. Conditional on $m(t)$ and $n(t)$, the observations greater than t are distributed as samples from $\{F(x) - F(t)\}I(x > t)/\bar{F}(t)$ and $\{G(y) - G(t)\}I(y > t)/\bar{G}(t)$, respectively. Thus, one can easily calculate

$$E\{\hat{e}_F(t) - \hat{e}_G(t)\} = e_F(t)[1 - \{F(t)\}^m] - e_G(t)[1 - \{G(t)\}^n]$$

and other exact moments of the numerator and denominator of $Z_{mn}(t)$. Such calculations suggest that in small samples we use $m(t) - 1$ and $n(t) - 1$ in the denominators of $S_m^2(t)$ and $S_n^2(t)$, respectively. Chiang (1960) pools $S_m^2(t)$ and $S_n^2(t)$ in (2.2) but we prefer the unpooled version since $e_F(t) = e_G(t)$ does not imply var $\hat{e}_F(t)$ = var $\hat{e}_G(t)$. In general we expect $Z_{mn}(t)$ to be used in fairly large samples so that the following theorem is relevant. The proofs of Theorem 1 and all other theorems are in the Appendix.

Let $\sigma_F^2(t) = \mathrm{var}(X_1 | X_1 > t)$ and $\sigma_G^2(t) = \mathrm{var}(Y_1 | Y_1 > t)$ with the convention that $\sigma_F(t) = e_F(t) = 0$ if $\bar{F}(t) = 0$. Also we break $Z_{mn}(t)$ into two parts,

$$\begin{aligned} Z_{mn}(t) &= m^{1/2}[\{\hat{e}_F(t) - e_F(t)\} - \{\hat{e}_G(t) - e_G(t)\}] \Big/ D_{mn}(t) \\ &\quad + m^{1/2}\{e_F(t) - e_G(t)\} \Big/ D_{mn}(t) \\ &\equiv A_{mn}(t) + B_{mn}(t) , \end{aligned}$$

where $D_{mn}^2(t) = [S_m^2(t)/m(t) + S_n^2(t)/n(t)]m$. Let z_α be the $1 - \alpha$ quantile of the standard normal distribution.

THEOREM 1. Let $X_1,\ldots,X_m$ and $Y_1,\ldots,Y_m$ be independent samples with $P(X_1 \le x) = F(x)$ and $P(Y_1 \le y) = G(y)$. If $EX_1^2 < \infty$ and $EY_1^2 < \infty$ and

$\min(m,n) \to \infty$, then

(i) $S_m^2(t) \to \sigma_F^2(t)$ and $S_n^2(t) \to \sigma_G^2(t)$ with probability one,

(ii) $A_{mn}(t) \to N(0,1)$ in distribution, where $N(0,1)$ is a standard normal random variable,

and thus

$$\text{(iii)}\quad P(Z_{mn}(t) \geq z_\alpha) \to \begin{cases} 0 & \text{if } e_F(t) < e_G(t) \\ \alpha & \text{if } e_F(t) = e_G(t) \\ 1 & \text{if } e_F(t) > e_G(t) \end{cases}.$$

Part (iii) tells us that the test of H_{0t} versus H_{at} which rejects H_{0t} if and only if $Z_{mn}(t) \geq z_\alpha$ is asymptotically level α and consistent. Of course when m(t) and n(t) are small, one should replace z_α by appropriate t distribution quantiles.

We now propose combining these individual tests via the intersection-union method to obtain the following test of H_0 versus H_a: reject H_0 if and only if

$$\bigcap_{t\in[T_1,T_2]} \{Z_{mn}(t) \geq z_\alpha\} \text{ occurs,} \tag{2.3}$$

or equivalently if

$$\inf_{t\in[T_1,T_2]} Z_{mn}(t) \geq z_\alpha \text{ occurs.}$$

This test is asymptotically level α. That is, for any (F,G) in H_0 with $e_F(t_0) \leq e_G(t_0)$, $t_0 \in [T_1,T_2]$, by Theorem 1 we have

$$P(\bigcap_{t\in[T_1,T_2]} Z_{mn}(t) \geq z_\alpha)$$

$$\leq P(Z_{mn}(t_0) > z_\alpha)$$

$$\leq P(A_{mn}(t_0) > z_\alpha) \to \alpha \quad \text{as} \quad \min(m,n) \to \infty .$$

Consistency of our test (2.3) is not nearly so easy to prove and requires weak convergence of the $A_{mn}(t)$ process on $[T_1,T_2]$. The details are given in the Appendix but the result is straightforward.

THEOREM 2 (consistency). Let $X_1,\ldots,X_m$ and $Y_1,\ldots,Y_n$ be independent samples from *continuous* distribution functions F(x) and G(y), respectively. Suppose that

(i) $E|X_1|^r < \infty$, $E|Y_1|^r < \infty$ for some $r > 2$.

(ii) $\bar{F}(T_2) > 0$, $\bar{G}(T_2) > 0$.

(iii) $e_F(t) > e_G(t)$ for all $t \in [T_1,T_2]$.

If $\min(m,n) \to \infty$ such that $m/(m + n) \to \lambda$, $0 \leq \lambda \leq 1$, then

$$P(\inf_{t\in[T_1,T_2]} Z_{mn}(t) > z_\alpha) \to 1 .$$

For a given pair $(F,G) \in H_o$, the probability of rejection using (2.3) often converges to a number which is strictly less than α. In this next theorem we show that the convergence is to α for those pairs (F,G) in H_o for which $e_F(t) = e_G(t)$ at only one point t_o. The intuitive reason for the result is that asymptotically $\inf_t Z_{mn}(t)$ can be replaced by $Z_{mn}(t_o)$ which converges to a standard normal random variable. The proof in the Appendix relies on a Skorohod representation of $A_{mn}(t)$. A related proof is found in Berger (1984, Theorem 1).

THEOREM 3 (asymptotic size α). Assume the conditions of Theorem 2 except that (iii) is replaced by (iii′) $e_F(t_o) = e_G(t_o)$ for some $t_o \varepsilon [T_1,T_2]$ and $e_F(t) > e_G(t)$ for all $t \varepsilon \{[T_1,T_2] - t_o\}$. If $\min(m,n) \to \infty$ such that $m/(m + n) \to \lambda$, $0 \leq \lambda \leq 1$, then

$$P\left(\inf_{t\varepsilon[T_1,T_2]} Z_{mn}(t) > z_\alpha \right) \to \alpha .$$

Finally, our last result in this section gives an asymptotic justification for computing approximate power in cases where $\{e_F(t) - e_G(t)\}/D(t)$ has a unique minimum at t_o. Here $D(t)$ is the probability limit of $D_{mn}(t)$. The idea is similar to the last theorem in that $\inf_t Z_{mn}(t)$ is replaced asymptotically by $A_{mn}(t_o) + m^{1/2}\{e_F(t_o) - e_G(t_o)\}/D(t_o)$. In order to keep the power from going to one, we must let the "true" distributions of $X_1,\ldots,X_m$ and $Y_1,\ldots,Y_n$ given by F^m and G^n, respectively, converge to F and G in such a way that $m^{1/2}\{e_{F^m}(t_o) - e_{G^n}(t_o)\}/D_{F^mG^n}(t_o)$ converges to a finite constant. Details of the proof are in the Appendix.

THEOREM 4 (asymptotic power). Let $X_1,\ldots,X_m$ and $Y_1,\ldots,Y_n$ be independent samples from continuous distribution functions $F^m(x)$ and $G^n(y)$, respectively. Suppose that

(i) $\max\{E|X_1|^r, E|Y_1|^r\} \leq M < \infty$ for some $r > 2$ and all $m = 1,2,\ldots$ and $n = 1,2,\ldots$.

(ii) $\sup_{t\varepsilon[T_1,T_2]} |F^m(t) - F(t)|$, $\int_0^\infty |F^m(t) - F(t)|dt$,

$\sup_{t\varepsilon[T_1,T_2]} |G^n(t) - G(t)|$, and $\int_0^\infty |G^n(t) - G(t)|dt$ each converge to 0 as $\min(m,n) \to \infty$.

(iii) $\bar{F}(T_2) > 0$, $\bar{G}(T_2) > 0$, $\sigma_F^2(T_2) > 0$, $\sigma_G^2(T_2) > 0$.

(iv) $L_{mn}(t) = m^{1/2}\{e_{F^m}(t) - e_{G^n}(t)\}/D_{F^mG^n}(t)$ has a unique minimum at t_{0mn}, and $t_{0mn} \to t_0$, $L_{mn}(t_{0mn}) \to L_0 < \infty$, and $L_{mn}(a_{mn}) \to \infty$ if a_{mn} is bounded away from t_0 as $\min(m,n) \to \infty$.

If $\min(m,n) \to \infty$ such that $m/(m+n) \to \lambda$, $0 \leq \lambda \leq 1$, then

$$P(\inf_{t\varepsilon[T_1,T_2]} Z_{mn}(t) > z_\alpha) \to P(N(0,1) > z_\alpha - L_0) .$$

Example 1. Let $\bar{F}(x) = \exp(-x/\sigma_1)$ and $\bar{G}(x) = \exp(-x/\sigma_2)$. Then $e_F(t) = \sigma_1$, $e_G(t) = \sigma_2$, $D^2(t) = D^2(\sigma_1,\sigma_2,t) = \sigma_1^2 \exp(t/\sigma_1) + \{\lambda/(1-\lambda)\}\sigma_2^2\exp(t/\sigma_2)$. Since $D^2(t)$ is strictly increasing in t, the minimum of $m^{1/2}\{e_F(t) - e_G(t)\}/D(t)$ on $[0,T]$ is always achieved at $t = T$ when $\sigma_1 > \sigma_2$. Thus if $\sigma_1 = b + c/m^{1/2}$ and $\sigma_2 = b - c/m^{1/2}$, we can use Theorem 4 to justify the approximate power calculation $P(N(0,1) > z_\alpha - 2b/D(b,b,T))$. In Table 1 we give a few such values for equal sample sizes $\lambda = 1/2$ but have used $D(\sigma_1,\sigma_2,T)$ in place of $D(b,b,T)$ because it seems more intuitive.

Example 2. Let $\bar{F}(x) = (1 + x)\exp(-x)$, a Gamma distribution with $\alpha = 2$, and let $\bar{G}(x) = \exp(-x/\sigma)$. Then $e_F(t) = (2 + t)/(1 + t)$ and

TABLE 1. Approximate Power Calculations for Comparing Mean Residual Life of Two Exponential Distributions on [0,T] at $\alpha = 0.10$.

	m = n = 50		m = n = 100	
	$\sigma_1 = 1.1$ $\sigma_2 = 0.9$	$\sigma_1 = 1.5$ $\sigma_2 = 0.5$	$\sigma_1 = 1.1$ $\sigma_2 = 0.9$	$\sigma_1 = 1.5$ $\sigma_2 = 0.5$
T = 0.5	0.31	0.73	0.43	0.92
T = 1.0	0.25	0.58	0.33	0.80
	$\sigma_1 = 2.1$ $\sigma_2 = 1.9$	$\sigma_1 = 2.5$ $\sigma_2 = 1.5$	$\sigma_1 = 2.1$ $\sigma_2 = 1.9$	$\sigma_1 = 2.5$ $\sigma_2 = 1.5$
T = 0.5	0.20	0.43	0.26	0.61
T = 1.0	0.19	0.38	0.23	0.54
T = 2.0	0.16	0.30	0.20	0.42

TABLE 2. Approximate Power Calculations for Comparing Mean Residual Life of Gamma (2) with Exponential (σ) on [0,T] at $\alpha = 0.10$.

	m = n = 50		m = n = 100	
	$\sigma = 1.0$	$\sigma = 2^{1/2}$	$\sigma = 1.0$	$\sigma = 2^{1/2}$
T = 0.25	0.97	0.51	1.00	0.71
T = 0.50	0.88	0.32	0.99	0.44
T = 1.00	0.61	0.15	0.82	0.17

$$D^2(\sigma,t) = \frac{(2+4t+t^2)}{(1+t)^3}\, e^t + \left(\frac{\lambda}{1-\lambda}\right)\sigma^2 e^{t/\sigma}.$$

As in the previous example, $D^2(\sigma,t)$ is increasing in t and thus $m^{1/2}\{e_F(t) - e_G(t)\}/D(\sigma,t)$ is minimized at the largest t value considered. Table 2 gives some asymptotic power calculations for several values of σ and intervals $[0,T]$.

3. CONFIDENCE PROCEDURES

In this section we describe three confidence procedures based on the statistic $Z_{mn}(t)$. Each procedure produces a confidence statement of the form "$e_F(t) > e_G(t)$ for all $t \in \hat{I}$" where $\hat{I} \subseteq [0,\infty)$ is a random interval computed from the data. We show in Theorem 5 that the asymptotic confidence level for each procedure is $1-\alpha$.

The three procedures differ in the form of the interval $\hat{I}$, but each requires that a "starting value" $T \geq 0$ be specified in advance of sampling. If $Z_{mn}(T) \leq z_\alpha$, then no confidence statement is made. If $Z_{mn}(T) > z_\alpha$, then the three procedures assert that $e_F(t) > e_F(t)$ for all $t \in \hat{I}$ where the $\hat{I}$ for each procedure is defined as follows:

Procedure 1: $\hat{I} = [T,\hat{\theta}_1)$ where $\hat{\theta}_1 = \inf\{t \geq T\colon Z_{mn}(t) \leq z_\alpha\}$.

Procedure 2: $\hat{I} = (\hat{\theta}_2,T]$ where $\hat{\theta}_2 = \sup\{t \leq T\colon Z_{mn}(t) \leq z_\alpha\}$.

Procedure 3: $\hat{I} = (\max(0,T - \hat{\delta}),\ T + \hat{\delta})$ where

$$\hat{\delta} = \sup\{d \geq 0\colon \inf_{t \in [\max(0,T-d),T+d]} Z_{mn}(t) > z_\alpha\} .$$

As mentioned in Section 2, we replace z_α by t distribution percentiles whenever m(t) and n(t) are small.

In the first type of interval, T might be chosen to be 0. Then $\hat{\theta}_1$ is the smallest value of t for which $Z_{mn}(t) \leq z_\alpha$. Figure 1 illustrates the use of all three types of intervals with simulated data from F lognormal such that log X is N(log2,¼) and G exponential with mean = 1.8. The sample sizes were m = n = 200. The statistic $Z_{mn}(t)$ is $\leq z_{.10} = 1.282$ for the first time at t = 0.479. Thus the first confidence procedure asserts with confidence 90% that $e_F(t) \succ e_G(t)$ for all $t \in [0,0.479)$.

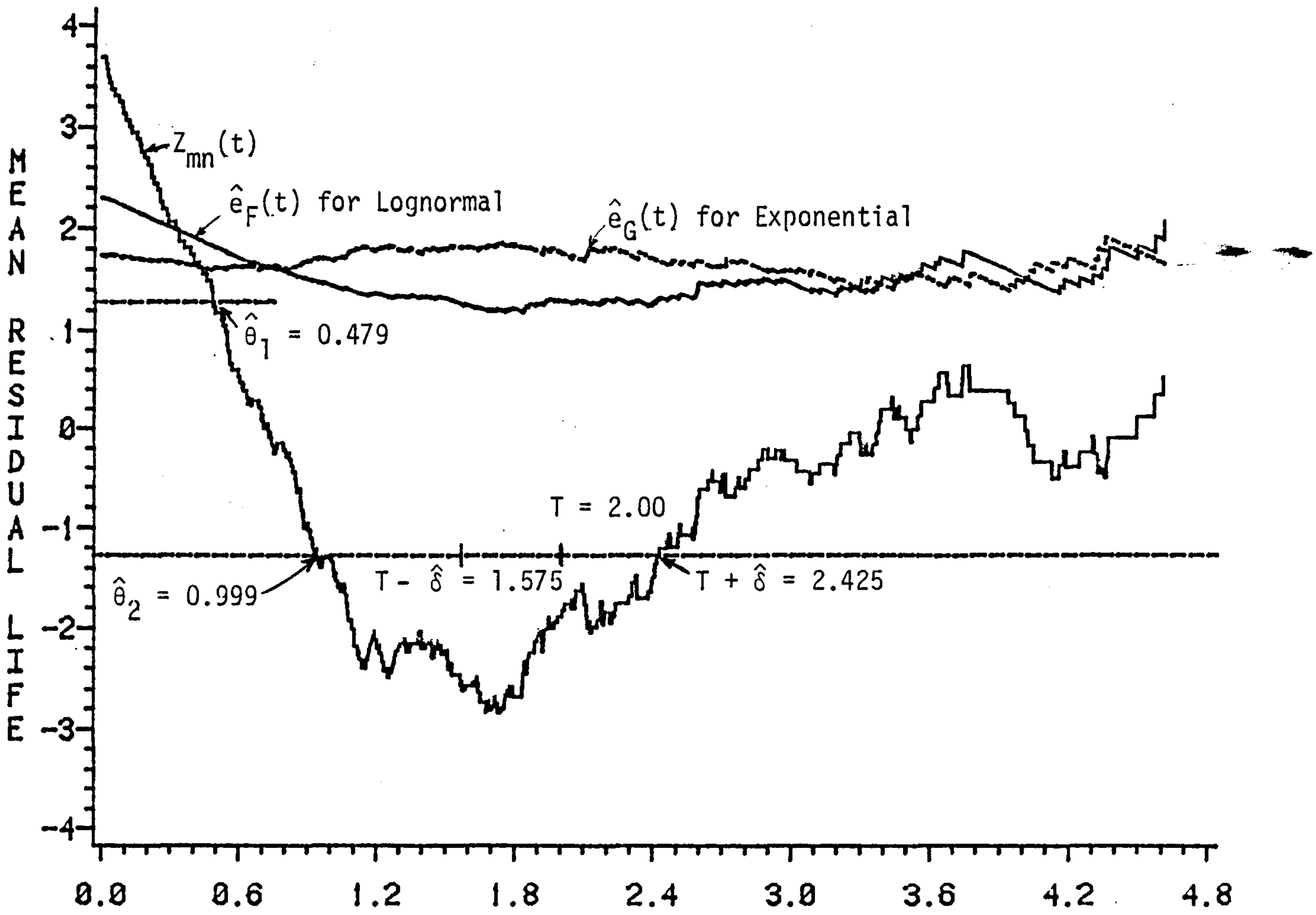
To illustrate the second type of interval, suppose that T = 2 was chosen as the upper limit of our interest. Reverse the roles of F and G in Figure 1. The largest $t \leq T = 2$ for which $Z_{mn}(t) \leq 1.282$ is t = 0.999. Actually $Z_{mn}(t) \geq -1.282$ in Figure 1. So the second confidence procedure asserts with confidence 90% that $e_G(t) \succ e_F(t)$ for all $t \in (0.999,2.000]$.

To illustrate the third type of interval, choose T = 2 as the center of the interval of interest. The value of t closest to T = 2 at which $Z_{mn}(t) \geq -1.282$ is t = 2.425. Thus $\hat{\delta} = 0.425$ and the third confidence procedures asserts, with confidence 90%, that $e_G(t) \succ e_F(t)$ for all $t \in (1.575, 2.425)$.

The asymptotic justification for these three procedures is given in Theorem 5, and its proof is in the Appendix.

THEOREM 5. Let $X_1,\ldots,X_m$ and $Y_1,\ldots,Y_n$ be independent samples from distribution functions F(x) and G(y), respectively. Suppose that

FIGURE 1. MEAN RESIDUAL LIFE FUNCTIONS
FROM SIMULATED LOGNORMAL AND EXPONENTIAL DISTRIBUTIONS



NOTE : HORIZONTAL DOTTED LINES ARE AT NORMAL PERCENTILES -1.282 AND 1.282

$EX_1^2 < \infty$ and $EY_1^2 < \infty$. Then

$$\lim_{\min(n,m)\to\infty} P(\text{Procedure 1 makes an incorrect statement}) \leq \alpha.$$

If in addition F(x) is a continuous function, then

$$\lim_{\min(m,n)\to\infty} P(\text{Procedure i makes an incorrect statement}) \leq \alpha \text{ for } i=2,3.$$

It is also easy to show using Theorem 1 that lim P(no statement is made) = lim $P(Z_{mn}(T) \leq z_\alpha) = 0$ if $e_F(T) > e_G(T)$. However, P(no statement is made) can be large when m and n are moderate and $e_F(t)$ exceeds $e_G(t)$ by only a small amount.

4. EXAMPLES

We illustrate the use of the hypothesis test and confidence procedures on data from two different experiments. The first experiment studied the lifelengths of guinea pigs after injection with different amounts of tubercle bacilli.

Guinea pigs are known to have a high susceptibility to human tuberculosis, which is one reason for choosing this species. Bjerkedal (1960) studies the acquisition of resistance in these animals and provides the data that we use here. His study labeled M is for animals in a single cage under the same regimen. The regimen number is the common log of the number of bacillary units in 0.5 *ml* of the challenge solution. E.g., regimen 4.3 corresponds to 2.2×10^4 bacillary units per 0.5 *ml*, $\log_{10}(2.2 \times 10^4) = 4.342$. Here we compare regimen 4.3 to that of regimen 5.5. It is reasonable *a priori* to hypothesize H_a of (1.1) with

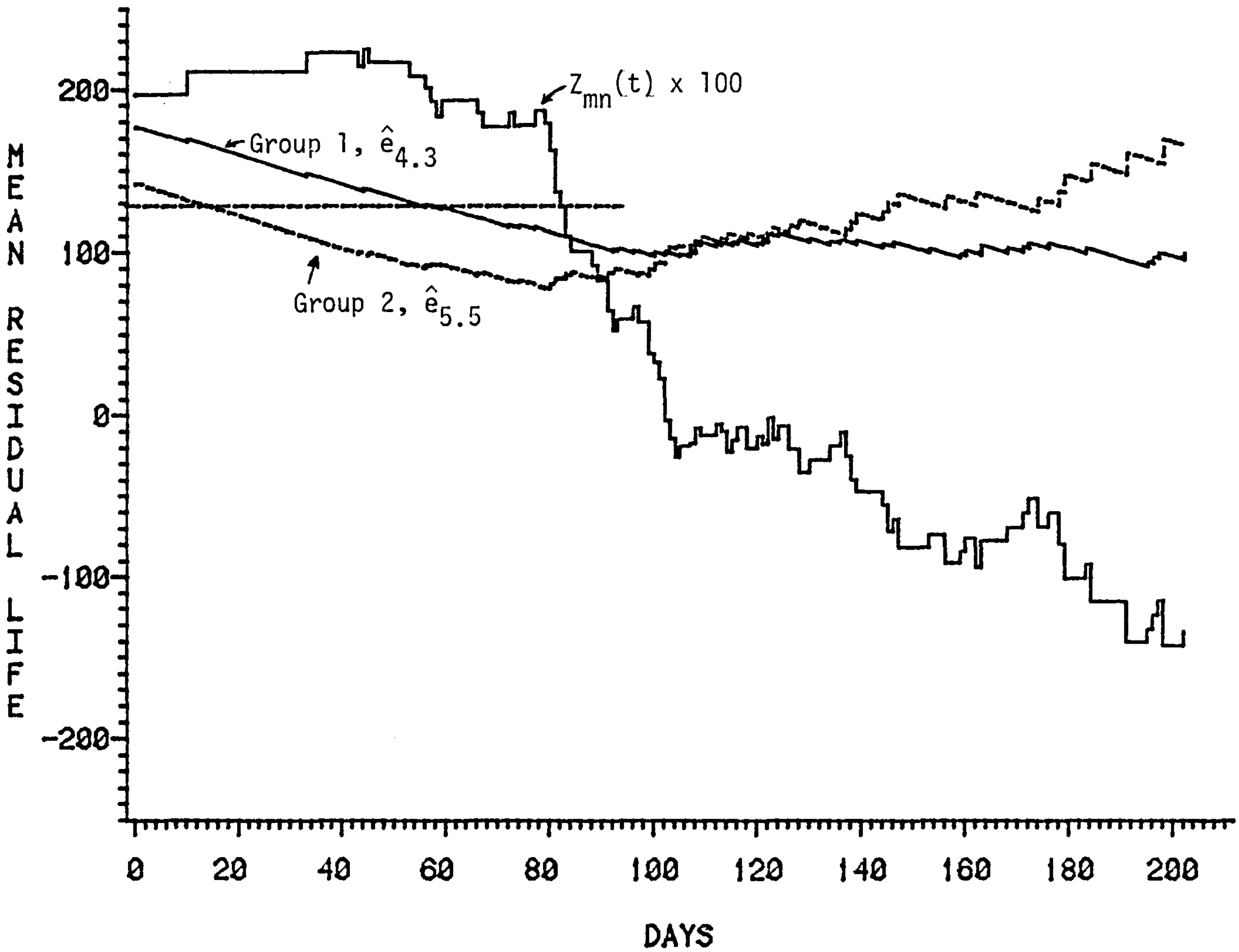
$T_1 = 0$, $e_F(t) = e_{4.3}(t)$, and $e_G(t) = e_{5.5}(t)$. Some natural T_2 values to consider are $T_2 = 30$, 60, or 90 days. Figure 2 shows the estimates $\hat{e}_{4.3}(t)$ and $\hat{e}_{5.5}(t)$ and the corresponding $Z_{mn}(t)$ function. At the approximate level $\alpha = 0.05$ we reject H_0 in favor of H_a for $T_2 = 30$ and for $T_2 = 60$. For $T_2 = 90$ we do not reject H_0.

The confidence procedures provide more insight into the testing results. If we specify $T = 0$ in the first confidence procedure, then we can assert with 90% confidence that $e_{4.3}(t) \succ e_{5.5}(t)$ for all $t \in [0,82)$. Note that $t = 82$ is the first time that $Z_{mn}(t)$ crosses $z_{.10} = 1.282$, the 90^{th} percentile of the standard normal distribution. If our main interest were in the region near $T = 60$, then the second confidence procedure would give the interval (0,60] and the third would give (38,82).

Note in Figure 2 that the empirical mean residual life functions cross each other. One explanation is that the shock of tubercle bacilli is more severe under regimen 5.5 initially, and thus $e_{4.3}(t) \succ e_{5.5}(t)$ for the first interval. The later reversal suggests that a more vigorous subgroup has been *screened* after surviving the initial disease state under regimen 5.5. Thus $e_{4.3}(t) \prec e_{5.5}(t)$ is reasonable for the later period. This example illustrates the flexibility of the techniques under natural crossings.

In the second example the empirical mean residual life functions do not cross. The context involves the influence of different diets on the aging process where research indicates that diet restriction promotes longevity. Yu, Masoro, Murato, Betrand, Lynd (1982) study the effects

FIGURE 2. MEAN RESIDUAL LIFE OF GUINEA PIGS INJECTED WITH DIFFERENT AMOUNTS OF TUBERCLE BACILLI. FROM BJERKEDAL(1960)



NOTE : HORIZONTAL DOTTED LINE IS AT 128.2 =(100) X 90TH PERCENTILE OF A STANDARD NORMAL DISTRIBUTION.

TABLE 3. Rat Lifelength Data in Days

Restricted Diet

105	605	811	931	1011	1073	1133	1190	1244
193	630	833	940	1012	1076	1136	1203	1258
211	716	868	957	1014	1085	1138	1206	1268
236	718	871	958	1017	1090	1144	1209	1294
302	727	875	961	1032	1094	1149	1218	1316
363	731	893	962	1039	1099	1160	1220	1327
389	749	897	974	1045	1107	1166	1221	1328
390	769	901	979	1046	1119	1170	1228	1369
391	770	906	982	1047	1120	1173	1230	1393
403	789	907	1001	1057	1128	1181	1231	1435
530	804	919	1008	1063	1129	1183	1233	
604	810	923	1010	1070	1131	1188	1239	

Ad libitum Diet

89	536	630	668	695	717	739	770	801
104	545	635	670	697	720	741	773	806
387	547	639	675	698	721	743	777	807
465	548	648	677	702	730	746	779	815
479	582	652	678	704	731	749	780	836
494	606	653	678	710	732	751	788	838
496	609	654	681	711	733	753	791	850
514	619	660	684	712	735	764	794	859
532	620	665	688	715	736	765	796	894
533	621	667	694	716	738	768	799	963

of a restricted diet on rates versus an *ad libitum* diet, i.e., free eating. Cf. also Witten (1985). Table 3 contains the data.

Let e_F correspond to the mean residual life function under the restricted diet and e_G under *ad libitum*. It is of interest to test (1.1) for $T_1 = 0$ and $T_2 = 730$ days, or equivalently 2 years. The test overwhelmingly rejects H_0 in favor of H_a at level $\alpha = 0.01$. In fact, $Z_{mn}(t) = 9.2$ at $t = 0$ days and $Z_{mn}(t) = 15.1$ at $t = 730$ days. Although a subjective graphical analysis would lead to similar conclusions, a distinct advantage of our procedures is the objective quantitative assessment of the data.

For a confidence statement let $T = 0$. Then we can assert with 99% confidence that $e_F(t) > e_G(t)$ for $t \in [0,894)$ days.

ACKNOWLEDGEMENTS

We gratefully acknowledge a helpful discussion with Dr. F. Proschan. We also thank Dr. B. P. Yu for his data in Table 3, which Dr. M. Witten provided, and Mr. J. Cho for the computations used in Section 4.

APPENDIX

<u>Proof of Theorem 1.</u> (i) Let $X_{(1)} \leq \ldots \leq X_{(m)}$ be the ordered X values. Then $S_m^2(t)$ can be written as a continuous function of $m^{-1}\Sigma_{i=1}^m I(X_i > t)X_i^2, m^{-1}\Sigma_{i=1}^m I(X_i > t)X_i$, $\bar{F}_m(t) = m^{-1}\Sigma_{i=1}^m I(X_i > t)$, and $I(X_{(m)} > t)$, each of which converges with probability one to the appropriate quantity. A similar result holds for $S_n^2(t)$. (ii) The Central Limit theorem applied to the sums $\Sigma_{i=1}^m I(X_i > t)\{X_i - t - e_F(t)\}$ and $\Sigma_{i=1}^n I(Y_i > t)\{Y_i - t - e_G(t)\}$ along with Slutsky's theorem and (i) yield

the result. Note that if $n/(m+n) \not\to \lambda$, $0 < \lambda < 1$, then a subsequence argument after assuming that (ii) is not true leads to a contradiction. (iii) follows from (ii) since $B_{mn}(t)$ converges in probability to $-\infty$ if $e_F(t) < e_G(t)$ and $B_{mn}(t)$ converges in probability to $+\infty$ if $e_F(t) > e_G(t)$.

<u>Proof of Theorem 2.</u> First we give two lemmas. Let $D[T_1,T_2]$ be the space of functions on $[T_1,T_2]$ that are right-continuous and have left-hand limits (see Billingsley, 1968, Ch. 3). We often write $\inf_t$ as shorthand notation for infimum over $t \in [T_1,T_2]$.

LEMMA A1. Under the assumptions of Theorem 2,

$$A_{mn}(t) \to A(t) \text{ weakly in } D[T_1,T_2]$$

with respect to the Skorohod topology, where $A(t)$ is a mean zero Gaussian process with continuous sample paths.

PROOF. Following Hall and Wellner (1979, p. 5) we note that $m^{1/2}\{\hat{e}_F(t) - e_F(t)\}$ has the same distribution as

$$\frac{1}{\bar{K}_m(F(t))}\left[-\int_t^\infty U_m(F(x))dx + e_F(t)U_m(F(t))\right]$$

$$\equiv Q(U_m,K_m,F;t)$$

where $U_m(t) = m^{1/2}\{K_m(t) - t\}$ and $K_m(t)$ is the empirical distribution of the specially constructed uniform random variables given in the appendix of Shorack (1972) which satisfy for each $0 < \ell < 1/2$

$$\rho_\ell(U_m,U_o) = \sup_{x\in[0,1]}\left|\frac{U_m(x) - U_o(x)}{\{x(1-x)\}^\ell}\right| \quad \text{converges in probability to 0}$$

and U_o is a Brownian Bridge. A similar representation is made for $n^{1/2}\{\hat{e}_G(t) - e_G(t)\}$ in terms of V_n, L_n, and G. Define $A(t) = Q(U_o, I, F; t) - \{\lambda/(1-\lambda)\}^{1/2} Q(V_o, I, G; t)$. Then one can shown that $\sup_t |A_{mn}(t) - A(t)|$ converges in probability to 0 using the convergence of $\rho_\ell(U_m, U_o)$ and $\rho_\ell(V_m, V_o)$ and that of $D_{mn}(t)$ given in Lemma A2 below. For example, the key step is

$$\sup_t |Q(U_m, K_m, F; t) - Q(U_0, I, F; t)|$$
$$\leq \int_0^\infty |U_m(F(x)) - U_0(F(x))| dx + \sup_t |e_F(t)| \sup_t |U_m(F(t)) - U_0(F(t))|$$
$$\leq \rho_\ell(U_m, U_o) \int_0^\infty [F(x)(1 - F(x))]^\ell dx + \sup_t |e_F(t)| \rho_\ell(U_m, U_0) .$$

The integral is finite because F has a finite moment >2 and $e_F(t)$ is bounded since $\bar{F}(T_2) > 0$. Now since $A_{mn}(t)$ converges in the uniform metric, it also converges in the Skorohod metric (Billingsley, p. 150), and convergence in probability implies convergence in distribution which is the conclusion of the lemma.

LEMMA A2. Under the conditions of Theorem 2 with $\lambda < 1$, we have

$$\sup_{t \varepsilon [T_1, T_2]} |D_{mn}(t) - D(t)| \text{ converges almost surely to } 0$$

where $D^2(t) = \sigma_F^2(t)/\bar{F}(t) + \{\lambda/(1 - \lambda)\}\sigma_G^2(t)/\bar{G}(t)$.

PROOF. The proof is similar to Proposition 3 of Hall and Wellner (1979). Let F_m be the empirical distribution function of the X's and note that

$$S_m^2(t) = \left[\frac{1}{\bar{F}_m(t)} \int_t^\infty x^2 dF_m(x) - \{\hat{e}_F(t) + t\}^2\right] I(X_{(m-1)} > t) .$$

The key step is the following bound after using integration by parts.

$$\left|\int_t^\infty x^2 d\{F_m(x) - F(x)\}\right| = \left|t^2\{F(t) - F_m(t)\} + 2\int_t^\infty \{F(x) - F_m(x)\}x dx\right|$$

$$\le T_2^2 \sup_t |F_m(t) - F(t)| + 2 \sup_t \left|\frac{F_m(t)-F(t)}{\{\bar{F}(t)\}^{\gamma+\frac{1}{2}}}\right| \int_t^\infty x\{\bar{F}(x)\}^{\gamma+\frac{1}{2}} dx .$$

where $r^{-1} < \gamma < \frac{1}{2}$. The Glivenko-Cantelli theorem handles the first term and the second term converges to zero in probability using Theorem 1 of Wellner (1977) and noting that $E|X_1|^r < \infty$ is enough to bound the integral. A similar result holds for $S_n^2(t)$. The conclusion of Lemma A2 then follows if $\lambda < 1$.

Now we can give the main argument for Theorem 2. Without loss of generality we assume $\lambda < 1$, otherwise we redefine $D_{mn}(t)$ with $n^{-\frac{1}{2}}$ factored out. Then

$$P(\inf_t Z_{mn}(t) \le z_\alpha) = P(\inf_t \{A_{mn}(t) + B_{mn}(t)\} \le z_\alpha)$$

$$\le P(\inf_t A_{mn}(t) + \inf_t B_{mn}(t) \le z_\alpha)$$

$$\le P(\inf_t A_{mn}(t) < z_\alpha - n^{\frac{1}{4}})$$

$$+ P(\inf_t B_{mn}(t) \le n^{\frac{1}{4}}) .$$

Now, $\inf_t(\cdot)$ is a continuous functional in $D[T_1,T_2]$ with respect to the Skorohod topology and thus Lemma A1 gives

$\inf_t A_{mn}(t)$ converges in distribution to $\inf_t A(t)$.

We can show that $\inf_t A(t)$ is finite-valued with probability one and thus $P(\inf_t A_{mn}(t) < z_\alpha - n^{1/4}) \to 0$. Using Lemma A2 and Condition (iii) of Theorem 2 we can show that $m^{-1/2}[\inf_t B_{mn}(t)]$ stays strictly above zero and thus $P(\inf_t B_{mn}(t) \leq n^{1/4}) \to 0$ and the result follows.

<u>Proof of Theorem 3.</u> Since $A_{mn}(t)$ and $D_{mn}(t)$ were shown to converge to $A(t)$ and $D(t)$ in probability in $D[T_1,T_2]$, we can assert that they converge jointly in $D[T_1,T_2] \times D[T_1,T_2]$ and use Skorohod's theorem to get representations with the same exact distribution and such that almost surely

$$\sup_t |A_{mn}(t) - A(t)| + \sup_t |D_{mn}(t) - D(t)| \to 0. \tag{1}$$

Let Ω_0 be the subset of the underlying probability space such that the above convergence holds. We will show that for each $\omega \in \Omega_0$

$$\inf_t Z_{mn}(t,\omega) = \inf_t \left[A_{mn}(t,\omega) + m^{1/2} \frac{\{e_F(t)-e_G(t)\}}{D_{mn}(t,\omega)} \right] \to A(t_0,\omega) , \tag{2}$$

where $A(t,\omega)$ is the sample path of $A(t)$ corresponding to ω. If (2) holds, then certainly $\inf_t Z_{mn}(t)$ converges in distribution to $A(t_0)$ and $A(t_0)$ is a standard normal random variable so that

$$P(\inf_t Z_{mn}(t) > z_\alpha) \to P(A(t_0) > z_\alpha) = \alpha .$$

To show (2), note first that since $B_{mn}(t_0) = 0$

$$\inf_t Z_{mn}(t,\omega) \leq Z_{mn}(t_0,\omega) = A_{mn}(t_0,\omega) \tag{3}$$

and thus

$$\overline{\lim_{m,n\to\infty}} \inf_t Z_{mn}(t,\omega) \leq \overline{\lim_{m,n\to\infty}} A_{mn}(t_0,\omega) = A(t_0,\omega) \tag{4}$$

by (1). Since $Z_{mn}(t)$ is a step function with a finite number of jumps, $\inf_t Z_{mn}(t,\omega)$ is attained for some $t_{mn} \varepsilon [T_1,T_2]$. We can show $t_{mn} \to t_0$ by a contradiction argument. We then have

$$\inf_t Z_{mn}(t,\omega) = A_{mn}(t_{mn},\omega) + \frac{m^{1/2}\{e_F(t_{mn})-e_G(t_{mn})\}}{D_{mn}(t_{mn},\omega)}$$

$$\geq A_{mn}(t_{mn},\omega)$$

since the second term is nonnegative. Thus

$$\underline{\lim}_{m,n\to\infty} \inf_t Z_{mn}(t,\omega) \geq \underline{\lim}_{m,n\to\infty} A_{mn}(t_{mn},\omega) = A(t_0,\omega) \tag{5}$$

using (1) and the convergence $t_{mn} \to t_0$. Putting (4) and (5) together yields (2).

Proof of Theorem 4. This proof is very similar to the proof of Theorem 3. Some minor inconveniences are caused by the contiguous sequences F^m and G^n, but Lemma A1 can be pushed through with the additional restrictions and an "in probability" version of Lemma A2 can be given based on adaptation of Theorem 10 of Ibragimov and Has'minskii (1980, p. 367).

Proof of Theorem 5. Let F and G be fixed distributions satisfying the condition of Theorem 5. All limits are as $\min(m,n) \to \infty$. For each of the three procedures if $e_F(T) \leq e_G(T)$, then by Theorem 1

$$\lim P(\text{Procedure } i \text{ makes an incorrect statement})$$

$$= \lim P(Z_{mn}(T) > z_\alpha) \leq \alpha .$$

Now assume $e_F(T) > e_G(T)$. For Procedure 1 define θ_1 =

$\inf\{t \geq T: \ e_F(t) \leq e_G(t)\}$. If $\theta_1 = \infty$ then P(Procedure 1 makes an incorrect statement) = 0. For the case $\theta_1 < \infty$ recall that mean residual life functions are right continuous. Thus, $e_F(t) - e_G(t)$ is also right continuous and $e_F(\theta_1) \leq e_G(\theta_1)$. Procedure 1 will make an incorrect statement only if $\hat{\theta}_1 > \theta_1$. Thus, by Theorem 1,

$$\lim P(\text{Procedure 1 makes an incorrect statement})$$

$$\leq \lim P(\hat{\theta}_1 > \theta_1) \leq \lim P(Z_{mn}(\theta_1) > z_\alpha) \leq \alpha \ .$$

For Procedure 2 and 3, define

$$\theta_2 = \sup\{t \leq T: \ e_F(t) \geq e_G(t)\}$$

and

$$\delta = \sup\{d \geq 0: \ \inf_{t \, \varepsilon [\max(0,T-d),T+d]} \{e_F(t) - e_G(t)\} > 0\} \ .$$

The proof of the validity of these procedures follows as for Procedure 1 except that the continuity of F is used to ensure that $e_F(\theta_2) \leq e_G(\theta_2)$ or $e_F(T - \delta) \leq e_G(T - \delta)$ if

$$\lim_{t \uparrow T-\delta} \{e_F(t) - e_G(t)\} \leq 0.$$

The continuity of F ensures the continuity of e_F. Also e_G, like any mean residual life function, is upper semi-continuous so that we have

$$\lim_{t \uparrow t^*} \{e_F(t) - e_G(t)\} \geq e_F(t^*) - e_G(t^*) \ .$$

REFERENCES

Berger, R. L. (1984). Testing whether one regression function is larger than another. *Comm. Statist.* A13, 1793-1810.

Berger, R. L. and Sinclair, D. F. (1984). Testing hypotheses concerning unions of linear subspaces. *J. Am. Statist. Assoc.* 79, 158-163.

Billingsley, P. (1968). *Convergence of Probability Measures.* New York: John Wiley.

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

Cheng, K. F. (1985). Tests for the equality of failure rates. *Biometrika* 72, 211-215.

Chiang, C. L. (1960). A stochastic study of the life table and its applications: II. Sample variance of the observed expectation of life and other biometric functions. *Human Biology* 32, 221-238.

Chikkagoudar, M. S. and Shuster, J. S. (1974). Comparison of failure rates using rank tests. *J. Am. Statist. Assoc.* 69, 411-413.

Cox, D. R. (1962). *Renewal Theory.* London: Meuthen.

Elandt-Johnson, R. C. and Johnson, N. L. (1980). *Survival Models and Data Analysis.* New York: Wiley.

Gleser, L. J. (1973). On a theory of intersection-union tests. *Inst. Math. Statist. Bull.* 2, 233. (Abstract)

Guess, F. and Proschan, F. (1985). Mean residual life: theory and applications. To appear in *Handbook of Statistics*, Volume 7, *Quality Control and Reliability.* Ed. P. R. Krishnaiah. Amsterdam: North Holland.

Hall, W. J. and Wellner, J. A. (1979). Estimation of mean residual life. University of Rochester Department of Statistics Technical Report.

Ibragimov, I. A. and Has'minskii, R. Z. (1981). *Statistical Estimation: Asympototic Theory.* New York: Springer-Verlag.

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

Kochar, S. C. (1979). Distribution-free comparison of two probability distributions with reference to their hazard rates. *Biometrika* 66, 437-441.

Kochar, S. C. (1981). A new distribution-free test for the equality of two failure rates. *Biometrika* 68, 423-426.

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

Stablein, D. M. and Koutrouvelis, I. A. (1985). A two-sample test sensitive to crossing hazards in uncensored and singly censored data. *Biometrics* 41, 643-652.

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

Witten, M. (1985). Reliability theoretic methods and aging: critical elements, hierarchies, and longevity - Interpreting survival curves. In *Molecular Biology of Aging*, Ed. A. D. Woodland, A. D. Plackett, and A. Hollaender, pp. 345-360. New York: Plenum Press.

Yu, B. P., Marsoro, E. J., Murata, I., Bertrand, H. A. and Lynd, F. T. (1982). Lifespan study of SPF Fisher 344 male rats fed ad libitum or restricted diets: longevity, growth, lean body mass and disease. *J. Gerontol.* 37, 130-141.