The Annals of Mathematical Statistics
1972, Vol. 43, No. 4, 1136-1146

TESTING WHETHER NEW IS BETTER THAN USED[1]

BY MYLES HOLLANDER AND FRANK PROSCHAN
The Florida State University

A U-statistic J_n is proposed for testing the hypothesis H_0 that a new item has stochastically the same life length as a used item of any age (i.e., the life distribution F is exponential), against the alternative hypothesis H_1 that a new item has stochastically greater life length ($\bar{F}(x)\bar{F}(y) \geqq \bar{F}(x+y)$, for all $x \geqq 0$, $y \geqq 0$, where $\bar{F} = 1 - F$). J_n is unbiased; in fact, under a partial ordering of H_1 distributions, J_n is ordered stochastically in the same way. Consistency against H_1 alternatives is shown, and asymptotic relative efficiencies are computed. Small sample null tail probabilities are derived, and critical values are tabulated to permit application of the test.

1. Introduction and summary. In performing reliability analyses, it has been found very useful to classify life distributions F(i.e., distributions for which $F(t) = 0$ for $t < 0$) according to the monotonicity properties of the failure rate, or alternately, the average failure rate. See Barlow and Proschan (1965), Barlow, Marshall, and Proschan (1963), Birnbaum, Esary, and Marshall (1966), and Esary, Marshall, and Proschan (1970a, b). (Additional references are presented in these papers.)

Recently, several new classes of life distributions have been shown to be fundamental in the study of replacement policies (Marshall and Proschan, 1970). Properties of such life distributions have been treated in Esary, Marshall, and Proschan (1970a, b).

DEFINITION 1.1. A life distribution F is *new better than used* (NBU) if

$$\bar{F}(x+y) \leqq \bar{F}(x)\bar{F}(y) \qquad \text{for all } x, y \geqq 0 , \tag{1.1}$$

where $\bar{F} \equiv 1 - F$. The corresponding concept of a *new worse than used* (NWU) distribution is defined by reversing the inequality in (1.1).

The NBU property defined in (1.1) has also been referred to as "positive aging" by Bryson and Siddiqui (1969).

Property (1.1) may be interpreted as stating that the chance $\bar{F}(x)$ that a new unit will survive to age x is greater than the chance $\bar{F}(x+y)/\bar{F}(y)$ that an unfailed unit of age y will survive an additional time x. That is, a new unit has stochastically greater life than a used unit of any age.

The boundary members of the NBU class, obtained by insisting on equality in (1.1), are of course the exponential distributions, for which used items are

Received June 7, 1971; revised November 1971.

[1] Research supported by Air Force Office of Scientific Research Grant AFOSR-71-2058, Office of Naval Research Contract NONR-988(08), NR 042-004, and National Science Foundation Grant GU 2612. The United States Government is authorized to reproduce and distribute reprints for governmental purposes notwithstanding any copyright notation hereon.

no worse (and no better) than new items. In this paper we consider the inferential problem of testing

$$H_0\colon F(x) = 1 - \exp(-\lambda x), \qquad x \geqq 0\,, \lambda > 0 \;(\lambda \text{ unspecified})\,, \tag{1.2}$$

versus

$$H_1\colon F \text{ is NBU (and not exponential)}\,, \tag{1.3}$$

on the basis of a random sample $X_1, X_2, \cdots, X_n$ from the distribution F. In the sequel, unless otherwise stated, F is assumed continuous.

The testing problem H_0 vs. H_1 is analogous to the testing problem of H_0 vs. H_1^* where H_1^* specifies that F is an increasing failure rate (IFR) distribution. The distribution F is said to be IFR if $-\ln \bar{F}(x)$ is convex. If F has a density f, this condition is equivalent to the condition that the failure rate $q(x) = f(x)/\bar{F}(x)$ is increasing in $x(x$ such that $\bar{F}(x) > 0)$. Tests of H_0 $(q(x) = \lambda, \lambda$ unspecified) vs. $H_1^*(q(x)$ is monotone increasing but nonconstant) include those considered by Barlow (1968), Barlow (1970), Barlow and Proschan (1969), Bickel (1969), Bickel and Doksum (1969), and Proschan and Pyke (1967). Since F is NBU if $-\ln \bar{F}(x)$ is superadditive, the IFR class is contained in the NBU class, and thus the test of H_0 vs. H_1 that we propose focuses on a larger class of alternative distributions than do the IFR tests. This will be appropriate, for example, when the underlying physical process suggests that new items are better than used ones but where we can expect the failure rate to fluctuate (and in particular not satisfy H_1^*).

As one example of a practical problem motivating the choice of the null hypothesis H_0 and alternative hypothesis H_1 above, consider a unit subject to shocks occurring successively in time according to a Poisson process. Since the occurrence of shocks and their effects cannot be directly observed, it is not known whether shocks already experienced by the unit make it more likely to fail under the impact of future shocks or not. However, if $\bar{P}_k$ is the probability that the unit survives the first k shocks, then it is believed that either

(a) $\bar{P}_k \equiv \bar{P}_{l+k}/\bar{P}_l$ for all $k, l \geqq 0$, or

(b) $\bar{P}_k \geqq \bar{P}_{l+k}/\bar{P}_l$ for all $k, l \geqq 0$.

Since under hypothesis (a), the lifelength is exponential, and under hypothesis (b), it is NBU (Esary, Marshall, Proschan (1970b) Theorem 3.1), a reasonable way to test (a) vs. (b) would be to test H_0 vs. H_1 above from lifelength observations.

Our test statistic is motivated by consideration of

$$\begin{aligned} \gamma(F) &=_{\text{def.}} \iint \{\bar{F}(x)\bar{F}(y) - \bar{F}(x+y)\}\, dF(x)\, dF(y) \\ &= \tfrac{1}{4} - \iint \bar{F}(x+y)\, d(F)\, dF(y) \\ &=_{\text{def.}} \tfrac{1}{4} - \Delta(F)\,. \end{aligned} \tag{1.4}$$

Viewing the parameter $\gamma(F)$ as a measure of the deviation of F from H_0, the classical nonparametric approach of replacing F by the empirical distribution function F_n suggests rejecting H_0 in favor of H_1 if $\iint \bar{F}_n(x+y)\,dF_n(x)\,dF_n(y)$ is too small. We find it more convenient to reject for small values of the asymptotically equivalent U-statistic

$$J_n = 2[n(n-1)(n-2)]^{-1} \textstyle\sum' \psi(X_{\alpha_1}, X_{\alpha_2} + X_{\alpha_3}) , \tag{1.5}$$

where

$$\begin{aligned} \psi(a, b) &= 1 \quad \text{if} \quad a > b \\ &= 0 \quad \text{if} \quad a \leqq b , \end{aligned} \tag{1.6}$$

and the $\sum'$ is over all $n(n-1)(n-2)/2$ triples $(\alpha_1, \alpha_2, \alpha_3)$ of three integers such that $1 \leqq \alpha_i \leqq n$, $\alpha_1 \neq \alpha_2$, $\alpha_1 \neq \alpha_3$, and $\alpha_2 < \alpha_3$. In the sequel, the test which rejects for small J_n values is referred to as the NBU test.

Section 2 demonstrates unbiasedness, asymptotic normality, and consistency of J_n. The NBU test is unbiased for NBU alternatives. In fact, a stronger result (Theorem 2.1) is established, namely, that when F is superadditive with respect to G (Definition 2.1), $J_n(\mathbf{X}) \leqq_{\text{st}} J_n(\mathbf{Y})$, where $\mathbf{Y} = (Y_1, Y_2, \cdots, Y_n)$ is a random sample from G. The asymptotic normality of J_n (Theorem 2.2) is a direct consequence of Hoeffding's (1948) U-statistic theory. The NBU test is consistent if and only if $\Delta(F)$, defined by (1.4), is strictly less than $\frac{1}{4}$, the latter being the value of Δ when F is exponential. In a result that parallels Theorem 2.1, we show (Theorem 2.3) that $\Delta(F) \leqq \Delta(G)$ when F is superadditive with respect to G. Also, consistency against NBU alternatives is established.

Section 3 considers the asymptotic relative efficiency of the NBU test. To the authors' knowledge, other tests for NBU alternatives have not yet been proposed. Thus, we take as competitors, tests designed for IFR alternatives. Since the NBU class contains the IFR class, one should expect the efficiencies (under IFR alternatives) to favor the IFR tests, and indeed this is the case. On the other hand, as is also to be expected, there are many NBU alternative distributions for which the IFR tests do not perform as well as the NBU test. The class of NBU alternatives $\mathscr{F}_{a,b}$, defined in Example 2.2, for which the NBU test is shown to have power equal to 1 (when $n \geqq 3$ and $\alpha \geqq \binom{2n-2}{n}^{-1}$) illustrates this point vividly. This is discussed in Section 3.

The small sample null distribution of the statistic $T_n = n(n-1)(n-2)J_n/2$ is considered in Section 4. Exact probabilities are computed in special cases, and in Table 4.1 lower and upper percentile points based on Monte Carlo sampling are given in the .01, .025, .05, .075, and .10 regions for $n = 4(1)20(5)50$.

2. Unbiasedness, asymptotic normality, and consistency. In this section we first show that the test which rejects H_0 if $J_n \leqq j_{n,\alpha}$, where $j_{n,\alpha}$ satisfies $P_0[J_n \leqq j_{n,\alpha}] = \alpha$, is unbiased. That is, $P_1[J_n \leqq j_{n,\alpha}] \geqq \alpha$, where $P_1(P_0)$ indicates the probability is computed for an F satisfying H_1 (H_0).

DEFINITION 2.1. Let F and G be continuous distributions, G be strictly increasing on its support, and $F(0) = 0 = G(0)$. Then F is said to be *super-*

additive with respect to G if $G^{-1}F$ is superadditive, that is,

$$G^{-1}F(x_1 + x_2) \geqq G^{-1}F(x_1) + G^{-1}F(x_2)\,, \qquad \text{for all } x_1, x_2 \geqq 0\,. \tag{2.1}$$

When the inequality in (2.1) is reversed, F is said to be *subadditive with respect to G*.

THEOREM 2.1. *Let F be superadditive with respect to G. Then* $J_n(\mathbf{X}) \leqq_{\text{st}} J_n(\mathbf{Y})$, *where* $\mathbf{X} = (X_1, \cdots, X_n)$ *is a random sample from F and* $\mathbf{Y} = (Y_1, \cdots, Y_n)$ *is a random sample from G.*

PROOF. Let $Y_i' = G^{-1}F(X_i)$, $i = 1, \cdots, n$. Then $(Y_1', \cdots, Y_n') =_{\text{st}} (Y_1, \cdots, Y_n)$. Now we show

$$X_3 \geqq X_1 + X_2 \Rightarrow Y_3' \geqq Y_1' + Y_2'\,. \tag{2.2}$$

The implication given by (2.2) can be seen as follows:

$$X_3 \geqq X_1 + X_2 \Rightarrow F(X_3) \geqq F(X_1 + X_2)$$
$$\Rightarrow G^{-1}F(X_3) \geqq G^{-1}F(X_1 + X_2) \geqq G^{-1}F(X_1) + G^{-1}F(X_2)\,,$$

where the last inequality is a consequence of the superadditivity of $G^{-1}F$. Equivalently, $Y_3' \geqq Y_1' + Y_2'$. From (2.2) we have $\psi(Y_3', Y_1' + Y_2') \geqq \psi(X_3, X_1 + X_2)$ and thus

$$J_n(\mathbf{X}) \leqq J_n(\mathbf{Y}') =_{\text{st}} J_n(\mathbf{Y})\,. \qquad \square$$

COROLLARY 2.1. *The* NBU *test is unbiased against* NBU *alternatives.*

PROOF. By taking G to be exponential, and noting that F is NBU if and only if F is superadditive with respect to the exponential (i. e., $-\ln \bar{F}(x)$ is a superadditive function for $x \geqq 0$), the result is a direct consequence of Theorem 2.1. $\square$

Some examples of parametric families of life distributions which are increasingly superadditive as the parameter θ increases are:

(a) *Weibull.* $F_\theta(t) = 1 - \exp(-\lambda t^\theta)$, $t \geqq 0$, $\lambda \geqq 0$.

(b) *Gamma.* $F_\theta(t) = \int_0^t \lambda^\theta x^{\theta-1} \exp(-\lambda x)/\Gamma(\theta)$, $t \geqq 0$, $\lambda \geqq 0$.

In each case, for fixed $\lambda \geqq 0$ and $0 < \theta_1 < \theta_2 < \infty$, F_{θ_2} is superadditive with respect to F_{θ_1}. It follows that the power function is an increasing function of the parameter θ.

The asymptotic normality of J_n is obtained by applying Hoeffding's (1948) U-statistic theory. Let

$$\Phi(x_1, x_2, x_3) = J_3 = 3^{-1}\{\psi(x, x_2 + x_3) + \psi(x_2, x_1 + x_3) + \psi(x_3\, x_1 + x_2)\}\,,$$

and set $\Phi_1(x_1) = E\Phi(x_1, X_2, X_3)$, $\Phi_2(x_1, x_2) = E\Phi(x_1, x_2, X_3)$, $\Phi_3(x_1, x_2, x_3) = \Phi(x_1, x_2, x_3)$, and $\xi_k = E\Phi_k^2(X_1, \cdots, X_k) - \Delta^2$, $k = 1, 2, 3$, where $\Delta(F)$ is defined by (1.4). Then $\operatorname{Var}(J_n) = \binom{n}{3}^{-1} \sum_{k=1}^3 \binom{3}{k}\binom{n-3}{3-k}\xi_k$, $\lim n \operatorname{Var}(J_n) = 9\xi_1$, and furthermore we may state

THEOREM 2.2. *If F is such that* $\xi_1(F) > 0$, *then the limiting distribution of*

$n^{\frac{1}{2}}(J_n - \Delta(F))$ *is normal with mean* 0 *and variance* $9\xi_1$.

Since $\psi(ca, cb) = \psi(a, b)$ for all $c > 0$, the statistic J_n is scale invariant, and hence in all null computations we may take the scale parameter of the exponential to be $\lambda = 1$. Straightforward calculations yield the hypothesis values $\Delta = \frac{1}{4}$ and $\xi_1 = 5/3888$, $\xi_2 = 7/1296$, $\xi_3 = 1/48$. From Theorem 2.2, we immediately obtain

COROLLARY 2.2. *Under* H_0, *the limiting distribution of* $n^{\frac{1}{2}}(J_n - \frac{1}{4})$ *is normal with mean* 0 *and variance* 5/432.

We next turn to consistency. From Theorem 2.2 it is easily seen that the NBU test is consistent if and only if $\Delta(F) < \frac{1}{4}$. We now prove

THEOREM 2.3. *Let F be superadditive with respect to G. Then* $\Delta(F) \leqq \Delta(G)$.

PROOF. Make the transformation $\bar{F}(x_i) = \bar{G}(y_i)$, $i = 1, 2$. Then

$$\begin{aligned}\Delta(G) &= \textstyle\int\int \bar{G}(y_1 + y_2)\, dG(y_1)\, dG(y_2) \\ &= \textstyle\int\int \bar{G}[\bar{G}^{-1}\bar{F}(x_1) + \bar{G}^{-1}\bar{F}(x_2)]\, dF(x_1)\, dF(x_2)\,.\end{aligned} \tag{2.3}$$

Since $\bar{G}^{-1}\bar{F}$ is superadditive, then

$$\bar{G}^{-1}\bar{F}(x_1) + \bar{G}^{-1}\bar{F}(x_2) \leqq \bar{G}^{-1}\bar{F}(x_1 + x_2)\,. \tag{2.4}$$

Combining (2.3) and (2.4) gives

$$\Delta(G) \geqq \textstyle\int\int \bar{G}[\bar{G}^{-1}\bar{F}(x_1 + x_2)]\, dF(x_1)\, dF(x_2) = \Delta(F)\,. \tag{2.5}$$

□

THEOREM 2.4. *If F is continuous*, NBU, *and not exponential, then the* NBU *test is consistent.*

PROOF. We need only show that the hypotheses imply $\Delta(F) < \frac{1}{4}$. Since F is assumed continuous, $\gamma(F) = \frac{1}{4} - \Delta(F)$ and we may equivalently prove $\gamma(F) > 0$. Set $D(x_1, x_2) = \bar{F}(x_1)\bar{F}(x_2) - \bar{F}(x_1 + x_2)$. Then $D(x_1, x_2) \geqq 0$ for all $x_1, x_2 \geqq 0$ since F is NBU and $D(x_1, x_2) \not\equiv 0$ since F is not exponential.

Assume that x_1^0, x_2^0 are such that $D(x_1^0, x_2^0) > 0$. Let $x_i' = \sup\{x : x \geqq x_i^0$ and $\bar{F}(x) = \bar{F}(x_i^0)\}$, $i = 1, 2$. Then

$$\begin{aligned}D(x_1', x_2') &\geqq \bar{F}(x_1')\bar{F}(x_2') - \bar{F}(x_1^0 + x_2^0) \\ &= \bar{F}(x_1^0)\bar{F}(x_2^0) - \bar{F}(x_1^0 + x_2^0) = D(x_1^0, x_2^0) > 0\,.\end{aligned}$$

Since F is continuous, D is continuous and there exist $\delta_1 > 0$, $\delta_2 > 0$, such that $D(x_1' + \delta_1, x_2' + \delta_2) > 0$. Also $F(x_i' + \delta_i) - F(x_i') > 0$, $i = 1, 2$, since x_1' and x_2' are points of increase of F. Thus $\gamma(F) > 0$. □

The NBU distribution used in the following example illustrates that the middle equality of (1.4), namely $\gamma(F) = \frac{1}{4} - \Delta(F)$, need not hold when the continuity assumption is removed. This example thus emphasizes that $\Delta(F)$, rather than $\gamma(F)$, is the basic consistency parameter of the NBU test, since the NBU test is consistent if and only if $\Delta(F)$ is less than $\frac{1}{4}$ (rather than if and only if $\gamma(F)$ is greater than zero.)

EXAMPLE 2.1. Let $\bar{F}(x) = \exp(-[\lambda x])$ for $x \geqq 0$, where $[x]$ denotes the largest integer less than or equal to x. We now show that $\gamma(F) = 0$ and $\Delta(F) = (e+1)^{-2} \doteq .072$. Since $\gamma(F)$ and $\Delta(F)$ are scale invariant (i. e., if $F_\beta(x) = F_1(x/\beta)$ for every $\beta > 0$, then $\gamma(F_\beta)$ and $\Delta(F_\beta)$ are constant in β) we may take $\lambda = 1$. Since $\bar{F}(i)\bar{F}(j) - \bar{F}(i+j) = e^{-i}e^{-j} - e^{-(i+j)} = 0$, we have

$$\gamma(F) = \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} \{\bar{F}(i)\bar{F}(j) - \bar{F}(i+j)\}\, dF(i)\, dF(j) = 0\,.$$

We next determine $\Delta(F)$. Since $\gamma(F) = 0$, we have

$$\Delta(F) = \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} \bar{F}(i)\bar{F}(j)\, dF(i)\, dF(j) = \{\sum_{i=1}^{\infty} \bar{F}(i)\, dF(i)\}^2\,.$$

But,

$$\sum_{i=1}^{\infty} \bar{F}(i)\, dF(i) = \sum_{i=1}^{\infty} e^{-i}(e^{-(i-1)} - e^{-i}) = (e-1)\cdot\sum_{i=1}^{\infty} e^{-2i} = (e+1)^{-1}\,.$$

Since $\Delta(F) = (e+1)^{-2} < \frac{1}{4}$, the NBU test is consistent against this alternative even though $\gamma(F) = 0$.

Example 2.2 below provides a class of NBU alternatives for which the NBU test is not only consistent, but for which the NBU test has power identically equal to 1 for every n. (We make the minor restriction that α_n, the level of the test based on small J_n values, exceeds $P_0[J_n = 0] = \binom{2n-2}{n}^{-1}$ (see (4.3)), so that when $J_n = 0$ we reject H_0 with probability 1.)

EXAMPLE 2.2. Let $\mathscr{F}_{a,b}$ denote the class of distributions with support $[a, b]$ where $b < 2a$. Then, by considering the three cases (i) $x < a$; $y < a$, (ii) $x \geqq a$; $y \geqq a$, and (iii) $x < a$; $y \geqq a$, and substituting in (1.1) one directly verifies that every $F \in \mathscr{F}_{a,b}$ is NBU. But for every $F \in \mathscr{F}_{a,b}$, $P_F[J_n = 0] = P_F[X_{(n)} < X_{(1)} + X_{(2)}] = 1$, and thus $P_F[\text{Rej } H_0] \geqq P_F[J_n = 0] = 1$.

3. Asymptotic relative efficiency and power. As far as we know, no other tests have as yet been proposed for testing against NBU alternatives. Thus in this section we compare the proposed NBU test with tests designed for a smaller class of alternatives, the IFR class. When the underlying distribution is actually IFR, it is to be expected that an IFR test will in general perform better than the NBU test. Switching the comparison to grounds where the NBU test should excel, we exhibit a class of NBU distributions for which the NBU test performs distinctly better than the IFR tests.

IFR tests that have been proposed include:

(i) Proschan and Pyke (1967): Define the normalized spacings $S_i = (n - i + 1)(X_{(i)} - X_{(i-1)})$, where $X_{(1)} \leqq \cdots \leqq X_{(n)}$ are the ordered X's with $X_{(0)} =_{\text{def.}} 0$. The Proschan-Pyke test rejects H_0 in favor of H_1^* for large values of

$$V_n = \sum_{i<j}^{n} V_{ij}\,, \tag{3.1}$$

where $V_{ij} = 1$ if $S_i \geqq S_j$, 0 otherwise.

(ii) Total time on test (cf. Epstein (1960), Barlow (1968), Bickel and Doksum

(1969), Bickel (1969), Barlow and Proschan (1969), Barlow (1970)): Reject H_0 in favor of H_1^* for large values of the cumulative total time on test statistic

$$K_n = \sum_{i=1}^{n-1} \sum_{j=1}^{i} S_j / \sum_{t=1}^{n} S_t . \tag{3.2}$$

(iii) Bickel and Doksum (1969): Reject H_0 in favor of H_1^* for large values of

$$W_n = \sum_{i=1}^{n} i \ln [1 - R_i/(n+1)] , \tag{3.3}$$

where R_i is the rank of S_i in the joint ranking of $S_1, \cdots, S_n$.

Let $\{F_{\theta_n}\}$ be a sequence of alternatives with $\theta_n = \theta_0 + kn^{-\frac{1}{2}}$, where k is an arbitrary positive constant and F_{θ_0} is exponential. From the results of Proschan and Pyke (1967), Bickel and Doksum (1969), and Theorem 2.2, we find the Pitman asymptotic relative efficiency of the NBU test with respect to the Proschan-Pyke test to be

$$e_F(J, V) = (12/5)\{\Delta'(\theta_0)/\mu'(\theta_0)\}^2 , \tag{3.4}$$

where

$$\Delta(\theta) = \iint \bar{F}_\theta(x+y)\, dF_\theta(x)\, dF_\theta(y) , \tag{3.5}$$

$$\mu(\theta) = \int_0^\infty \int_x^\infty q_\theta(y)[q_\theta(x) + q_\theta(y)]^{-1} f_\theta(x) f_\theta(y)\, dy\, dx , \tag{3.6}$$

are the asymptotic means of J_n and V_n respectively for the alternative F_θ, the factor (12/5) in (3.4) equals $\lim_n \{\mathrm{Var}_0(V_n)/\mathrm{Var}_0(J_n)\}$, and $\Delta'(\theta_0)(\mu'(\theta_0))$ is the derivative of $\Delta(\theta)(\mu(\theta))$ with respect to θ, evaluated at $\theta = \theta_0$. For simplicity, we have used the V_n test in our efficiency calculations. Bickel and Doksum (1969) have shown $e_F(V, W) = \frac{3}{4}$ and $e_F(K, W) = 1$, for all F, and thus $e_F(J, K) = e_F(J, W) = (\frac{3}{4})e_F(J, V)$. Consider the IFR Weibull and linear failure rate alternatives given respectively by $F_1(x) = 1 - \exp(-x^\theta)$, $\theta \geqq 1$, $x \geqq 0$, and $F_2(x) = 1 - \exp(-\{x + (\theta x^2/2)\})$, $\theta \geqq 0$, $x \geqq 0$. For F_1, H_0 is achieved at $\theta = \theta_0 = 1$ whereas for F_2, H_0 is achieved at $\theta = \theta_0 = 0$. Direct calculations yield

$$e_{F_1}(J, V) = 1.25 , \qquad e_{F_1}(J, W) = .937 ;$$
$$e_{F_2}(J, V) = .60 , \qquad e_{F_2}(J, W) = .45 .$$

Next we compare the power of the NBU test J_n with the power of the IFR tests V_n and W_n for the class $\mathscr{F}_{a,b}$ of NBU distributions introduced in Example 2.2. We showed there that the J_n test has power 1 for every $n \geqq 3$, for every $F \in \mathscr{F}_{a,b}$, as long as $\alpha \geqq \binom{2n-2}{n}^{-1}$. Consider the V_n and W_n tests based on the normalized spacings. For simplicity, take $n = 3$ and $\alpha = \frac{1}{6}$. Then both V_n and W_n reject H_0 when $A = [S_1 > S_2 > S_3]$ occurs. It is easily seen that for every $F \in \mathscr{F}_{a,b}$, $P_F[S_1 > S_2,\ S_1 > S_3] = 1$, but for many distributions in $\mathscr{F}_{a,b}$, $P_F[A] < 1$, implying that for these distributions the power of the V_n (and W_n) test is less than 1. Here the $\alpha = \frac{1}{4}$ test based on J_n has power 1. The case $n = 3$ was chosen for convenience. It is clear that for larger n we can exhibit F's $\in \mathscr{F}_{a,b}$ for which the powers of the V_n and W_n tests are less than 1 (the

TABLE 4.1
Critical values of the T_n statistic

	Lower Tail					Upper Tail				
$n \backslash \alpha$	.01	.025	.05	.075	.10	.01	.025	.05	.075	.10
4				0(.067)	1(.103)					
5	0(.018)	1(.028)	2(.040)	3(.072)	4					10(.177)
6	2	5(.028)	7	8(.061)	9(.086)			20(.053)		19(.167)
7	7	11	14(.045)	16(.072)	18(.105)	35		34(.047)		33(.103)
8	15	21	25	28	30	55	54	53	52(.083)	52(.083)
9	27	34	40	44	47	81	80(.019)	78	77	76
10	42	52	60	65	69	114	112	110	109(.067)	107(.107)
11	63	76	86	92	97	155	152(.028)	150(.046)	148(.070)	146
12	89	105	117	125	131	205	201	198(.047)	195	193
13	122	141	157	167	174	264	260	255	252(.072)	249
14	162	185	204	215	223	334	328	322	318	315
15	209	236	259	272	282	415	408	401	396	392
16	266	298	323	338	350	508	499	491	484	480
17	330	368	397	415	429	613	603	593	586	580
18	405	446	480	502	518	732	720	709	700	693
19	490	538	577	601	619	866	852	838	828	820
20	594	642	685	713	732	1015	998	982	970	961
25	1250	1351	1427	1472	1507	2018	1986	1956	1934	1918
30	2320	2463	2574	2651	2704	3521	3461	3411	3375	3349
35	3850	4064	4215	4316	4394	5625	5546	5464	5403	5359
40	5947	6214	6434	6593	6700	8418	8301	8189	8113	8049
45	8665	9040	9341	9543	9686	12012	11864	11709	11600	11502
50	12170	12661	13020	13255	13439	16519	16310	16085	15937	15823

latter value being the power of the J_n test) and at the same time where the corresponding Type I errors of V_n and W_n exceed that of J_n.

Of course, the class $\mathscr{F}_{a,b}$ contains F's that are IFR (e. g. uniform on $[a, b]$) and ones that are not. Thus this class simultaneously provides (i) F's which are NBU but not IFR for which the NBU test is better (as is to be expected) and (ii) F's which are IFR for which the NBU test is better!

4. The null distribution. Define

$$T_n = n(n-1)(n-2)J_n/2 = \sum' \psi(X_{\alpha_1}, X_{\alpha_2} + X_{\alpha_3}) . \tag{4.1}$$

Let $X_{(1)} \leqq \cdots \leqq X_{(n)}$ denote the ordered X's. Since $i \leqq \max(j, k)$ implies $\psi(X_{(i)}, X_{(j)} + X_{(k)}) = 0$, we can rewrite T_n as

$$T_n = \sum_{i>j>k} \psi(X_{(i)}, X_{(j)} + X_{(k)}) . \tag{4.2}$$

Note that T_n has possible values $0, 1, \cdots, n(n-1)(n-2)/6$. Exact percentile points for the NBU test can be obtained from the distribution of T_n, calculated under the assumption that the X's are exponential. For even moderate n, these calculations are prohibitive. We now obtain exact probabilities for some special cases. (Where exact probabilities are available, they show excellent agreement with the Monte Carlo values given in Table 4.1.)

Define the spacings $A_i = X_{(i)} - X_{(i-1)}$, $i = 1, \cdots, n$. Then, for $n \geqq 3$, we have

$$\begin{aligned} P_0[T_n = 0] &= P_0[X_{(n)} < X_{(1)} + X_{(2)}] = P_0[\textstyle\sum_{i=3}^n A_i < A_1] \\ &= n! \int_0^\infty \int_0^\infty \cdots \int_{a_3+\cdots+a_n}^\infty \prod_{i=1}^n \exp(-(n-i+1)a_i)\, da_i = \binom{2n-2}{n}^{-1} . \end{aligned} \tag{4.3}$$

For $n = 3$, $P_0[T_3 \leqq 1] = 1$, and calculations similar to (4.3) yield, for $n \geqq 4$,

$$P_0[T_n \leqq 1] = \binom{2n-3}{n-3}^{-1} \left\{ \frac{(3n-1)(n-2)}{(2n-2)(2n-1)} \right\} . \tag{4.4}$$

The complete H_0 distribution of T_4 is found, by direct calculation, to be $P_0[T_4 = 0] = 7/105$, $P[T_4 = 1] = 4/105$, $P_0[T_4 = 2] = 16/105$, $P_0[T_4 = 3] = 33/105$, $P_0[T_4 = 4] = 45/105$.

We also have, for $n \geqq 3$

$$\begin{aligned} P_0[T_n = n(n-1)(n-2)/6] &= P_0[X_{(3)} > X_{(1)} + X_{(2)}, X_{(4)} > X_{(2)} \\ &\quad + X_{(3)}, \cdots, X_{(n)} > X_{(n-1)} + X_{(n-2)}] \\ &= n! \int_0^\infty \int_{x_1}^\infty \int_{x_1+x_2}^\infty \cdots \int_{x_{n-2}+x_{n-3}}^\infty \int_{x_{n-1}+x_{n-2}}^\infty \prod_{i=1}^n \exp(-x_{n-i+1})\, dx_{n-i+1} . \end{aligned} \tag{4.5}$$

To evaluate the n-fold integral of (4.5), let $c_j(d_j)$ denote the coefficient of $-x_{n-j}(-x_{n-j-1})$ in the exponent after the jth integration. Then $c_j = c_{j-1} + d_{j-1}$ and $d_j = c_{j-1} + 1$, with $c_0 = d_0 = 1$. Thus the desired probability is seen to be equal to

$$P_0[T_n = n(n-1)(n-2)/6] = n! \left[\prod_{j=1}^{n-1} c_{j-1}^{-1}\right][c_{n-2} + d_{n-2}]^{-1} . \tag{4.6}$$

We now obtain expressions for c_j and d_j. Define the y sequence by $y_{j+2} = d_j$

with the initial values $y_0 = 0$, $y_1 = 1$. This sequence satisfies $y_{j+2} = y_{j+1} + y_j$; this is the famous Fibonacci sequence for which (cf. Brand (1966) page 381) $y_n = (5)^{-\frac{1}{2}}\{s^n - t^n\}$, where $s = \{1 + (5)^{\frac{1}{2}}\}/2$, $t = \{1 - (5)^{\frac{1}{2}}\}/2$. If we define the z sequence by $z_{j+2} = c_j$ with initial values $z_0 = z_1 = 0$, then the relations between the c's and d's yield $z_{n-1} = y_n - 1$. Solving for the c's and d's in terms of the y's and z's and substituting into (4.6) yields

$$(4.7) \qquad P_0[T_n = n(n-1)(n-2)/6] = \frac{n!\prod_{j=1}^{n-1}[-1 + 5^{-\frac{1}{2}}(s^{j+2} - t^{j+2})]^{-1}}{-1 + 5^{-\frac{1}{2}}(s^{n+1} + s^n - t^{n+1} - t^n)} .$$

To make the NBU test practical, we need more tail probabilities than those available via direct calculations. Table 4.1, based on Monte Carlo sampling, gives lower and upper critical points of T_n in the $\alpha = .01, .025, .05, .075$ and $.10$ regions for $n = 4(1)20(5)50$. For $n \leqq 19$, each value is based on 100,000 replications, for $n > 19$, on 10,000 replications. In Table 4.1, the lower tail should be used for tests of H_0 versus F NBU, the upper tail for tests of H_0 versus F NWU. The lower tail values are integers $C_\alpha{}^L$ for which the estimated probabilities $\hat{P}_0[T_n \leqq C_\alpha{}^L]$ are closest to α, and similarly the upper tail values are integers $C_\alpha{}^U$ for which the estimated probabilities $\hat{P}_0[T_n \geqq C_\alpha{}^U]$ are closest to α. Parenthetical entries adjacent to critical points give the Monte Carlo estimated tail probabilities for those estimated probabilities that are not within .002 of the nominal α. For $n \geqq 25$, all estimated probabilities agree with the nominal α(to three decimal places). For $n > 50$, use the normal approximation (see Section 2) keeping in mind that (i) lower tail probabilities of events of the form $[T_n \leqq a]$ are underestimated using the normal approximation and upper tail probabilities of events $[T_n \geqq b]$ are overestimated, and (ii) for fixed n, the approximation improves as α increases, $0 \leqq \alpha < \frac{1}{2}$.

We close with a comparison of some exact probabilities and their Monte Carlo estimates. Letting $B_n = \{T_n = 0\}$, $C_n = \{T_n \leqq 1\}$, $D_n = \{T_n = n(n-1)(n-2)/6\}$, the values of $(P_0\{B_n\}, \hat{P}_0\{B_n\})$ are, for $n = 4, 5, 6, 7$, respectively (.067, .066), (.018, .018), (.005, .005), (.001, .001); the corresponding values of $(P_0\{C_n\}, \hat{P}_0\{C_n\})$ are (.105, .103), (.028, .028), (.007, .007), (.002, .002) and those of $(P_0\{D_n\}, \hat{P}_0\{D_n\})$ are (.429, .431), (.179, .177), (.054, .053), (.011, .011). The very close agreement inspires confidence that the entries in Table 4.1 are accurate.

Acknowledgment. We are indebted to Douglas E. Whitten for programming the Monte Carlo calculations which led to Table 4.1, to Ramesh M. Korwar for checking the efficiency calculations, and to the referee and associate editor for helpful suggestions.

REFERENCES

[1] BARLOW, R. E. (1968). Likelihood ratio tests for restricted families. *Ann. Math. Statist.* **39** 547-560.

[2] BARLOW, R. E. (1970). Isotonic tests for convex orderings. Univ. of California: Berkeley Report ORC 70-22.

[3] BARLOW, R. E., MARSHALL, A. W., and PROSCHAN, F. (1963). Properties of probability distributions with monotone hazard rate. *Ann. Math. Statist.* **34** 375-389.

[4] BARLOW, R. E. and PROSCHAN, F. (1965). *Mathematical Theory of Reliability*. Wiley, New York.

[5] BARLOW, R. E. and PROSCHAN, F. (1969). A note on tests for monotone failure rate based on incomplete data. *Ann. Math. Statist.* **40** 595-600.

[6] BICKEL, P. (1969). Tests for monotone failure rate II. *Ann. Math. Statist.* **40** 1250-1260.

[7] BICKEL, P. and DOKSUM, K. (1969). Tests on monotone failure rate based on normalized spacings. *Ann. Math. Statist.* **40** 1216-1235.

[8] BIRNBAUM, Z. W., ESARY, J. D., and MARSHALL, A. W. (1966). A stochastic characterization of wear-out for components and systems. *Ann. Math. Statist.* **37** 816-825.

[9] BRAND, L. (1966). *Differential and Difference Equations*. Wiley, New York.

[10] BRYSON, M. C. and SIDDIQUI, M. M. (1969). Some criteria for aging. *J. Amer. Statist. Assoc.* **64** 1472-1483.

[11] EPSTEIN, B. (1960). Tests for the validity of the assumption that the underlying distribution of life is exponential. *Technometrics* **2** Part I, 83-101 and Part II, 167-183.

[12] ESARY, J. D., MARSHALL, A. W., and PROSCHAN, F. (1970a). Some reliability applications of the hazard transform. *SIAM J. Appl. Math.* **18** 849-860.

[13] ESARY, J. D., MARSHALL, A. W., and PROSCHAN, F. (1970b). Shock models and wear processes. Florida State University Statistics Report M194.

[14] HOEFFDING, W. A. (1948). A class of statistics with asymptotically normal distribution. *Ann. Math. Statist.* **19** 293-325.

[15] MARSHALL, A. W. and PROSCHAN, F. (1972). Classes of distributions applicable in replacement, with renewal theory implications. To appear in *Proc. Sixth Berkeley Symp. Math. Statist. Prob.*

[16] PROSCHAN, F. and PYKE, R. (1967). Tests for monotone failure rate. *Proc. Fifth Berkeley Symp. Math. Statist. Prob.* **3** 293-312.