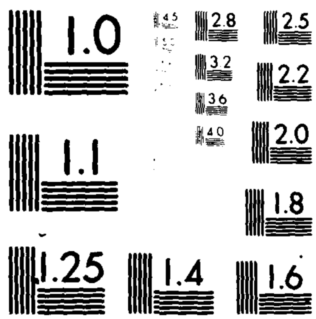
AD-A089 687 FLORIDA STATE UNIV TALLAHASSEE DEPT OF STATISTICS F/G 12/1
TESTING WHETHER NEW IS BETTER THAN USED WITH RANDOMLY CENSORED --ETC(U)
JUN 80 Y CHEN, M HOLLANDER, N A LANGBERG AFOSR-78-3678
UNCLASSIFIED FSU-STATISTICS-M503 AFOSR-TR-80-0943 NL

1 OF 1

END
DATE FILMED
10-80
DTIC



MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS 1963-A

UNCLASSIFIED

LEVEL II

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE	READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER AFOSR-TR-80-0943 / 2. GOVT ACCESSION NO. AD-A089 687	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) TESTING WHETHER NEW IS BETTER THAN USED WITH RANDOMLY CENSORED DATA	5. TYPE OF REPORT & PERIOD COVERED Interim rept.
	6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) Yuan-Yan Chen, Naftali Langberg, Myles Hollander	8. CONTRACT OR GRANT NUMBER(s) AFOSR-78-3678
9. PERFORMING ORGANIZATION NAME AND ADDRESS Florida State University, Department of Statistics, Tallahassee, FL 32306	10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS 61102F 2304/A5
11. CONTROLLING OFFICE NAME AND ADDRESS Air Force Office of Scientific Research/NM, Bolling AFB, Washington, DC 20332	12. REPORT DATE June 1980
	13. NUMBER OF PAGES 25
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office)	15. SECURITY CLASS. (of this report) UNCLASSIFIED
	15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

FSU-STATISTICS-M543

DTIC ELECTE SEP 30 1980 S D E

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

TR-9-10-AF

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

New better than used, exponentiality, hypothesis test, consored data

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

See back for abstract.

AD A089687

DDC FILE COPY

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

UNCLASSIFIED 400277

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE(When Data Entered)

A life distribution F is new better than used (NBU) if $\bar{F}(x + y) \le \bar{F}(x)\bar{F}(y)$ for all $x, y \ge 0$ ($\bar{F} \equiv 1 - F$). Using a randomly censored sample of size n from F, we propose a test of H_0: F is exponential, versus H_1: F is NBU, but not exponential. Our test is based on the statistic $J_n^c = \iint \bar{F}_n(x + y)dF_n(x)dF_n(y)$, where F_n is the product limit estimator of F, introduced by Kaplan and Meier (1958).

Under mild regularity on the amount of censoring, the asymptotic normality of J_n^c, suitably normalized, is established. Then using a consistent estimator of the null standard deviation of $n^{1/2}J_n^c$, an asymptotically distribution-free test is obtained. Finally, using tests for the censored and uncensored models we develop a measure of the efficiency loss due to the presence of censoring.

Accession For	
NTIS GRA&I	☒
DDC TAB	☐
Unannounced	☐
Justification	
By	
Distribution/	
Availability Codes	
Dist.	Avail and/or special
A	

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE(When Data Entered)

AFOSR-TR- 80-0943

Testing Whether New is Better Than Used With Randomly Censored Data

by

Yuan-Yan Chen[1], Myles Hollander[1], and Naftali A. Langberg[1]

FSU Statistics Report No. M503
AFOSR Technical Report No. 78-12

June 1980

The Florida State University
Department of Statistics
Tallahassee, Florida 32306

[1]Research supported by the Air Force Office of Scientific Research, AFSC, USAF, under Grant AFOSR 78-3678.

Testing Whether New is Better Than Used
With Randomly Censored Data

by

Yuan-Yan Chen, Myles Hollander, and Naftali A. Langberg

Abstract.

A life distribution F is new better than used (NBU) if $\bar{F}(x+y) \leq \bar{F}(x)\bar{F}(y)$ for all $x, y \geq 0$ ($\bar{F} \equiv 1 - F$). Using a randomly censored sample of size n from F, we propose a test of H_0: F is exponential, versus H_1: F is NBU, but not exponential. Our test is based on the statistic $J_n^c = \iint \bar{F}_n(x+y)dF_n(x)dF_n(y)$, where F_n is the product limit estimator of F, introduced by Kaplan and Meier (1958).

Under mild regularity on the amount of censoring, the asymptotic normality of J_n^c, suitably normalized, is established. Then using a consistent estimator of the null standard deviation of $n^{1/2}J_n^c$, an asymptotically distribution-free test is obtained. Finally, using tests for the censored and uncensored models we develop a measure of the efficiency loss due to the presence of censoring.

Key words: New better than used, exponentiality, hypothesis test, censored data.

1. Introduction and Summary. A life distribution F (a distribution function (d.f.) such that $F(x) = 0$ for $x < 0$), with survival function $\boxed{\bar{F} \equiv 1 - F}$, the superior limit of F is approximately 1 - F

is new better than used (NBU), ~~if~~

$$(1.1) \qquad \bar{F}(x+y) \leq \bar{F}(x)\bar{F}(y) \text{ for } x, y \in [0, \infty).$$

The dual notion of a new worse than used (NWU) life d.f. is defined by reversing the inequality, ~~in (1.1).~~ The boundary members of the NBU and NWU classes, obtained by insisting an equality, ~~in (1.1)~~ are the exponential d.f.'s.

The NBU class of life distributions has proved to be very useful in performing analyses of lifelengths. These d.f.'s provide readily interpretable models for describing wearout, play a fundamental role in studies of replacement policies (Marshall, Proschan, 1972), and shock models (Esary, Marshall, and Proschan, 1973), and have desirable closure properties (c.f. Barlow and Proschan, 1975).

Hollander and Proschan (HP) (1972) develop a test of

$$(1.2) \quad H_0: \; F(x) = 1 - \exp\{-x/\mu\}, \; x \geq 0, \; \mu > 0 \; (\mu \text{ unspecified}),$$

versus

$$(1.3) \qquad H_1: \; F \text{ is NBU, but not exponential},$$

based on a random sample $X_1, \ldots, X_n$ from a continuous life distribution F. The hypothesis H_0 asserts that a new item has stochastically the same lifelength as a used item of any age, where the alternative H_1 states that a new item has stochastically greater lifelength than a used item of any age. The HP (1972) test is motivated by considering the parameter

$$\gamma(F) = \int_0^\infty \int_0^\infty [\overline{F}(x)\overline{F}(y) - \overline{F}(x+y)]dF(x)dF(y) = \tag{1.4}$$

$$= 1/4 - \int_0^\infty \int_0^\infty \overline{F}(x+y)dF(x)dF(y) \stackrel{\text{def}}{=} 1/4 - \Delta(F).$$

Viewing $\gamma(F)$ as a measure of the deviation of F from exponentiality towards NBU [or NWU] alternatives, HP (1972) replace F by G_n, the empirical d.f. of $X_1, \ldots, X_n$, and suggest rejecting H_0 in favor of H_1 if $\Delta(G_n)$ is too small [H_0 is rejected, in favor of H_1': F is NWU, but not exponential, if $\Delta(G_n)$ is too large.] For further details about the test see HP (1972), Hollander and Wolfe (1973), Cox and Hinkley (1974), and Randles and Wolfe (1979).

In this paper we consider a randomly censored model where we do not get to observe a complete sample of X's. Let $X_1, X_2, \ldots$ be independent identically distributed (i.i.d.) random variables (r.v.'s) having a common continuous life d.f. F. The X's represent lifelengths of identical items. Let $Y_1, Y_2, \ldots$ be i.i.d. r.v.'s having a common continuous d.f. H. The Y's represent the random times to right-censorship. Throughout we assume the X's and Y's are mutually independent and the pairs $(X_1, Y_1), (X_2, Y_2), \ldots$ are defined on a common probability space $(\Omega, \mathcal{B}, P)$. Further, let $I(A)$ denote the indicator function of the set A, and for $i = 1, \ldots, n$, let $Z_i = \min\{X_i, Y_i\}$, and $\delta_i = I(X_i \leq Y_i)$. Based on the incomplete data set $(Z_1, \delta_1), \ldots, (Z_n, \delta_n)$ we test H_0, given by (1.2), against H_1, given by (1.3). The censoring d.f. H is assumed to be unknown and is treated as a nuisance parameter. Due to the censoring, the empirical d.f. G_n corresponding to F cannot be computed. Thus, we propose to reject H_0 in favor of H_1 for small values of

$$\Delta(F_n) \stackrel{\text{def}}{=} J_n^c = \int_0^\infty \int_0^\infty \overline{F}_n(x+y)dF_n(x)dF_n(y), \tag{1.5}$$

where F_n is the Product Limit Estimator (PLE) of F, introduced by Kaplan and Meier (1958):

(1.6) $$\bar{F}_n(x) \overset{\text{def}}{=} 1 - F_n(x) = \prod_{\{i:\ Z_{(i)} \le x\}} [(n-i)(n-i+1)^{-1}]^{\delta_{(i)}},$$

where $Z_{(1)} < \ldots < Z_{(n)}$ denote the ordered Z's, $Z_{(0)} = 0$, $\delta_{(0)} = 1$, and $\delta_{(1)}, \ldots, \delta_{(n)}$ are the δ's corresponding to $Z_{(1)}, \ldots, Z_{(n)}$ respectively. In (1.6), we treat $Z_{(n)}$ as a death (whether or not it actually is) so that $\delta_{(n)} = 1$. Furthermore, although our assumptions preclude the possibility of ties, in practice ties will occur. When censored observations are tied with uncensored observations, our convention, when forming the list of the ordered Z's, is to treat uncensored members of the tie as preceding the censored members of the tie.

For computational purposes, it is convenient to write J_n^c as:

$$J_n^c = \Sigma_{i=1}^n \Sigma_{j=1}^n \{[\prod_{\{k:\ Z_{(k)} \le Z_{(i)} + Z_{(j)}\}} \{(n-k)(n-k+1)^{-1}\}^{\delta_{(k)}}]$$

$$[\prod_{r=1}^{i-1} \{(n-r)(n-r+1)^{-1}\}^{\delta_{(r)}}][\prod_{s=1}^{j-1} \{(n-s)(n-s+1)^{-1}\}^{\delta_{(s)}}]$$

$$(n-i+1)^{-1}(n-j+1)^{-1}\delta_{(i)}\delta_{(j)}\}.$$

In clinical trials the X's may be times, measured from date of diagnosis of a disease, to relapse. NBU alternatives may be preferable to increasing failure rate alternatives because the latter class insists upon a non-decreasing failure rate whereas, in this medical context, we may expect the failure rate to increase (at least for a short period of time) after treatment begins. Incomplete observations can arise at the time of data analysis due to, for example, dropout or patients who have not yet relapsed. In this situation, it is appropriate to use J_n^c to test H_0 vs. H_1.

Marshall and Proschan (1972) consider age replacement policies and block replacement policies. Under an age replacement policy, a unit is replaced upon failure or upon reaching a specified age T, whichever comes first. Under a block replacement policy, a replacement is made whenever a failure occurs, and additionally at specified times T, 2T, 3T, ..., .

Marshall and Proschan show that a necessary and sufficient condition for failure-free intervals to be stochastically larger (smaller) under age replacement than under a policy of replacement at failure only is that the underlying distribution be NBU (NWU). Marshall and Proschan also show that a necessary and sufficient condition that the number of failures in a specified [0, t] be stochastically smaller (larger) under age replacement than under a policy of replacement at failure only is that the underlying distribution be NBU (NWU). Similar comparisons hold for block replacement. Thus in reaching a decision as to whether to use an age (block) replacement policy or not, it is important to investigate whether or not the underlying distribution is NBU. If lifelength times are censored, the test based on J_n^c facilitates such an investigation.

Other references describing situations where it is important to know whether the underly distribution is NBU are Esary, Marshall and Proschan (1973) in the context of shock models, and El-Neweihi, Proschan, and Sethuraman (1978) in the context of multiple coherent systems.

In Section 2 we establish the asymptotic normality of the sequence $n^{1/2}\{J_n^c - \Delta(F)\}$ under the assumptions:

(A.1) The supports of F and H are equal to $[0, \infty)$,

(A.2) $$\sup\{[\bar{F}(x)]^{1-\epsilon}[\bar{H}(x)]^{-1}, x\epsilon[0,\infty)\} < \infty$$

for some nonnegative real number ϵ.

and

(A.3) The processes $\{n^{1/2}\{F_n(t) - F(t)\}, t\epsilon(-\infty,\infty)\}$ converge weakly to a Gaussian process with mean zero and covariance kernel given by (2.1).

Condition (A.2) restricts the amount of censoring allowed in the model. To see this in a simple case, consider the proportional hazards model where $\bar{H} = [\bar{F}]^{\beta}$ for some $\beta > 0$. Then $P\{X_1 \leq Y_1\} = (\beta + 1)^{-1}$, and condition (A.2) implies that $\beta \leq 1$. Thus, in the proportional hazards model, the J_n^c test is inappropriate when the expected amount of censoring $P\{Y_1 < X_1\}$ exceeds 50%.

The null asymptotic mean of J_n^c is 1/4, independent of the nuisance parameters μ and H. However, the null asymptotic variance of $n^{1/2}J_n^c$ does depend on μ and H and must be estimated from the data. A consistent estimator, $\hat{\sigma}_n^2$, given by (3.3), is derived in Section 3. The approximate α-level test rejects H_0 is favor of H_1 if $n^{1/2}\{J_n^c - (1/4)\}\hat{\sigma}_n^{-1} \leq - z_{\alpha}$, where z_{α} is the upper α-percentile of a standard normal distribution. In Section 3 we also show that this asymptotically distribution-free test is, under suitable regularity, consistent against all continuous NBU alternatives.

Section 4 develops a measure of the loss in efficiency due to the presence of censoring. This measure is derived using the HP (1972) NBU test and its generalization herein proposed based on J_n^c. In certain instances, this measure assumes values close to $P(X_1 < Y_1)$; the latter also being a (rough) measure of the loss of information due to censoring.

Section 5 contains an application of the J_n^c statistic to some survival data.

2. Asymptotic Normality of the NBU Test Statistic. In this section we establish the asymptotic normality of the test statistic J_n^c, defined by (1.5).

Let $\bar{K}(t) = \bar{F}(t)\bar{H}(t)$, $t\epsilon(-\infty,\infty)$, and let $\{\phi(t), t\epsilon(-\infty,\infty)\}$ be a Gaussian process with mean zero and covariance kernel given by:

$$(2.1)\quad E\phi(t)\phi(s) = \begin{cases} \overline{F}(t)\overline{F}(s)\int_0^s [\overline{K}(z)\overline{F}(z)]^{-1} dF(z), & 0 \le s \le t < \infty, \\ 0 & , s < 0 \text{ or } t < 0. \end{cases}$$

Unless otherwise specified, all limits are evaluated as $n \to \infty$, and all integrals range over $(-\infty, \infty)$.

First we state the main result of this section.

Theorem 2.1. Assume that conditions (A.1), (A.2), and (A.3), given in Section 1, hold. Then $n^{1/2}\{J_n^c - \Delta(F)\}$ converges in distribution to a normal r.v. with mean zero and variance σ^2, given by:

$$(2.2)\quad \sigma^2 = \iiiint E\{[\phi(t+s) - 2\phi(t-s)][\phi(u+v) - 2\phi(u-v)]\} dF(s)dF(t)dF(u)dF(v).$$

Note that for $n = 1, 2, \ldots$,

$$\begin{aligned} J_n^c - \Delta(F) = {} & \{\iint[\overline{F}_n(x+y) - \overline{F}(x+y)]dF_n(x)dF_n(y) \\ & - \iint[\overline{F}_n(x+y) - \overline{F}(x+y)]dF_n(x)dF(y)\} \\ & + \{\iint[\overline{F}_n(x+y) - \overline{F}(x+y)]dF_n(x)dF(y) \\ & - \iint[\overline{F}_n(x+y) - \overline{F}(x+y)]dF(x)dF(y)\} \\ & + \iint[\overline{F}_n(x+y) - \overline{F}(x+y)]dF(x)dF(y) \\ & + \iint\overline{F}(x+y)dF_n(x)dF_n(y) - \iint\overline{F}(x+y)dF(x)dF_n(y) \\ & + \iint\overline{F}(x+y)dF(x)dF_n(y) - \iint\overline{F}(x+y)dF(x)dF(y). \end{aligned}$$

Upon integration by parts and change of variable we obtain that:

$$\begin{aligned} & \iint\overline{F}(x+y)dF_n(x)dF_n(y) - \iint\overline{F}(x+y)dF(x)dF_n(y) \\ & = - \iint[\overline{F}_n(x-y) - \overline{F}(x-y)]dF_n(x)dF(y), \quad n = 1, 2, \ldots, \end{aligned}$$

and that

$$\iint \bar{F}(x + y)dF(x)dF_n(y) - \iint \bar{F}(x + y)dF(x)dF(y)$$
$$= -\iint [\bar{F}_n(x - y) - \bar{F}(x - y)]dF(x)dF(y), \; n = 1, 2, \ldots .$$

Thus for $n = 1, 2, \ldots,$

$$n^{1/2}\{J_n^c - \Delta(F)\} = B_{n,1} + B_{n,2} - B_{n,3} + B_{n,4},$$

where

$$B_{n,1} = \iint n^{1/2}[\bar{F}_n(x + y) - \bar{F}(x + y)]dF_n(x)dF_n(y) -$$
$$- \iint n^{1/2}[\bar{F}_n(x + y) - \bar{F}(x + y)]dF_n(x)dF(y),$$

$$B_{n,2} = \iint n^{1/2}[\bar{F}_n(x + y) - \bar{F}(x + y)]dF_n(x)dF(y) -$$
$$- \iint n^{1/2}[\bar{F}_n(x + y) - \bar{F}(x + y)]dF(x)dF(y),$$

$$B_{n,3} = \iint n^{1/2}[\bar{F}_n(x - y) - \bar{F}(x - y)]dF_n(x)dF(y) -$$
$$- \iint n^{1/2}[\bar{F}_n(x - y) - \bar{F}(x - y)]dF(x)dF(y),$$

and

$$B_{n,4} = \iint n^{1/2}[\bar{F}_n(x + y) - \bar{F}(x + y) - 2\{\bar{F}_n(x - y) - \bar{F}(x - y)\}]dF(x)dF(y).$$

Consequently to prove the result of Theorem 2.1 it suffices, by Slutsky's Theorem [Billingsley (1968), p. 49], to show that $B_{n,1}$, $B_{n,2}$, and $B_{n,3}$ converge in probability to zero, and that $B_{n,4}$ converges in distribution to a normal r.v. with mean zero and variance σ^2, given by (2.2).

First we prove that $B_{n,1}$, $B_{n,2}$, and $B_{n,3}$ converge in probability to zero, and that $B_{n,4}$ converges in distribution. Then we prove that the limiting d.f. of $B_{n,4}$ is normal with mean zero and variance σ^2.

To establish the convergence of $B_{n,1}$ through $B_{n,4}$ we introduce a notation and four lemmas. Let D = {ψ: ψ is real valued, bounded, and right-continuous function defined on $(-\infty, \infty)$, with finite left-hand limits at each $t \epsilon (-\infty,\infty)$, and finite limits at $t = \pm \infty\}$. Throughout we view D as a metric space with the Skorohod metric [Billingsley (1968), p. 112].

Lemma 2.2. Let $\psi \epsilon$ D, and $y \epsilon (-\infty,\infty)$. Then $\int\psi(x + y)dF_n(x)$, and $\int\psi(x - y)dF_n(x)$ converge w.p.1, to $\int\psi(x + y)dF(x)$, and to $\int\psi(x - y)dF(x)$, respectively.

Proof. There is a set $\Omega_1 \epsilon \mathcal{B}$, $P\{\Omega_1\} = 1$, such that $F_n(x, \omega)$ converges to $F(x)$ for every $x \epsilon (\infty,\infty)$, and $\omega \epsilon \Omega_1$, [c.f. Peterson (1977), Th. 3.3, or Langberg, Proschan, and Quinzi (1980), Th. 4.9]. Note that the sets of discontinuities of the functions $\psi(\cdot + y)$, and $\psi(\cdot - y)$ are countable [Billingsley (1968), p. 110]. Since F is continuous these sets have F-measure zero. Consequently the desired results follow by the Helly-Bray Lemma [Breiman (1968), p. 163, Th. 8.12]. ||

Lemma 2.3. Let $\psi \epsilon$ D. Then $\int\int\psi(x + y)dF_n(x)dF(y)$, and $\int\int\psi(x - y)dF_n(x)dF(y)$, converge w.p.1 to $\int\int\psi(x + y)dF(x)dF(y)$, and $\int\int\psi(x - y)dF(x)dF(y)$, respectively.

Proof. Note that $\int\psi(x + \cdot)dF_n(x)$, and $\int\psi(x - \cdot)dF_n(x)$ are sequences of bounded functions. By Lemma 2.2 these sequences converge w.p.1 to $\int\psi(x + \cdot)dF(x)$, and $\int\psi(x - \cdot)dF(x)$, respectively. Consequently the desired results follow by the Dominated Convergence Theorem. ||

Lemma 2.4. Let ψ be a continuous function in D. Then $\int\int\psi(x + y)dF_n(x)dF_n(y) - \int\int\psi(x + y)dF_n(x)dF(y)$, and $\int\int\psi(x - y)dF_n(x)dF_n(y) - \int\int\psi(x - y)dF_n(x)dF(y)$, converge w.p.1 to zero.

Proof. To prove the desired results it suffices, by Lemma 2.3, to show that $\int\int\psi(x + y)dF_n(x)dF_n(y)$, and $\int\int\psi(x - y)dF_n(x)dF_n(y)$, converge w.p.1. to $\int\int\psi(x + y)dF(x)dF(y)$, and to $\int\int\psi(x - y)dF(x)dF(y)$, respectively.

We now prove the preceding two statements. There is a set $\Omega_1 \epsilon \mathcal{B}$, $P\{\Omega_1\} = 1$, such that $F_n(x, \omega)$ converges to $F(x)$ for all $x \epsilon (-\infty, \infty)$, and $\omega \epsilon \Omega_1$. Consequently the preceding two statements follow by the Helly-Bray Lemma [Billingsley (1968), p. 11, Th. 2.1 (ii)]. ||

Lemma 2.5. Assume (A.1) holds. Then the Gaussian process $\{\phi(t), t \epsilon (-\infty,\infty)\}$, with mean zero and covariance kernel given by (2.1), has continuous paths w.p.1.

Proof. Let $g(t) = \int_0^t [\bar{K}(z)\bar{F}(z)]^{-1} dF(z)$, $t \epsilon [0,\infty)$. By (A.1) and the continuity of F, g is strictly increasing and continuous. Let $\phi_1(t)$ be a stochastic process given by:

$$\phi_1(t) = \begin{cases} [\bar{F}(g^{-1}(t))]^{-1}\phi(g^{-1}(t)), & t \epsilon [0,\infty), \\ 0 & , t \epsilon (-\infty, 0). \end{cases}$$

Clearly $\{\phi_1(t), t \epsilon (-\infty,\infty)\}$ is a Gaussian process with mean zero, and covariance kernel given by:

$$E\phi_1(t)\phi_1(s) = \begin{cases} s, & 0 \leq s \leq t < \infty \\ 0, & s < 0 \text{ or } t < 0. \end{cases}$$

Thus $\{\phi_1(t), t \epsilon [0,\infty)\}$ is a standard Wiener process.

Note that g^{-1} is continuous and strictly increasing to ∞, $g^{-1}(0) = 0$, and that $\{\phi_1(t), t \epsilon (-\infty,\infty)\}$ has continuous paths w.p.1 [Breiman (1968), p. 257]. Consequently the desired result follows from the definition of $\{\phi_1(t), t \epsilon (-\infty,\infty)\}$. ||

Breslow and Crowley (1974) and Peterson (1977) prove that the processes $\{n^{1/2}(\bar{F}_n(t) - \bar{F}(t)), t \epsilon (-\infty, T)\}$ converge weakly to the process $\{\phi(t), t \epsilon (-\infty, T)\}$

for all $T \epsilon (-\infty, \infty)$, provided (A.1) holds. To prove that $B_{n,1}$, $B_{n,2}$, and $B_{n,3}$ converge in probability to zero, and that $B_{n,4}$ converges in distribution to a normal r.v. with mean zero and variance σ^2, we must *assume* that Breslow and Crowley's result holds for $T = \infty$. Since (A.2) restricts the amount of censoring allowed in the model we *conjecture* that under (A.1), (A.2), $\{n^{1/2}\{\overline{F}_n(t) - \overline{F}(t)\}, t \epsilon (-\infty, \infty)\}$ converges weakly to $\{\phi(t), t \epsilon (-\infty, \infty)\}$. We assume this conjecture to prove Lemmas 2.6, 2.7 and 2.9 that follow.

Now we proceed to prove that $B_{n,1}$, $B_{n,2}$, and $B_{n,3}$ converge in probability to zero, and that $B_{n,4}$ converges in distribution. A lemma is needed.

Lemma 2.6. Assume (A.3) holds. Then the processes $\{n^{1/2}\{\overline{F}_n(x + y) - \overline{F}(x + y) - 2[\overline{F}_n(x - y) - \overline{F}(x - y)]\}, x, y \epsilon (-\infty, \infty)\}$, converge weakly to the process $\{\phi(x + y) - 2\psi(x - y), x, y \epsilon (-\infty, \infty)\}$.

Proof. Let $D^2 = \{< \psi_1, \psi_2 >, \psi_1, \psi_2 \in D\}$, be a mertic space with the metric induced by the one of D. By a standard argument the bivariate processes $\{n^{1/2} < \overline{F}_n(t) - \overline{F}(t), \overline{F}_n(s) - \overline{F}(s) >, t, s \epsilon (-\infty, \infty)\}$ converge weakly to the bivariate process $\{< \phi(t), \phi(s) >, t, s \epsilon (-\infty, \infty)\}$. Thus, by the Continuous Mapping Theorem (Billingsley (1968), P. 30, Th. 5.1) the processes $\{n^{1/2}\{\overline{F}_n(t) - \overline{F}(t) - 2[\overline{F}_n(s) - \overline{F}(s)]\}, t, s \epsilon (-\infty, \infty)\}$ converge weakly to the process $\{\phi(t) - 2\phi(s), t, s \epsilon (-\infty, \infty)\}$. Consequently the desired result follows. ||

We now establish the convergence of $B_{n,1}$ through $B_{n,4}$. Some notation is useful. Let Q^1, Q^2, Q_n^1, Q_n^2, be the probability measures on D induced by the processes $\{\phi(t), t \epsilon (-\infty, \infty)\}$, $\{\phi(x + y) - 2\phi(x - y), x, y \epsilon (-\infty, \infty)\}$,

$\{n^{1/2}\{\bar{F}_n(t) - \bar{F}(t)\}, t\epsilon(-\infty,\infty)\}$, and $\{n^{1/2}\{\bar{F}_n(x + y) - \bar{F}(x + y) - 2[\bar{F}_n(x - y) - \bar{F}(x - y)]\}, x, y\epsilon(-\infty, \infty)\}$, $n = 1, 2, \ldots$, respectively. Let S_1, S_2 be two sets, let A be a subset of S_2 and let ξ be a mapping from S_1 to S_2; then $\xi^{-1}(A) = \{s: \; s \in S_1, \xi(s) \in A\}$.

Lemma 2.7. Assume (A.1), (A.2), and (A.3) hold. Then:

(a) $B_{n,1}$, B_{n2}, and $B_{n,3}$ converge in probability to zero,

and

(b) $B_{n,4}$ converges in distribution to the r.v. $\iint[\phi(x + y) - 2\phi(x - y)]dF(x)dF(y)$.

Proof. For $\psi \in$ D, and $n = 1, 2, \ldots$ let

$$\xi_{n,1}(\psi) = \iint\psi(x + y)dF_n(x)dF_n(y) - \iint\psi(x + y)dF_n(x)dF(y),$$

$$\xi_{n,2}(\psi) = \iint\psi(x + y)dF_n(x)dF(y) - \iint\psi(x + y)dF(x)dF(y),$$

$$\xi_{n,3}(\psi) = \iint\psi(x - y)dF_n(x)dF(y) - \iint\psi(x - y)dF(x)dF(y),$$

and

$$\xi_4(\psi) = \iint[\psi(x + y) - 2\psi(x - y)]dF(x)dF(y).$$

The probabilities Q_n^1 converge weakly to Q^1. By Lemma 2.6 Q_n^2 converges weakly to Q^2. By Lemma 2.5 the supports of Q^1 and Q^2 coincide with the set of all continuous functions in D. By the definitions of the mappings and probability measures:

$$Q_n^1\xi_{n,q}^{-1}\{(-\infty, x]\} = P\{B_{n,q} \leq x\}, \; x\epsilon(-\infty,\infty), \; q = 1, 2, 3, \; n, = 1, 2, \ldots,$$

$$Q_n^2\xi^{-1}\{(-\infty.x]\} = P\{B_{n,4} \leq x\}, \; x\epsilon(-\infty,\infty), \; n = 1, 2, \ldots$$

and

$$Q^1\xi^{-1}\{(-\infty, x]\} = P\{\iint[\phi(u + v) - 2\phi(u - v)]dF(u)dF(v) \leq x\}, \; x\epsilon(-\infty,\infty).$$

Thus to obtain the desired results it suffices to show, by the Extended Continuous Mapping Theorem [Billingsley (1968), p. 34, Th. 5.5], that for every sequence $\psi_n \epsilon D$ that converges to a continuous function $\psi \epsilon D$, $\lim \xi_{n,q}(\psi_n) = 0$, w.p.1 for $q = 1, 2, 3$, and $\lim \xi(\psi_n) = \xi(\psi)$.

We now prove the preceding statements. Let $\psi_n \epsilon D$, $n = 1, 2, \ldots$, and let ψ be a continuous function in D. Assume $\lim \psi_n = \psi$. By a well-known result [Billingsley (1968), p. 112]:

$$\lim \sup\{|\psi_n(x) - \psi(x)|, x\epsilon(-\infty,\infty)\} = 0.$$

By Lemma 2.4 $\lim \xi_{n,1}(\psi) = 0$, w.p.1. By Lemma 2.3 $\lim \xi_{n,q}(\psi) = 0$, $q = 2, 3$. Consequently by simple integral evaluations we obtain that $\lim \xi_{n,q}(\psi_n) = 0$, w.p.1. for $q = 1, 2, 3$, and that $\lim \xi(\psi_n) = \xi(\psi)$. ||

We are ready to show that the limiting d.f. of $B_{n,4}$ is normal with mean zero and variance σ^2, given by (2.2). First, we show that under (A.2), $\sigma^2 < \infty$.

Lemma 2.8. Assume (A.2) holds. Then σ^2, given by (2.2), is finite.

Proof. Note that for $a, b\epsilon(-\infty,\infty)$, $(a - b)^2 \leq 2(a^2 + b^2)$. Thus, by the Cauchy-Schwartz Inequality:

$$\sigma^2 \leq \iint E[\phi(s + t) - 2\phi(s - t)]^2 dF(t)dF(s) \leq$$

$$\leq 2\iint[E\{\phi(s + t)\}^2 + 4E\{\phi(s - t)\}^2]dF(s)dF(t)$$

$$\leq 10 \sup\{E\{\phi(t)\}^2, t\epsilon[0,\infty)\}.$$

Hence to prove the desired result it suffices to show that $\sup\{[\bar{F}(t)]^2\int_0^t [\bar{K}(z)\bar{F}(z)]^{-1}dF(z), t\epsilon[0,\infty)\} < \infty$. By (A.2), $[\bar{H}(z)]^{-1} \leq c[\bar{F}(z)]^{\epsilon-1}$ for all $z\epsilon[0,\infty)$, some $c\epsilon(0,\infty)$, and some nonnegative real number ϵ. Consequently

$$[\bar{F}(t)]^2\int_0^t[\bar{K}(z)\bar{F}(z)]^{-1}d\bar{F}(z) \le c[\bar{F}(t)]^2\int_0^t[F(z)]^{\epsilon-3}dF(z),\ t\epsilon[0,\infty).$$

Since F is continuous

$$[\bar{F}(t)]^2\int_0^t[\bar{F}(z)]^{\epsilon-3}dF(z) = \begin{cases}(\epsilon - 2)^{-1}\{[\bar{F}(t)]^2 - [\bar{F}(t)]^{\epsilon}\}, & \epsilon \neq 2, \\ [\bar{F}(t)]^2 - [\bar{F}(t)]^2\ln\bar{F}(t) & , \epsilon = 2,\end{cases} \quad \text{for } t\epsilon[0,\infty).$$

The desired result follows now by simple limiting arguments. ||

Finally, we identify the limiting d.f. of $B_{n,4}$.

Lemma 2.9. Assume (A.1),(A.2), and (A.3) hold. Then $B_{n,4}$ converges in distribution to a normal r.v. with mean 0 and variance σ^2, given by (2.2).

Proof. By Lemma 2.7 (b) it suffices to show that the r.v. $\int\int[\phi(x + y) - 2\phi(x - y)]dF(x)dF(y)$ is normal with mean zero and variance σ^2. Since the process $\{\phi(x + y) -2\phi(x - y),\ x,\ y\epsilon(-\infty,\infty)\}$ is Gaussian the desired result follows by the theory of stochastic integration [c.f. Parzen (1962), p. 78]. ||

3. Consistency. Let $f(z) = z^3[1 + 4\ln z + 4(\ln z)^2]/16$, $0 < z \le 1$, and $= 0$ for $z = 0$, let $\mu = \int xdF(x)$, and let $\eta = [P\{X_1 \le Y_1\}]^{-1}EZ_1$. Further, let $\hat{\mu}_n = [\Sigma_{i=1}^n\delta_i]^{-1}\Sigma_{i=1}^nZ_i$, and $\bar{K}_n(t) = n^{-1}\Sigma_{i=1}^nI(Z_i > t)$, $n = 1, 2, \ldots$, $t\epsilon(-\infty,\infty)$. Finally, let

$$\sigma^2(\theta) = \int_0^1 f(z)[\bar{K}(-\theta\ln z)]^{-1}dz,\ \theta\epsilon(0,\infty), \tag{3.1}$$

$$\sigma_n^2(\theta) = \int_0^1 f(z)[\bar{K}_n(-\theta\ln z)]^{-1}I(-\ln z < \theta^{-1}Z_{(n)})dz,\ \theta\epsilon(0,\infty),\ n = 1, 2, \ldots, \tag{3.2}$$

and

$$\hat{\sigma}_n^2 = \hat{\sigma}_n^2(\hat{\mu}_n),\ n = 1, 2, \ldots . \tag{3.3}$$

For computational purposes $\hat{\sigma}_n^2$ can be written as

$$\hat{\sigma}_n^2 = (128)^{-1} + \Sigma_{i=1}^{n-1} n(n - i + 1)^{-1}(n - i)^{-1}[(128)^{-1} - (32)^{-1}Z_{(i)}(\hat{\mu}_n)^{-1} + (16)^{-1}Z_{(i)}^2(\hat{\mu}_n)^{-2}]\exp\{-4Z_{(i)}(\hat{\mu}_n)^{-1}\}] - n[(128)^{-1} - (32)^{-1}Z_{(n)}(\hat{\mu}_n)^{-1} + (16)^{-1}Z_{(n)}^2(\hat{\mu}_n)^{-2}] \exp\{-4Z_{(n)}(\hat{\mu}_n)^{-1}\}.$$

In this section we show that, under H_0, $\hat{\sigma}_n^2$ is a consistent estimator of σ^2 provided $\sigma^2(\theta)$ is finite in a neighborhood of μ. We then show, under (A.1), (A.2) and (A.3), and the assumptions: $\mu < \infty$ and $\sigma^2(\theta) < \infty$ in a neighborhood of η, that the approximate α-level test, which rejects H_0 in favor of H_1 if $n^{1/2}\{J_n^c - (1/4)\}\hat{\sigma}_n^{-1} < - z_\alpha$, is consistent against all continuous NBU alternatives. We conclude the section by presenting a sufficient condition for $\sigma^2(\theta)$ to be finite at $\theta\epsilon(0,\infty)$.

Now we show the consistency of $\hat{\sigma}_n^2$ under H_0. The proof of consistency uses several lemmas. We first show that $\hat{\sigma}_n^2(\theta)$ converges in probability to $\sigma^2(\theta)$, provided $\sigma^2(\theta) < \infty$. Then we show that $\hat{\sigma}_n^2$ converges in probability to $\sigma^2(\eta)$, provided $\sigma^2(\theta)$ is finite in an interval containing η. Finally, using the previous results, we obtain the consistency of $\hat{\sigma}_n^2$ under H_0.

To show that $\hat{\sigma}_n^2(\theta)$ converges to $\sigma^2(\theta)$ we need a well-known proposition, stated for the sake of completeness, and a lemma.

<u>Proposition 3.1</u>. [David (1970), p. 18] Let $U_{(1)} < \ldots < U_{(n)}$ be the order statistics of a sample of size n, n = 2, 3, ..., taken from a continuous d.f. G. Further let $u\epsilon(\inf\{s\colon\ G(s) > 0\}, \infty)$, and $G_u(t) = G(t)[G(u)]^{-1}$, $t\epsilon(-\infty,\ )$. Then the conditional random vector $\{(U_{(1)}, \ldots, U_{(n-1)})|U_{(n)} = u\}$, is stochastically equal to the order statistics of a sample of size n - 1 taken from the *d.f.* G_u.

<u>Lemma 3.2</u>. Let $\theta\epsilon(0,\infty)$, and $\delta\epsilon[0,1)$. Then

$$E\int_0^\delta f(z)[\overline{K}_n(-\theta\ell nz)]^{-1}I(-\ell nz < \theta^{-1}Z_{(n)})dz \le$$

$$\le 3\int_0^\delta f(z)[\overline{K}(-\theta\ell nz)]^{-1}dz, \; n = 1, 2, \ldots .$$

<u>Proof</u>. Note that for $n = 1, 2, \ldots$,

$$E\int_0^\delta f(z)[\overline{K}_n(-\theta\ell nz)]^{-1}I(-\ell nz \le \theta^{-1}Z_{(n)})dz =$$

$$= \int_0^\delta f(z)\{\int_{-\theta\ell nz}^\infty E\{[\overline{K}_n(-\theta\ell nz)]^{-1}|Z_{(n)} = u\}dP\{Z_{(n)} \le u\}\}dz.$$

Now let $z\epsilon(0,1)$, $u\epsilon(-\theta\ell nz,\infty)$, $q(u, z) = [K(u)]^{-1}K(-\theta\ell nz)$, and $p(u, z) = 1 - q(u, z)$. By Proposition 3.1 the conditional r.v. $\{[\overline{K}_n(-\theta\ell nz)]^{-1}|Z_{(n)} = u\}$ is stochastically equal to $n[B(n - 1, p(u, z)) + 1]^{-1}$, where $B(n - 1, p(u, z))$ is a binomial r.v. with parameters $n - 1$, and $p(u, z)$. Thus

$$E\{[\overline{K}_n(-\theta\ell nz)]^{-1}|Z_{(n)} = u\} =$$

$$= n\,\Sigma_{j=0}^{n-1}(j + 1)^{-1}\binom{n-1}{j}[p(u, z)]^j[q(u, z)]^{n-1-j}$$

$$= [p(u, z)]^{-1}(1 - [q(u, z)]^n).$$

Let $m(n) = n$ for $n = 2, 4, 6, \ldots$, and $= n + 1$ for $n = 1, 3, 5, \ldots$.
Then

$$[p(u, z)]^{-1}(1 - [q(u, z)]^n) < [p(u, z)]^{-1}(1 - [q(u, z)]^{m(n)})$$

$$\le 2[p(u, z)]^{-1}(1 - [q(u, z)]^{m(n)/2}) = 2\Sigma_{i=0}^{m(n)/2}[q(u, r)]^i, \; n = 1, 2, \ldots .$$

By direct evaluation we obtain that:

$$\int_{-\theta \ell n z}^{\infty} E\{[\overline{K}_n(-\theta \ell n z)]^{-1} | Z_{(n)} = u\} dP\{Z_{(n)} \leq u\}$$

$$\leq n\Sigma_{i=0}^{m(n)/2}(n - i)^{-1}[\overline{K}(-\theta \ell n z)]^{i} \leq 3[\overline{K}(-\theta \ell n z)]^{-1}.$$

Consequently the desired result follows. ||

We are ready to prove that $\hat{\sigma}_n^2(\theta)$ converges in probability to $\sigma^2(\theta)$.

Lemma 3.3. Let $\theta \epsilon (0,\infty)$. Assume $\sigma^2(\theta)$, given in (3.1), is finite. Then

$$p - \lim \hat{\sigma}_n^2(\theta) = \sigma^2(\theta).$$

Proof. Let λ, $\delta \epsilon (0,\infty)$. Then

$$P\{|\hat{\sigma}_n^2(\theta) - \sigma^2(\theta)| > \lambda\} \leq P\{\int_0^{\delta} f(z)[\overline{K}_n(-\theta \ell n z)]^{-1} I(-\ell n z < \theta^{-1} Z_{(n)}) dz > \lambda/3\}$$

$$+ P\{\int_0^{\delta} f(z)[\overline{K}(-\theta \ell n z)]^{-1} dz > \lambda/3\}$$

$$+ P\{\int_0^{\delta} f(z)|[\overline{K}_n(-\theta \ell n z)]^{-1} I(-\ell n z < \theta^{-1} Z_{(n)}) - [\overline{K}(-\theta \ell n z)]^{-1}| dz > \lambda/3\}.$$

By the Glivenko-Cantelli Lemma:

$$\lim \int_{\delta}^{1} (f(z)[\overline{K}_n(-\theta \ell n z)]^{-1} I(-\ell n z < \theta^{-1} Z_{(n)}) dz = \int_{\delta}^{1} f(z)[\overline{K}(-\theta \ell n z)]^{-1} dz, \text{ w.p.1.}$$

Thus by Lemma 3.2, and the Chebyshev Inequality:

$$\overline{\lim}\ P\{|\hat{\sigma}_n^2(\theta) - \sigma^2(\theta)| > \lambda\} \leq 6\lambda^{-1} \int_0^{\delta} f(z)[\overline{K}(-\theta \ell n z)]^{-1} dz.$$

Now since $\sigma^2(\theta) < \infty$

$$\lim_{\delta \to 0^+} \int_0^{\delta} f(z)[\overline{K}(-\theta \ell n z)]^{-1} dz = 0.$$

Consequently the desired result follows. ||

We now prove that $\hat{\sigma}_n^2$ converges in probability to $\sigma^2(\eta)$.

Lemma 3.4. Assume that $\eta < \infty$ and that $\sigma^2(\theta)$ is finite in an open interval that contains η. Then

$$p - \lim \hat{\sigma}_n^2 = \sigma^2(\eta).$$

Proof. Let $0 < \delta < \delta_0 < \eta$, $\lambda \epsilon (0,\infty)$, and let $A_n = \{\omega: \ |\eta_n(\omega) - \eta| > \delta\}$. Assume $\sigma^2(\theta) < \infty$ for $\theta \epsilon [\eta - \delta_0, \eta + \delta_0]$. By the monotonicity of $\hat{\sigma}_n^2(\theta)$ in θ for $n = 1, 2, \ldots$, we obtain by some simple computations that:

$$\begin{aligned} P\{|\hat{\sigma}_n^2 - \sigma^2(\eta)| > \lambda\} &\le P\{|\hat{\sigma}_n^2(\eta + \delta) - \sigma^2(\eta + \delta)| > \lambda/4\} \\ &+ P\{|\hat{\sigma}_n^2(\eta - \delta) - \sigma^2(\eta - \delta)| > \lambda/4\} + 4\lambda^{-1}|\sigma^2(\eta + \delta) - \sigma^2(\eta)| \\ &+ 4\lambda^{-1}|\sigma^2(\eta - \delta) - \sigma^2(\eta)| + P\{A_n\},\ n = 1, 2, \ldots . \end{aligned}$$

By the Weak Law of Large Numbers, $\lim P\{A_n\} = 0$.

Thus by Lemma 3.3:

$$\overline{\lim}\, P\{|\hat{\sigma}_n^2 - \sigma^2(\eta)| > \lambda\} \le 4\lambda^{-1}|\sigma^2(\eta + \delta) - \sigma^2(\eta)| + 4\lambda^{-1}|\sigma^2(\eta - \delta) - \sigma^2(\eta)|.$$

Consequently the desired result follows from the continuity of $\sigma^2(\theta)$ in $[\eta - \delta_0, \eta + \delta_0]$, by letting $\delta \to 0^+$. ||

We obtain now the consistency of $\hat{\sigma}_n^2$ under H_0.

Theorem 3.5. Assume $\sigma^2(\theta)$ is finite in an interval that contains μ. Then under H_0

$$p - \lim \hat{\sigma}_n^2 = \sigma^2.$$

Proof. Note that under H_0, $\mu = \eta$ and that $\sigma^2 = \sigma^2(\eta)$. Consequently the desired result follows by Lemma 3.4. ||

Next we show that our test is consistent.

Theorem 3.6. Assume (A.1), (A.2), and (A.3) hold. Further, assume that $\mu < \infty$, and that $\sigma^2(\theta)$ is finite in an interval that contains η. Then the test, which rejects H_0 in favor of H_1 if $n^{1/2}\{J_n^c - (1/4)\}\hat{\sigma}_n^{-1} < -z_\alpha$, is consistent against all continuous NBU alternatives.

Proof. Note that

$$P\{n^{1/2}\{J_n^c - (1/4)\}\hat{\sigma}_n^{-1} \leq -z_\alpha\} = P\{n^{1/2}\{J_n^c - \Delta(F)\} \leq -z_\alpha\hat{\sigma}_n + n^{1/2}((1/4) - \Delta(F))\},$$

that under H_1 $(1/4) - \Delta(F) > 0$, and that by Lemma 3.4 p- lim $\hat{\sigma}_n = \sigma(\eta) < \infty$. Consequently the desired result follows by Theorem 2.1. ||

Finally, we present a sufficient condition for $\sigma^2(\theta) < \infty$ at $\theta\epsilon(0,\infty)$.

Lemma 3.1. Let $\theta\epsilon(0,\infty)$. Assume

$$\lim_{z\to 0^+} z^{4-\beta}[\overline{K}(-\theta\ell nz)]^{-1} < \infty \text{ for some } \beta\epsilon(0,\infty). \tag{3.4}$$

Then $\sigma^2(\theta) < \infty$.

Proof. To obtain the desired result is sufficies to show that there is a $\delta\epsilon(0,1)$, such that

$$\int_0^\delta f(z)[\overline{K}(-\theta\ell nz)]^{-1}dz < \infty.$$

We show now the preceding inequality. There is a $\delta\epsilon(0,1)$ and a $d\epsilon(0,\infty)$, such that $z^{4-\beta} \leq d\overline{K}(-\theta\ell nz)$, $z\epsilon[0,\delta)$. Thus:

$$\int_0^\delta f(z)[\overline{K}(-\theta\ell nz)]^{-1}dz \leq d\int_0^\delta f(z)z^{\beta-4}dz.$$

Consequently the desired result follows by an evaluation of $\int_0^\delta f(z)z^{\beta-4}dz$. ||

4. Efficiency loss due to censoring. Recall that the J_n^c test is a generalization of the HP (1972) test for the uncensored model based on the statistic J_n (see equation (1.5) of HP (1972)). In this section we study the efficiency loss due to the presence of censoring by comparing the power of the J_n test based on n observations in the uncensored model with the power of the J_n^c test based on n* observations in the randomly censored model.

Let F_γ be a parametric family within the NBU class with F_{γ_0} being exponential with scale parameter 1 (for example, one such family is the Weibull $F_\gamma(x) = 1 - \exp\{-(x)^\gamma\}$, $\gamma \geq 1$ and $\gamma_0 = 1$) and assume the randomly censored model with $F = F_\gamma$ and with censoring distribution H. Consider the sequence of alternatives $\gamma_n = \gamma_0 + cn^{-1/2}$, with $c > 0$. Let $\beta_n(\gamma_n)$ be the power of the approximate α-level J_n test based on n observations in the uncensored model and let $\beta_{n^*}(\gamma_n)$ denote the power of the (approximate) α-level test based on J_n^c for n* observations in the randomly censored model. Consider $n^* = h(n)$ such that $\lim \beta_n(\gamma_n) = \lim \beta_{n^*}(\gamma_n)$, where the limiting value is strictly between 0, and 1, and let $k = \lim n/n^*$. The value of k can be viewed as a measure of the efficiency loss due to censoring. The value of k is adopted from Pitman's (cf. Noether, 1955) measure of asymptotic relative efficiency but the interpretation of k must be modified because J_n and J_n^c are not competing tests which are both applicable in the randomly censored model. Roughly speaking, for large n and NBU alternatives close to the null hypothesis of exponentiality, the J_n^c test requires n/k observations from the randomly censored model to do as well as the J_n test applied to n observations from the uncensored model. It can be shown that since J_n and J_n^c have the same asymptotic means, k reduces to

$$k \overset{\text{def}}{=} e_H(J^c, J) = (5/432)/\sigma^2(1) \tag{4.1}$$

where (5/432) is the null asymptotic variance of $n^{1/2}J_n$ and $\sigma^2(1)$, given by (3.1), is the null asymptotic variance of $n^{1/2}J_n^c$. Thus note that k depends only on the censoring distribution H, and not on the parametric family F_γ of NBU alternatives. Hence we use the notation $e_H(J^c, J)$, rather than $e_{F,H}(J^c, J)$, in (4.1).

We consider the cases (i) where the censoring distribution is exponential, $\overline{H}_1(x) = 1$ for $x < 0$, $\overline{H}_1(x) = \exp(-\lambda x)$, $x > 0$, and (ii) where the censoring distribution is piecewise exponential, $\overline{H}_2(x) = 1$ for $x < 0$, and for $\gamma = 1, \ldots, m$, $\overline{H}_2(x) = c_r \exp(-\lambda_r x)$, $s_{r-1} < x \leq s_r$, and $\overline{H}_2(x) = c_{m+1}\exp(-\lambda_{m+1}x)$, $s_m < x$ where $c_r = \exp\left(-\sum_{i=1}^{r-1} \lambda_i(s_i - s_{i-1} - 1) + \lambda_r s_{r-1}\right)$, and $s_0 = 0$.

For H_1, we see that (A.2) is satisfied with $\varepsilon = 0$ and thus we impose the restriction $\lambda \leq 1$. Then from (3.1) and (4.1) we find

$$e_{H_1}(J^c, J) = 5(3 - \lambda)^3/\{27(\lambda^2 - 2\lambda + 5)\}. \tag{4.2}$$

Values of $e_{H_1}(J^c, J)$ are given in Table 4.1. From (4.2) we note that, as is to be expected, as λ tends to 0 (corresponding to the case of no censoring), $e_{H_1}(J^c, J)$ tends to 1.

In order to provide a reference point to the amount of censoring, and thereby facilitate the interpretation of $e_{H_1}(J^c, J)$, we also include in Table 4.1 the value of $p_{H_1} = P(X_1 < Y_1) = (1 + \lambda)^{-1}$, the probability of obtaining an uncensored observation when X_1 is exponential with scale parameter 1 and Y_1 is independent of X_1 and has the censoring distribution H_1.

When the censoring distribution is H_2, straightforward but tedious calculations yield

$$(4.3)\quad e_{H_2}(J^c, J) = \sum_{r=1}^{m+1} 27(\lambda_r^2 - 2\lambda_r + 5)(5c_r)^{-1}(3 - \lambda_r)^{-3}[\exp\{-(3 - \lambda_r)s_{r-1}\}$$
$$- \exp\{-(3 - \lambda_r)s_r\}] - \sum_{r=1}^{m+1} 108(1 - \lambda_r)(5c_r)^{-1}(3 - \lambda_r)^{-2}[s_{r-1}\exp\{-(3 - \lambda_r)s_{r-1}\}$$
$$- s_r\exp\{-(3 - \lambda_r)s_r\}] + \sum_{r=1}^{m+1} 108(5c_r)^{-1}(3 - \lambda_r)^{-1}[s_{r-1}^2\exp\{-(3 - \lambda_r)s_r\}$$
$$- s_r^2\exp\{-(3 - \lambda_r)s_r\}],$$

where $s_{m+1} = \infty$. Again, with the censoring distribution H_2, (A.2) can include the case $\epsilon = 0$ and thus here we have $\lambda_m \leq 1$.

Values of $e_{H_2}(J^c, J)$ are also given in Table 4.1. Again, as a reference point for the amount of censoring under the censoring distribution H_2, we include in Table 4.1 values of $p_{H_2} = (X_1 < Y_1)$ when X_1 is exponential with scale parameter 1, and Y_1 is independent of X with distribution H_2. Direct calculations show

$$p_{H_2} = 1 - \sum_{r=1}^{m+1} c_r\lambda_r(1 + \lambda_r)^{-1}[\exp\{-(\lambda_r + 1)s_{r-1}\} - \exp\{-(\lambda_r + 1)s_r\}], \text{ where } s_{m+1} = \infty.$$

TABLE 4.1

Efficiency loss under exponential (H_1) and piecewise exponential (H_2) censoring.

		(H_1)			
λ:	1	1/2	1/3	1/4	1/10
$e_{H_1}(J^c,J)$:	.371	.681	.790	.844	.939
p_{H_1}:	.500	.667	.750	.800	.909

(H_2)

m=1	s_1:	1	1	1	2
	(λ_1, λ_2):	(1/2, 1)	(1, 1/2)	(1/2, 1/3)	(1/2, 1)
	$e_{H_2}(J^c, J)$:	.529	.498	.723	.642
	p_{H_2}:	.630	.523	.685	.675
m=2	(s_1, s_2):	(1/2, 1)	(1/2, 1)	(1/2, 1)	
	$(\lambda_1, \lambda_2, \lambda_3)$:	(1, 1/2, 1/3)	(1/3, 1/2, 1)	(1/2, 1/3, 1/4)	
	$e_{H_2}(J^c, J)$:	.617	.597	.772	
	p_{H_2}:	.576	.667	.718	

5. An Example. The data in Table 5.1 are found in Hollander and Proschan (1979) and are an up-dated version of data given by Koziol and Green (1976). The data correspond to 211 state IV prostate cancer patients treated with estrogen in a Veterans Administration Cooperative Urological Research Group study. At the March, 1977 closng date there were 90 patients who died of prostate cancer, 105 who died of other diseases, and 16 still alive. Those observations corresponding to deaths due to other causes and those corresponding to the 16 survivors are treated as censored observations (withdrawals). As reported by Koziol and Green (1976), there is a basis for suspecting that had the patients not been treated with estrogen, their survival distribution for deaths from cancer of the prostate wuld be exponential with mean 100 months.

Hollander and Proschan (1979) developed a goodness-of-fit procedure for testing, in the randomly censored model, that F is a certain (completely specified) distribution. They applied their test, and competing procedures of Koziol and Green (1976) and Hyde (1977), to the data of Table 5.1. The hypothesized F was taken to be exponential with mean 100. The two-sided P values obtained were consistent with the hypothesis. However, Gregory (1979) has proposed some goodness-of-fit tests which (for certain alternatives) may be more powerful than the tests of Hollander and Proschan (1979), Koziol and Green (1976), and Hyde (1977). Gregory's tests, applied to the data of Table 5.1, strongly indicate a deviation from the postulated exponential, with mean 100, distribution.

Possible alternative models include an exponential distribution with a mean different than 100, or a distribution, such as an NBU distribution, that could represent "wearout." To explore the possiblity of the latter type of alternative, it is reasonable to apply the test based on J_n^c.

Applying our NBU test to the data of Table 5.1, we obtain $J_{211}^c = .193$, $\hat{\sigma}_{211}^2 = .105$ and $(211)^{1/2}\{J_{211}^c - (1/4)\}\hat{\sigma}_{211}^{-1} = -2.56$ with a corresponding one-sided P value of .0052. Thus the test indicates strong evidence of wearout and suggests that an NBU model is preferable to an exponential model.

TABLE 5.1

Survival times and withdrawal times in months for 211 patients (with number of ties given in parentheses)

Survival times: 0(3), 2, 3, 4, 6, 7(2), 8, 9(2), 11(3), 12(3), 15(2), 16(3), 17(2), 18, 19(2), 20, 21, 22(2), 23, 24, 25(2), 26(3), 27(2), 28(2), 29(2), 30, 31, 32(3), 33(2), 34, 35, 36, 37(2), 38, 40, 41(2), 42(2), 43, 45(3), 46, 47(2), 48(2), 51, 53(2), 54(2), 57, 60, 61, 62(2), 67, 69, 87, 97(2), 100, 145, 158.

Withdrawal times: 0(6), 1(5), 2(4), 3(3), 4, 6(5), 7(5), 8, 9(2), 10, 11, 12(3), 13(3), 14(2), 15(2), 16, 17(2), 18(2), 19(3), 21, 23, 25, 27, 28, 31, 32, 34, 35, 37, 38(4), 39(2), 44(3), 46, 47, 48, 49, 50, 53(2), 55, 56, 59, 61, 62, 65, 66(2), 72(2), 74, 78, 79, 81, 89, 93, 99, 102, 104(2), 106, 109, 119(2), 125, 127, 129, 131, 133(2), 135, 136(2), 138, 141, 142, 143, 144, 148, 160, 164(3).

Acknowledgement. We are grateful to Professor Jayaraman Sethuraman for many helpful discussions.

References

[1] Barlow, R. E. and Proschan, F. (1975). Statistical Theory of Reliability and Life Testing Probability Models. Holt, Rinehart and Winston, New York.

[2] Billingsley, P. (1968). Convergence of Probability Measures. Wiley, New York.

[3] Breiman, L. (1968). Probability. Addison-Wesley, London.

[4] Breslow, N. and Crowley, J. (1974). A large sample study of the life table and product limit estimates under random censoring. Ann. Statist. 2 437-453.

[5] Cox, D. R. and Hinkley, D. V. (1974). Theoretical Statistics. Chapman and Hall, London.

[6] David, H. A. (1970). Order Statistics. Wiley, New York.

[7] Esary, J. D., Marshall, A. W. and Proschan F. (1973). Shock models and wear processes. Ann. Prob. 1 627-649.

[8] Gregory, G. G. (1979). Goodness-of-fit with randomly censored data: Tests based on the conditional distribution. University of Texas at El Paso, Department of Mathematics Technical Report.

[9] Hollander, M. and Proschan, F. (1972). Testing whether new is better than used. Ann. Math. Statist. 43 1136-1146

[10] Hollander, M. and Proschan, F. (1979). Testing to determine the underlying distribution using randomly censored data. Biometrics 35 393-401.

[11] Hollander, M. and Wolfe, D. A. (1973). Nonparametric Statistical Methods. Wiley, New York.

[12] Hyde, J. (1977). Testing survival under right censoring and left truncation. Biometrika 64 225-230.

[13] Kaplan, E. L. and Meier, P. (1958). Nonparametric estimation from incomplete observations. J. Amer. Statist. Assoc. 53 457-481.

[14] Koziol, J. A. and Green, S. B. (1976). A Cramér-von Mises statistic for randomly censored data. Biometrika 63 465-474.

[15] Langberg, N. A., Proschan, F. and Quinzi, A. J. (1980). Estimating dependent lifelengths, with applications to the theory of competing risks. Ann. Statist. 8 (to appear).

[16] Marshall, A. W. and Proschan F. (1972). Classes of distributions applicable in replacement, with renewal theory implications. Proceedings of the 6th Berkeley Symposium on Mathematical Statistics and Probability, vol. I, ed. by L. LeCam, J. Neyman, and E. L. Scott, pp. 395-415, Berkeley, California: University of California Press.

[17] Noether, G. E. (1955). On a theorem of Pitman. Ann. Math. Statist. 26 64-68.

[18] Parzen, E. (1962). Stochastic Processes. Holden-Day, San Francisco.

[19] Peterson, A. V. (1977). Expressing the Kaplan-Meier estimator as a function of empirical subsurvival functions. J. Amer. Statist. Assoc. 72 854-858.

[20] Randles, R. H. and Wolfe, D. A. (1979). Introduction to the Theory of Nonparametric Statistics. Wiley, New York.

UNCLASSIFIED

Security Classification of this Page

REPORT DOCUMENTATION PAGE

1. REPORT NUMBERS FSU No. M503 AFOSR No. 78-12	2. GOVT. ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE Testing Whether New is Better Than Used With Randomly Censored Data		5. TYPE OF REPORT AND PERIOD COVERED Technical Report 6. PERFORMING ORG. REPORT NUMBER FSU Statistics Report M503
7. AUTHOR(s) Yuan-Yan Chen, Myles Hollander, Naftali Langberg		8. CONTRACT OR GRANT NUMBER(s) AFOSR 78-3678
9. PERFORMING ORGANIZATION NAME AND ADDRESS The Florida State University Department of Statistics Tallahassee, Florida 32306		10. PROGRAM ELEMENT, PROJECT, TASK AREA AND WORK UNIT NOS.
10. CONTROLLING OFFICE NAME AND ADDRESS The U.S. Air Force Air Force Office of Scientific Research Bolling Air Force Base, D.D. 20332		12. REPORT DATE June, 1980 13. NUMBER OF PAGES 25
14. MONITORING AGENCY NAME AND ADDRESS (if different from Controlling Office)		15. SECURITY CLASS (of this report) Unclassified 15a. DECLASSIFICATION/DOWNGRADING SCHEUDLE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release: distribution unlimited

17. DISTRIBUTION STATEMENT (of the abstract, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS

New better than used, exponentiality, hypothesis test, censored data

20. ABSTRACT

A life distribution F is new better than used (NBU) if $\bar{F}(x + y) \leq \bar{F}(x)\bar{F}(y)$ for all $x, y \geq 0$ ($\bar{F} \equiv 1 - F$). Using a randomly censored sample of size n from F, we propose a test of H_0: F is exponential, versus H_1: F is NBU, but not exponential. Our test is based on the statistic $J_n^c = \iint \bar{F}_n(x + y)dF_n(x)dF_n(y)$, where F_n is the product limit estimator of F, introduced by Kaplan and Meier (1958).

Under mild regularity on the amount of censoring, the asymptotic normality of J_n^c, suitably normalized, is established. Then using a consistent estimator of the null standard deviation of $n^{1/2}J_n^c$, an asymptotically distribution-free test is obtained. Finally, using tests for the censored and uncensored models we develop a measure of the efficiency loss due to the presence of censoring.