The Annals of Mathematical Statistics
1969, Vol. 40, No. 4, 1216–1235

TESTS FOR MONOTONE FAILURE RATE BASED ON NORMALIZED SPACINGS[1]

By Peter J. Bickel[2] and Kjell A. Doksum

University of California, Berkeley

1. Introduction and summary. Let F be a distribution with density f, and let $q(t) = f(t)/[1 - F(t)]$ be the failure rate of F. Tests for constant versus monotone increasing failure rate based on the ranks of the normalized spacings between the ordered observations have been considered by Proschan and Pyke (1967). They show that these statistics are asymptotically normally distributed for fixed alternatives F and compute the ratios of the efficacies of one of their rank tests to the best statistics for Weibull and Gamma alternatives.

In this paper, it is shown that asymptotic normality holds also for sequences of alternatives $\{F_{\theta_n}\}$ that approach the H_0 distribution $1 - \exp(-\lambda t)$, $t \geqq 0$, as $n \to \infty$; and that the above mentioned ratios of efficacies are in fact Pitman efficiencies.

Let $R_1, \cdots, R_n$ be the ranks of the normalized spacings, $T_1 = \sum iR_i$ and $T_2 = -\sum i \log[1 - R_i/(n + 1)]$. Then T_1 is asymptotically equivalent to the Proschan Pyke statistic. It is shown that the Pitman efficiency satisfies

$$e(T_1, T_2) \equiv \tfrac{3}{4} \tag{1.1}$$

for all sequences of alternatives $\{F_{\theta_n}\}$ and thus T_1 is asymptotically inadmissible.

Statistics that are linear in the normalized spacings and asymptotically most powerful for parametric alternatives $\{F_{\theta_n}\}$ if the scale parameter λ is known, are derived, and it is shown that the rank statistics that are asymptotically most powerful in the class of linear rank tests, are nowhere most powerful in the class of all tests, when λ is known.

If λ is unknown, studentizing of the linear normalized spacing tests which are asymptotically most powerful for λ known leads to procedures which have only the same asymptotic power as the most powerful linear rank tests.

Unbiasedness is shown for tests that are monotone in the normalized spacings, and Monte Carlo power estimates are used to compare the various statistics with the likelihood ratio tests considered by Barlow (1967).

2. Tests monotone in the normalized spacings. Let $X_1, \cdots, X_n$ be a random sample from a population with a continuous distribution F satisfying $F(0) = 0$,

Received 11 November 1967; revised 3 February 1969.
[1] This research has been partially supported by the Office of Naval Research under Contracts Nonr-3656(18) and Nonr-222(83) with the University of California; Reproduction in whole or in part is permitted for any purpose of the United States Government.
[2] Professor Bickel's research was partially supported by the Office of Naval Research under Contract N00014-67-A-0114-0004.

and let $0 = X_{(0)} < X_{(1)} < \cdots < X_{(n)}$ be the order statistics. The normalized sample spacings $D_1, \cdots, D_n$ are defined by $D_i = (n - i + 1)\ (X_{(i)} - X_{(i-1)})$, $i = 1, \cdots, n$; and R_i denotes the rank of D_i among $D_1, \cdots, D_n$. The problem is to test H_0:"$-\log[1 - F(x)] = \lambda x$ on $[0, \infty)$ for some positive constant λ" against H_1:"$-\log[1 - F(x)]$ is convex and not of the form λx on $[0, \infty)$." Note that the only distributions satisfying H_0 are

$$K_\lambda(x) = 1 - e^{-\lambda x}, \qquad x \geqq 0, \lambda > 0. \tag{2.1}$$

Under H_0 it is well known that $(D_1, \cdots, D_n)$ has the same distribution as a random sample from a population with distribution K_λ, while under the alternative there is a downward trend in the sense that $P(D_i \geqq D_j) < \frac{1}{2}$ whenever $i > j$ (see Proschan and Pyke (1967). One thus defines a test $\phi = \phi(D_1, \cdots, D_n)$ to be *monotone in the D's* if

$$\begin{aligned} &\phi(D_1', \cdots, D_n') \leqq \phi(D_1, \cdots, D_n) \quad \text{for all} \\ &(D_1, \cdots, D_n) \quad \text{and} \quad (D_1', \cdots, D_n') \quad \text{such that} \\ &i < j \quad \text{and} \quad D_i' \geqq D_j' \quad \text{implies} \quad D_i \geqq D_j. \end{aligned} \tag{2.2}$$

Following van Zwet (1964), one defines a distribution F_1 to have a *more slowly increasing failure rate* than F, written $F_1 >_c F$, if $F_1^{-1}F$ is convex. Here $F_1^{-1}F$ is defined by $P(F_1^{-1}F(X) \leqq x \mid F) = F_1(x)$, $x \geqq 0$.

THEOREM 2.1. *Monotone tests have monotone power*, i.e., *if ϕ is a monotone test and if $F_1 >_c F$, then*

$$E(\phi \mid F_1) \leqq E(\phi \mid F). \tag{2.3}$$

PROOF. Since $F_1^{-1}F$ is increasing, $X_{(i)}' = F_1^{-1}F(X_{(i)})$ is the ith order statistic in a random sample from a population with distribution F_1. Let $D_i' = (n - i + 1)(X_{(i)}' - X_{(i-1)}')$, $i = 1, \cdots, n$. Since $F_1^{-1}F$ is convex, $i < j$ and $D_i' \geqq D_j'$ implies $D_i \geqq D_j$. From (2.2) one obtains

$$\phi(D_i', \cdots, D_n') \leqq \phi(D_1, \cdots, D_n), \tag{2.4}$$

and (2.3) follows upon taking expectations in (2.4).

Note that a test ϕ is *similar* if $E(\phi \mid K_\lambda)$ is independent of λ. Thus all rank tests are similar.

COROLLARY 2.1. *All monotone tests are rank tests and they are unbiased.*

PROOF. The rank property follows since (2.2) implies $\phi(D_1, \cdots, D_n) = \phi(g(D_1), \cdots, g(D_n))$ for all continuous and strictly increasing functions g. Unbiasedness follows by letting $F_1 = K_\lambda$ in Theorem 2.1.

COROLLARY 2.2. *If $-c_n(i)$ and $J_n(i)$ are nondecreasing in $i = 1, \cdots, n$, then the test that rejects when*

$$\textstyle\sum_{i=1}^n c_n(i) J_n(R_i) \geqq C_1 \tag{2.5}$$

has monotone power and is unbiased.

PROOF. Using the notation of Theorem 2.1, define R_i' to be the rank of D_i'

among $D_1', \cdots, D_n'$. Then $i < j$ and $R_i' \geqq R_j'$ implies $R_i \geqq R_j$. From Corollary 2 of Lehmann (1966) one obtains

$$\sum_{i=1}^n c_n(i) J_n(R_i') \leqq \sum_{i=1}^n c_n(i) J_n(R_i). \tag{2.6}$$

It follows that the test is monotone, and Theorem 2.1 and Corollary 2.1 apply.

REMARK 2.1. The results of this section can be used to obtain bounds on the power of monotone tests. For instance, if $\beta(V_n, F_3)$ denotes the power of the Proschan Pyke statistic V_n = "number of pairs (i, j) with $i < j$ and $D_i \geqq D_j$" for the Weibull distribution F_3, then the power satisfies $\beta(V_n, F) \geqq \beta(V_n, F_3)$ for all distributions F such that $F_3^{-1}F$ is convex.

REMARK 2.2. Barlow and Proschan [(1966) Theorem 3.12(v)] have shown that if $c_n(1) \geqq \cdots \geqq c_n(n)$, the test that rejects when

$$\sum_{i=1}^n c_n(i) D_i / \sum_{i=1}^n D_i \geqq C_2$$

is unbiased. Similarity follows at once from scale invariance.

REMARK 2.3. Monotone tests as defined by (2.2) have been considered earlier in a different context by Lehmann (1966) and Bell and Doksum (1967). The fact that all monotone tests must be rank tests (the first part of Corollary 2.1) was discovered independently by the referee and the authors while the paper was being refereed. The referee also makes several other interesting remarks listed below. The notation used is $D = (D_1, \cdots, D_n)$, $D' = (D_1', \cdots, D_n')$, $R = (R_1, \cdots, R_n)$, etc.

(i) Change the definition (2.2) to read

$$\phi(D') \leqq \phi(D) \quad \text{for all} \quad D \text{ and } D' \quad \text{such that} \quad D_i'/D_i \text{ is nondecreasing in } i, \tag{A}$$

then Theorem 2.1 continues to hold and the class of tests satisfying definition (A) is larger than the class of tests satisfying (2.2) (see (iv) below). Theorem 2.1 continues to hold since the convexity of $F_1^{-1}F$ implies that $(X'_{(i)} - X'_{(i-1)})/(X_{(i)} - X_{(i-1)}) \leqq (X'_{(j)} - X'_{(j-1)})/(X_{(j)} - X_{(j-1)})$ for $i < j$, and thus D_i'/D_i is nondecreasing in i. Corollary 2.1 now becomes: "All (A)-monotone tests are similar and unbiased." Similarity follows since (A) implies $\phi(D_1, \cdots, D_n) = \phi(\lambda D_1, \cdots, \lambda D_n)$.

(ii) The class of (2.2)-monotone tests equals the class of tests satisfying

$$\begin{aligned} &\phi(R') \leqq \phi(R) \quad \text{whenever there are } i, j \text{ and } m \text{ such that} \\ &i < j,\ R_i' = R_j = m, \quad R_j' = R_i = m + 1 \quad \text{and} \\ &R_k = R_k' \quad \text{for} \quad k \neq i, j; \end{aligned} \tag{B}$$

i.e., $\phi(R') \leqq \phi(R)$ whenever R can be obtained from R' by interchanging two consecutive integers that occur in ascending order (namely $m = R_i'$ and $(m + 1) = R_j'$, $i < j$).

To see this, first note that the class of (B)-monotone tests contains the class of (2.2)-monotone tests. For the converse, suppose R and R' are such that $i < j$ and $R_i' \geqq R_j'$ imply $R_i \geqq R_j$. To complete the proof, it is enough to show that R' can be turned into R by a series of interchanges of two consecutive integers. Take any pair (if it exists) $i < j$ with $R_i' = m$, $R_j' = m + 1$ and $R_i > R_j$, and interchange R_i' and R_j'. Repeat this process, raising the number of inversions in R' by one each step, until no more steps are possible (the total number of steps could be zero). The result is again called R'. It is claimed that $R = R'$. To see this, note that we still have that $i < j$ and $R_i' > R_j'$ imply $R_i > R_j$. Moreover, we now have that $i < j$, $R_i' = m$ and $R_j' = m + 1$ imply $R_i < R_j$. Together these two assertions yield that $R_i' = m$ and $R_j' = m + 1$ imply $R_i < R_j$ regardless of the order of i and j. Repeated application of this property yields that $R_i' < R_j'$ implies $R_i < R_j$ for any pair (i, j). Hence $R_i' < R_j'$ is equivalent to $R_i < R_j$ for any pair i, j and thus $R = R'$.

The condition (B) has essentially been considered by Hájek ((1968), Chapter 2) in a different context. Note that Corollary 2.2 follows at once from Definition (B).

(iii) "The class of (2.2)-monotone tests equals the class of (A)-monotone rank tests." The second class includes the first since $i < j$, $R_i' \geqq R_j'$, and D_i'/D_i increasing in i imply that $R_i \geqq R_j$. To show that the inclusion also goes the other way, one performs for given R, R', i, j and m as in (B) a construction of D and D' with ranks R and R' and such that D_k'/D_k is nondecreasing. The details are as follows: Let $\epsilon > 0$ and choose D in such a way that $(D_i/D_j) < 1 + \epsilon$ whereas for all other pairs (s, t) where we need $D_s > D_t$ to get the rank ordering R, D is chosen so that $(D_s/D_t) > 1 + \epsilon$.
The choice $D_k' = D_k$ for $k \leqq i$ and $D_k' = (1 + \epsilon)D_k$ for $k > i$ works.

(iv) Definition (A) provides a proof of the unbiasedness of the tests in Remark 2.2. To see this, let Z be a discrete random variable taking on the values $1, \cdots, n$ and let P_0 and P_1 be probabilities such that $P_0(Z = k) = D_k/\sum D_i$ and $P_1(Z = k) = D_k'/\sum D_i'$. Let D and D' be such that D_k'/D_k is nondecreasing in k, then P_0 and P_1 have a monotone likelihood ratio in k. It follows from Lehmann (1955) that $E_{P_0}(-c_n(Z)) = -\sum c_n(i)D_i/\sum D_i \leqq -\sum c_n(i) \cdot D_i'/\sum D_i' = E_{P_1}(-c_n(Z))$; thus the tests defined in Remark 2.2 are (A)-monotone.

3. Asymptotically most powerful tests. Let $\{f_{\theta,\lambda} : \theta \geqq 0, \lambda > 0\}$ be a class of densities such that $f_{0,\lambda}(x) = \lambda \exp(-\lambda x)$, $x \geqq 0$, and such that

$$h_\lambda(x) = (\partial/\partial\theta) \log f_{\theta,\lambda}(x) \mid_{\theta=0} \tag{3.1}$$

exists.

We suppose λ is a scale parameter, i.e.,

$$P_{\theta,\lambda}(\lambda X \leqq t) = P_{\theta,1}(X \leqq t). \tag{3.2}$$

For testing "$\theta = 0$" versus "$\theta > 0$", the locally most powerful test rejects "$\theta = 0$" for large values of $T_n(h_\lambda)$ where,

$$T_n(h) = n^{-\frac{1}{2}}\textstyle\sum_1^n h(X_i). \tag{3.3}$$

We will consider sequences of alternatives $\{f_{\theta_n,\lambda}\}$ such that

$$(3.4) \qquad \lim n^{\frac{1}{2}}\theta_n = b \qquad \text{for some} \quad 0 \leqq b < \infty .$$

A sequence $\{f_{\theta_n,\lambda}\}$ is said to be *contiguous* to $f_{0,\lambda}$ (in the sense of LeCam-Hájek) if for any sequence of random variables $\hat{R}_n(X_1, \cdots, X_n)$, $\hat{R}_n \to 0$ in P_0 probability implies $\hat{R}_n \to 0$ in P_{θ_n} probability, where P_θ denotes the probability distribution of $X_1, \cdots, X_n$ if $f_{\theta,\lambda}$ is true. The following condition implying contiguity for sequences as in (3.4) can be obtained from LeCam (1966).

(a) $\partial f_{\theta,\lambda}(x)/\partial\theta \neq 0 \qquad \text{whenever} \quad f_{\theta,\lambda}(x) > 0.$

(b) For some $\delta > 0$ and all $\theta \varepsilon [0, \delta]$,

$$(3.5) \qquad 0 < H(\theta) = \int_0^\infty [\delta f_{\theta,\lambda}(x)/\partial\theta]^2 [f_{\theta,\lambda}(x)]^{-1} \, dx < \infty,$$

and $H(\theta)$ is continuous in θ.

(c) $\lim_{|\epsilon|\to 0} \int_0^\infty \{h^{-1}(f^{\frac{1}{2}}_{\theta+\epsilon,\lambda}(x) - f^{\frac{1}{2}}_{\theta,\lambda}(x)] - \frac{1}{2}[\partial f_{\theta,\lambda}(x)/\partial\theta] f^{-\frac{1}{2}}_{\theta,\lambda}(x)\}^2 dx = 0$

for $\theta \varepsilon [0, \delta]$, some $\delta > 0$.

An easy sufficient condition implying (3.5) (b) and (c) is

$$(3.6) \qquad \int_0^\infty \sup \{[\partial f_{\theta,\lambda}(x)/\partial\theta]^2 [f_{\theta,\lambda}(x)]^{-1} : 0 \leqq \theta \leqq \delta\} \, dx < \infty \quad \text{for some } \delta > 0.$$

It also follows from LeCam's work that under condition (3.5) we have

$$(3.7) \qquad [bn^{-\frac{1}{2}} \textstyle\sum_{i=1}^n h_\lambda(X_i) - \frac{1}{2}b^2 E_{0,\lambda}(h_\lambda^2(X_1))] - \textstyle\sum_{i=1}^n \{\log f_{\theta_n,\lambda}(X_i) - \log f_{0,\lambda}(X_i)\} \to 0$$

in P_0 probability and hence in P_{θ_n} probability.

According to Wald (1941), a sequence of level α tests $\{\phi_n\}$ is said to be *asymptotically most powerful* (λ known) if

$$(3.8) \qquad \lim_n \sup \{E_{\theta,\lambda}(\psi_n) - E_{\theta,\lambda}(\phi_n) : \theta > 0, E_{0,\lambda}(\psi_n) \leqq \alpha\} = 0.$$

Using contiguity we can prove Wald's (1941) main theorem under weaker conditions.

THEOREM 3.1. *Suppose* (3.5) *holds and,*

$$(3.9) \qquad T_n(h_\lambda) \to \infty .$$

in $P_{\theta_n,\lambda}$ probability if $n^{\frac{1}{2}}\theta_n \to \infty$. Then the sequence of level α tests ϕ_n which reject for suitably large values of $T_n(h_\lambda)$ is asymptotically most powerful.

PROOF. The result is an easy consequence of the Neyman-Pearson Lemma, (3.7) and (3.9). □

REMARK 3.1. Under the conditions of the theorem it follows that if $\{S_n\}$ is any sequence of statistics such that $S_n - T_n(h_\lambda) \to 0$ in $P_{0,\lambda}$ probability, then S_n and $T_n(h_\lambda)$ have the same asymptotic distribution under both hypothesis and contiguous alternatives. If, furthermore,

$$(3.10) \qquad S_n \to \infty$$

in $P_{\theta_n,\lambda}$ probability when $n^{\frac{1}{2}}\theta_n \to \infty$, we may conclude that the natural sequence of level α tests which reject for large values of S_n is asymptotically most powerful (λ known). An important class of such S_n is discussed in the following section.

4. Linear approximations to locally and asymptotically most powerful statistics. As we have seen, locally and asymptotically most powerful test statistics for parametric alternatives are of the form $T_n(h) = n^{-\frac{1}{2}} \sum h(X_i)$, for some function h on $[0, \infty)$. It will be shown that for each fixed, regular function h (h is not necessarily related to the functions h_λ of Section 3), $T_n(h)$ can be approximated by a statistic linear in the D_i's, given by,

$$S_n(h) = n^{-\frac{1}{2}} \textstyle\sum_{i=1}^{n} a_h(i/(n+1))(D_i - 1) \tag{4.1}$$

where,

$$a_h(u) = (1-u)^{-1} \int_{-\log(1-u)}^{\infty} h'(x) e^{-x}\, dx, \qquad 0 < u < 1. \tag{4.2}$$

When h equals h_1 of Section 3, a_{h_1} and $S_n(h_1)$ will be called a and S_n.

THEOREM 4.1. *Let h be any function such that*

(i) *h' is continuous on* $(0, \infty)$,

(ii) $\int_0^\infty h(t)e^{-t}\, dt = 0, 0 < \int_0^\infty h^2(t) e^{-t}\, dt < \infty$,

(iii) *either one of the following holds*

(a) *$h'(-\log(1-u))$, $0 < u < 1$, satisfies assumption E of Chernoff, Gastwirth and Johns* (1967) *and* $\int_0^\infty |h'(v)|\, e^{-v/2}(1 - e^{-v})^{\frac{1}{2}}\, dv < \infty$.

(b) *$h'(t)$ changes sign only a finite number of times as $t \to 0$ or ∞ and h' does not vanish infinitely often.*

Moreover, suppose that the X's have the exponential density $\exp(-x)$, $x \geqq 0$. *Then $T_n - S_n$ tends to zero in probability.*

The proof is deferred to the appendix, Section 7.

REMARK 4.1. It may also be shown that if (i), (ii) and (iii) (b) hold and $\lambda = 1$, then $E_{0,1}(T_n - S_n)^2 \to 0$.

The following remarks are readily obtained from the results of this section and the appendix.

REMARK 4.2. Generalize a_h and $S_n(h)$ to

$$a_{h,\lambda}(u) = a_{h(\cdot/\lambda)}(u) = (1-u)^{-1} \int_{-\log(1-u)}^{\infty} \lambda^{-1} h'(x/\lambda) e^{-x}\, dx,$$
$$S_{n,\lambda}(h) = n^{-\frac{1}{2}} \lambda \textstyle\sum a_{h,\lambda}(i/(n+1))(D_i - \lambda^{-1}).$$

Then the theorem continues to hold if the X's have the density $\lambda \exp(-\lambda x)$, $x \geqq 0$, $\lambda > 0$, and we use $S_{n,\lambda}(h)$ as the approximand. Moreover, $S_{n,\lambda}(h)$ can be written

$$S_{n,\lambda}(h) = \lambda S_{n,1}(h(\cdot/\lambda)) + n^{-\frac{1}{2}}(\lambda - 1) \textstyle\sum a_{h,\lambda}(i/(n+1)). \tag{4.3}$$

REMARK 4.3. Let $S_{n,\lambda}$ denote $S_{n,\lambda}(h_\lambda)$, then since $h_\lambda(t) = h_1(\lambda t)$,

$$S_{n,\lambda} = \lambda S_n(h_1) + n^{-\frac{1}{2}}(\lambda - 1) \textstyle\sum a(i/(n+1)). \tag{4.4}$$

Therefore, if the conditions of Theorem 3.1 and 4.1 hold and $S_{n,\lambda}$ satisfies (3.10), then the level α tests which reject for large values of $S_n = S_n(h_1)$ are asymptotically most powerful for each value of λ.

5. Asymptic normality and inefficiency of rank statistics. Let Pitman asymptotic efficiency be defined as usual (e.g., Hodges and Lehmann (1961)). In this section we shall show the asymptotic normality of statistics of the form

$$W_n = n^{-\frac{1}{2}}\sum_i^n (c_i - \bar{c})J(R_i/(n+1)), \qquad \bar{c} = n^{-1}\sum_i^n c_i, \tag{5.1}$$

and compute their Pitman efficiencies with respect to the asymptotically most powerful statistics of Section 4 Let h_λ be as defined in (3.1) and let $S_{n,\lambda}$ be the asymptotically most powerful statistics of Section 4. It will be clear in the sequel that for most purposes we can take $\lambda = 1$ in which case we shall write S_n for $S_{n,\lambda}$. Also, $a(u)$ will henceforth always be defined with $h = h_1$. Thus $a(u)$ depends on the distribution of the X's through h_1. Finally, we set $a_i = a(i/(n+1)), i = 1, \cdots, n$.

THEOREM 5.1. *If* (3.4) *and* (3.5) *hold, if* h_λ *satisfies the conditions of Theorem* 4.1 *of this paper, and if* $\{c_i\}$ *and* $\{J(i/(n+1))\}$ *satisfy the conditions of Theorem* 4.1 *of Hájek* (1961), *then for the sequence of alternatives given in* (3.4) $[W_n - \mu(W_n)]/\sigma(W_n)$ *has asymptotically a standard normal distribution, where*

$$\mu(W_n) = b[n^{-1}\sum_1^n a_i(c_i - \bar{c})]\int_0^1 J(u)[-\log(1-u) - 1]\,du, \tag{5.2}$$

and

$$\sigma^2(W_n) = [n^{-1}\sum_1^n (c_i - \bar{c})^2]\,[\int_0^1 J^2(u)\,du - (\int_0^1 J(u)\,du)^2]. \tag{5.3}$$

PROOF. Since the ranks are scale invariant, assume without loss of generality that $\lambda = 1$. The results of Hájek (1961) and the contiguity condition imply that

$$W_n - Q_n \to 0 \quad \text{in } P_{\theta_n,1} \text{ probability, where,} \tag{5.4}$$

$$Q_n = n^{-\frac{1}{2}}\sum_1^n (c_i - \bar{c})J(1 - \exp(-D_i)). \tag{5.5}$$

It follows from the Lindeberg-Feller theorem that under H_0, the joint limiting distribution of bS_n and Q_n is the bivariate normal distribution with means zero, variances $\lim b^2\sigma^2(S_n)$ and $\lim \sigma^2(W_n)$ and covariance

$$\lim b[n^{-1}\sum_1^n a_i(c_i - \bar{c})]\,[\int_0^1 J(u)[-\log(1-u) - 1]\,du]$$

where

$$\sigma^2(S_n) = n^{-1}\sum_1^n a_i^2.$$

The result now follows from LeCam's third lemma (see Hájek and Šidák (1967), p. 208) and the results of Sections 3 and 4.

For each vector $c = (c_1, \cdots, c_n)$, define

$$V_n(c) = n^{-1}\sum_1^n (c_i - \bar{c})^2 \tag{5.6}$$

and for two vectors c and a define

$$\text{Cor}_n^2(a, c) = [n^{-1}\sum a_i(c_i - \bar{c})]^2/[V_n(a)V_n(c)]. \tag{5.7}$$

Then Theorem 5.1 yields

COROLLARY 5.1. *Under the conditions of Theorem* 5.1, *the Pitman asymptotic relative efficiency of* W_n *to* S_n *is*

$$(5.8) \quad e(W, S) = \mathrm{Cor}^2(J(U), -\log(1-U)) \cdot \lim_{n\to\infty} [\mathrm{Cor}_n^2(a, c) V_n(a)/(n^{-1}\textstyle\sum_1^n a_i^2)^{-1}]$$

where U *is a uniform* $(U(0, 1))$ *random variable.*

It is clear from (5.8) that one obtains the most efficient linear rank statistic, call it $W_n^{(1)}$, by taking $J(u) = -\log(1-u)$ and $c_i = a_i$. Note that the choice of J is independent of the alternative densities $\{f_{\theta,\lambda}\}$. In this case we have

$$(5.9) \qquad e(W^{(1)}, S) = \mathrm{Var}(a(U))/E(a^2(U)).$$

Evidently, if $E(a(U)) = \int_0^\infty x h_1(x) e^{-x}\, dx \neq 0$, then $e(W, S) < 1$. This is equivalent to having $\bar{X}_n = n^{-1}\sum_1^n X_i$ correlated with the locally most powerful statistic T_n. It will be shown in the proof of the next result that $\mathrm{Cor}(\bar{X}_n, T_n) < 0$ when the failure rate is increasing. Let

$$(5.10) \qquad q_\lambda(x, \theta) = f_{\theta,\lambda}(x)\,[1 - F_{\theta,\lambda}(x)]^{-1}, \qquad x \geqq 0,$$

denote the failure rate of $F_{\theta,\lambda}$. We assume in what follows that $\partial q_\lambda(x, \theta)/\partial\theta$ exists and is continuous in (x, θ) for $0 \leqq \theta \leqq \delta$, $x > 0$, some $\delta > 0$. Let

$$(5.11) \qquad L(x) = \partial q_\lambda(x, \theta)/\partial\theta \mid_{\theta=0}.$$

Increasing failure rate clearly implies

$$(5.12) \qquad L(x) \geqq 0$$

for all $x > 0$. We have

THEOREM 5.2. *Suppose the conditions of Theorems* 5.1 *hold, that* (5.12) *is satisfied with strict inequality for* x *in some set of positive measure and that,*

$$(5.13) \qquad \int_0^\infty t e^{-t} L(t)\, dt < \infty.$$

Then each linear rank statistic of the form (5.1) *is inefficient*, i.e.,

$$(5.14) \qquad e(W^{(1)}, S) < 1.$$

PROOF. Without loss of generality let $\lambda = 1$. Now,

$$(5.15) \qquad f_{\theta,1}(x) = q_1(x, \theta) \exp\{-\textstyle\int_0^x q_1(t, \theta)\, dt\}.$$

Since $\partial q_1(x, \theta)/\partial\theta$ is continuous in (x, θ) we have, differentiating under the integral sign,

$$(5.16) \qquad h_1(x) = L(x) - \int_0^x L(t)\, dt.$$

Using Fubini's theorem and (5.13) we get if $\lambda = 1$,

$$(5.17) \quad \mathrm{cov}\,(X_1, h_1(X_1)) = \int_0^\infty x h_1(x) e^{-x}\, dx = -\int_0^\infty e^{-x} L(x)\, dx < 0$$

by our assumptions. □

EXAMPLE 5.1. An alternative for which there is an efficient rank test is provided by

$$f_{\theta,1}(x) = (1-\theta)^{-1}[1+\theta(-2x+\tfrac{1}{2}x^2)]e^{-x}, \qquad \theta < \tfrac{1}{2}. \tag{5.18}$$

For this density, the failure rate is not monotone. It is easy to see that Cor $(X_1, h(X_1)) = 0$. The efficient rank test rejects for large values of

$$n^{-\frac{1}{2}}\textstyle\sum_1^n [\log(1 - i/(n+1)) + 1][-\log(1 - R_i/(n+1))]. \tag{5.19}$$

It follows from Hájek (1961) that if

$$\begin{aligned} W_n^{(2)}(r_1, \cdots, r_n) &= E_{0,1}(S_n \mid r_1, \cdots, r_n) \\ &= n^{-\frac{1}{2}}\textstyle\sum_{i=1}^n a(i/(n+1))E_{0,1}(X_{(r_i)}), \end{aligned} \tag{5.20}$$

then

$$E_{0,1}(W_n^{(1)} - W_n^{(2)})^2 \to 0, \tag{5.21}$$

where $W_n^{(2)} = W_n^{(2)}(R_1, \cdots, R_n)$.

It is not difficult to see by using the method of Hoeffding (1951) that the locally most powerful rank test rejects the null hypothesis for large values of the statistic $W_n^{(3)}(R_1, \cdots, R_n)$, where

$$\begin{aligned} W_n^{(3)}(r_1, \cdots, r_n) &= E_{0,1}(T_n \mid r_1, \cdots, r_n) \\ &= n^{-\frac{1}{2}}\textstyle\sum_{i=1}^n E_{0,1}[h_1(\sum_{j=1}^i X_{(r_j)}(n-j+1)^{-1})]. \end{aligned} \tag{5.22}$$

By Remark 4.1 and since a projection argument yields

$$E_{0,1}(W_n^{(2)} - W_n^{(3)})^2 \leq E_{0,1}(S_n - T_n)^2, \tag{5.23}$$

then $W_n^{(2)}$ and $W_n^{(3)}$ (and hence $W_n^{(1)}$) are asymptotically equivalent.

These rank statistics are of course usable even if λ is unknown, the situation which primarily concerns us, while the optimal statistics of Section 4 depend on λ and do not lead to similar tests. If we use the method of Barlow (1968) and Nadler and Eilbott (1967) and consider studentized statistics of the form

$$S_n^* = n^{\frac{1}{2}}\textstyle\sum_{i=1}^n a(i/(n+1))D_i(\sum_{i=1}^n D_i)^{-1} - n^{-\frac{1}{2}}\sum_{i=1}^n a(i/(n+1)), \tag{5.24}$$

it is not difficult to show that for any fixed λ under the null hypothesis,

$$S_n^* - n^{-\frac{1}{2}}\lambda\textstyle\sum_{i=1}^n (a(i/(n+1)) - \bar{a})(D_i - \lambda^{-1}) \to 0 \tag{5.25}$$

in probability, and hence by (5.4) the best studentized test of this form is *asymptotically equivalent* under contiguous alternatives to the rank tests based on $W_n^{(1)}$, $W_n^{(2)}$, and $W_n^{(3)}$.

It may be shown that the asymptotically most powerful linear rank tests, asymptotically equivalent to the studentized asymptotically most powerful linear spacings tests, are in fact asymptotically equivalent to the level α tests which are most powerful among all tests which are similar and level α. For the hypothesis H_0, λ unknown, this and related results are given in [5].

REMARK 5.1. If L is defined by (5.11) with $\lambda = 1$, then $W_n^{(1)}$ can be written $W_n^{(1)} = n^{-\frac{1}{2}}\sum_{i=1}^{n} - L(-\log(1 - i/(n+1)) [-\log(1 - R_i/(n+1))]$. This form is used to derive the statistics $W_i(i = 1, 2, 3)$ considered in the next section.

6. Applications of the theory and Monte Carlo results. The results of the previous sections will now be applied to specific alternatives and specific statistics. The weights c_i of (5.1) will be of the form

$$c_i = c(i(n+1)^{-1}) \quad \text{for some function} \quad c \quad \text{on} \quad (0, 1) \tag{6.1}$$

and the efficiency (5.8) will be

$$e(W, S) = \text{Cor}^2(J(U), -\log(1 - U)) \cdot \text{Cor}^2(a(U), c(U)) \,\text{Var}(a(U))/E(a^2(U)) \tag{6.2}$$

where U is an uniform ($U(0, 1)$) random variable.

The statistics to be considered are:

$$W_0 = \textstyle\sum_1^n - (i/(n+1))\,(R_i/(n+1))$$
$$W_1 = \textstyle\sum_1^n - (i/(n+1))\,[-\log(1 - R_i/(n+1))]$$
$$W_2 = \textstyle\sum_1^n [-\log(1 - i/(n+1))]\,[-\log(1 - R_i/(n+1))]$$
$$W_3 = \textstyle\sum_1^n - \{\log[-\log(1 - i/(n+1))]\}\,[-\log(1 - R_i/(n+1))]$$
$$W_4 = \textstyle\sum_1^n g(i/(n+1))\,[-\log(1 - R_i/(n+1))]$$
$$S_1 = \textstyle\sum_1^n - (i/(n+1))\,D_i$$
$$S_2 = \textstyle\sum_1^n [\log(1 - i/(n+1))]\,D_i$$
$$S_3 = \textstyle\sum_1^n - \{\log[-\log(1 - i/(n+1))]\}\,D_i$$
$$S_4 = \textstyle\sum_1^n g(i/(n+1))\,[-\log(1 - R_i/(n+1))]$$

where $g(t) = (1-t)^{-1}\int_{-\log(1-t)}^{\infty} x^{-1}e^{-x}\,dx$.

Large values are significant.

The alternative densities to be considered are listed below for $\lambda = 1$; to obtain the general form, make the transformation $f(x) \to \lambda f(\lambda x)$.

$$f_\theta^{(1)}(x) = [1 + \theta(1 - e^{-x})] \exp\{-[x + \theta(x + e^{-x} - 1)]\} \qquad \text{(Makeham)},$$

$$f_\theta^{(2)}(x) = (1 + \theta x) \exp\{-(x + \tfrac{1}{2}\theta x^2)\} \qquad \text{(linear F.R.)},$$

$$f_\theta^{(3)}(x) = (1 + \theta)x^\theta \exp\{-x^{(1+\theta)}\} \qquad \text{(Weibull)},$$

$$f_\theta^{(4)}(x) = [x^\theta e^{-x}]/\Gamma(1 + \theta) \qquad \text{(Gamma)}.$$

For each density, $x \geqq 0$, $\theta \geqq 0$; and the null hypothesis is obtained for $\theta = 0$. Each of these densities have increasing failure rates for $\theta > 0$. Note that $f_\theta^{(2)}$

has the linear FR (failure rate) $1 + \theta x$, while $f_\theta^{(1)}$ has the failure rate $1 + \theta(1 - e^{-x})$. Our Makeham distribution is a special case of the general Makeham distribution which has a failure rate of the form $a + b\exp(cx)$.

From Theorem 3.1 it follows that S_i is asymptotically most powerful for $f_\theta^{(i)}$ when λ is known, $i = 1, 2, 3, 4$. Theorem 5.1 implies that W_i is asymptotically most powerful for $f_\theta^{(i)}$ in the class of linear rank statistics, $i = 1, 2, 3, 4$.

W_0 is asymptotically equivalent to the Proschan-Pyke statistic V_n (see Remark 2.1). It is uniformly improved asymptotically by W_1. We now list the efficiencies of W_i to the asymptotically most powerful statistic of Section 3. The efficiencies are given in general as functions of $h(x) = h_1(x) = [\partial \log f_{\theta,1}(x)/\partial\theta]_{\theta=0}$ They are always independent of λ.

$$
\begin{aligned}
e(W_0) &= 9[\textstyle\int_0^\infty h'(x)e^{-x}(\frac{1}{2}x - 1 + e^{-x})\,dx]^2\sigma^{-2}(h)\\
&= 9[\textstyle\int_0^\infty h(x)e^{-x}(\frac{1}{2}x - \frac{3}{2} + 2e^{-x})\,dx]^2\sigma^{-2}(h),\\
e(W_1) &= \tfrac{4}{3}e(W_0),\\
e(W_2) &= [\textstyle\int_0^\infty h'(x)e^{-x}(\frac{1}{2}x^2 - x)\,dx]^2\sigma^{-2}(h),\\
e(W_3) &= [\textstyle\int_0^\infty h'(x)e^{-x}x(\log x + \gamma - 1)\,dx]^2(\frac{1}{6}\pi^2 - 2\gamma + 1)^{-1}\sigma^{-2}(h),\\
e(W_4) &= \textstyle\int_0^\infty x^{-1}e^{-x}\int_0^{1-e^{-x}}(1 - t)^{-1}(a(t) - \bar{a})\,dt\,dx(\frac{1}{6}\pi^2 - 1)^{-1}\sigma^{-2}(h),
\end{aligned}
$$

where $\gamma \doteq .5772$ is Euler's constant, $\sigma^2(h) = \int_0^\infty h^2(x)e^{-x}\,dx$, $a(t)$ is defined by (4.2), and $\bar{a} = \int_0^1 a(t)\,dt$.

As remarked in Section 5, if

$$
S_i^* = S_i/\textstyle\sum_{i=1}^n D_j, \qquad i = 1, 2, 3, 4, \tag{6.3}
$$

are the studentized linear spacings statistics, then S_i^* have the same efficiency as W_i, i.e. $e(S_i^*) = e(W_i)$, $i = 1, 2, 3, 4$. These efficiencies are given in Table 6.1.

The last row in this table gives the factor by which the efficiencies have to be multiplied in order to obtain the efficiencies with respect to the best linear rank tests (or the best studentized linear spacings tests). The efficiencies agree with those given by Proschan and Pyke (1967) and Lewis (1965) (where comparable).

Tables 6.2 deal with a Monte Carlo study of the powers of the W_i and S_i^* statistics for the linear F.R., Weibull and Gamma distributions, $1 \leqq i \leqq 3$.

From the tables, the following conclusions are apparent: (1) The rank tests are uniformly less powerful than the corresponding studentized linear spacings tests. (2) Of all the tests considered in the above tables, the total time on test statistic S_1^* and the "Weibull optimal" statistic S_3^* are generally best on the basis of both asymptotic efficiency and Monte Carlo power. (3) On the basis of Monte Carlo power alone, S_3^* is best. Barlow [1] has shown that for Weibull and Gamma alternatives, S_1^* is much better than the IFR likelihood ratio test and slightly worse than the IFRA likelihood ratio test (on the basis of Monte Carlo power).

The agreement between the asymptotic theory and the Monte Carlo power

TABLE 6.1

$e(W_i) = e(S_i^*)$

Statistic	$f_\theta^{(i)}$ Makeham	Linear F. R	Weibull	Gamma
W_0	$\frac{3}{16} = .1875$	$\frac{9}{32} \doteq .2813$	$9(\log 2)^2(4\gamma_0)^{-1} \doteq .5927$	$54(\log 2 - \frac{1}{2})^2\pi^{-2} \doteq .2041$
W_1	$\frac{1}{4} = .25$	$\frac{3}{8} = .375$	$3(\log 2)^2\gamma_0^{-1} \doteq .7903$	$72\pi^{-2}(\log 2 - \frac{1}{2})^2 \doteq .2721$
W_2	$\frac{3}{16} = .1875$	$\frac{1}{2} = .50$	$\gamma_0^{-1} \doteq .5483$	$\frac{3}{2}\pi^{-2} \doteq .1520$
W_3	$\frac{3}{4}(\log 2)^2(\gamma_0 - \gamma^2)^{-1} \doteq .2416$	$\frac{1}{2}(\gamma_0 - \gamma^2)^{-1} \doteq .3355$	$1 - \gamma^2\gamma_0^{-1} \doteq .8173$	$6\gamma^2(\pi^2 - 6)\pi^{-4}\gamma_0^{-1} \doteq .0435$
W_4	$18(\log 2 - \frac{1}{2})^2(\pi^2 - 6)^{-1} \doteq .1735$	$\frac{3}{4}(\pi^2 - 6)^{-1} \doteq .1938$	$6\gamma^2(\gamma_0 - \gamma^2)\ \pi^{-2}\gamma_0^{-2} \doteq .0908$	$1 - 6\pi^{-2} \doteq .3921$
Rank efficiency factor	4	2	$\gamma_0(\gamma_0 - \gamma^2)^{-1} \doteq 1.2235$	$\pi^2(\pi^2 - 6)^{-1} \doteq 2.5505$

$\gamma \doteq .5772 \ldots =$ Euler's constant, $\gamma_0 = (\gamma - 1)^2 + \pi^2 \doteq 1.8237$.

TABLE 6.3

Monte Carlo power based on 1,000 trials for samples of size 30 from the Weibull distribution, for significance levels α = .01, .05 and .10

α	.01	.05	.10	.01	.05	.10	.01	.05	.10	.01	.05	.10	.01	.05	.10	01	.05	.10
$\theta + 1$	1.10			1.25			1.40			1.75			2.00			2.50		
W_1	.056	.161	.262	.180	.386	.552	.365	.642	.783	.797	.942	.977	.916	.986	.999	.973	.992	.999
S_1^*	.045	.165	.268	.195	.416	.565	.441	.706	.820	.923	.985	.995	.995	1.00	1.00	1.00	1.00	1.00

TABLE 6.2

Monte Carlo power based on 2,000 trials for samples of size 10 from Linear F. R., Weibull and Gamma distributions for significance level $\alpha = .01, .05$ *and* $\cdot 1$

LINEAR F. R.

α	.01	.05	.10	.01	.05	.10	.01	.05	.10	.01	.05	.10	.01	.05	.10	.01	.05	.10
θ		.50			1.00			2.50			4.00			6.00			10.00	
W_1	.022	.104	.198	.032	.136	.263	.061	.211	.359	.068	.248	.415	.085	.300	.478	.094	.334	.518
W_2	.022	.105	.190	.032	.126	.236	.053	.201	.342	.059	.222	.387	.074	.263	.435	.077	.294	.482
W_3	.016	.109	.200	.022	.153	.261	.047	.226	.361	.054	.259	.420	.066	.322	.470	.070	.350	.516
S_1^*	.026	.105	.198	.046	.159	.265	.082	.262	.394	.101	.302	.453	.121	.350	.522	.152	.414	.582
S_2^*	.019	.112	.179	.036	.157	.249	.064	.263	.369	.080	.295	.421	.090	.341	.485	.117	.405	.548
S_3^*	.028	.109	.206	.048	.156	.271	.079	.249	.404	.106	.306	.468	.120	.342	.523	.164	.408	.579

WEIBULL

α	.01	.05	.10	.01	.05	.10	.01	.05	.10	.01	.05	.10	.01	.05	.10	.01	.05	.10	.01	.05	.10	.01	.05	.10
$\theta + 1$		1.25			1.50			1.75			2.00			2.50			3.00			3.50			4.50	
W_1	.029	.136	.251	.073	.280	.477	.130	.420	.621	.217	.562	.757	.318	.727	.880	.405	.805	.932	.482	.832	.943	.538	.863	.954
W_2	.021	.112	.227	.055	.255	.426	.106	.340	.538	.174	.477	.670	.252	.605	.791	.317	.675	.846	.369	.731	.869	.403	.774	.895
W_3	.023	.150	.252	.059	.318	.506	.107	.470	.672	.188	.617	.800	.286	.797	.924	.388	.879	.963	.470	.702	.971	.529	.927	.985
S_1^*	.039	.149	.269	.112	.337	.507	.238	.546	.719	.400	.734	.869	.721	.941	.982	.905	.994	.999	.977	.999	1.00	1.00	1.00	1.00
S_2^*	.025	.146	.226	.079	.306	.448	.168	.513	.657	.310	.686	.813	.621	.916	.967	.837	.989	.997	.946	.998	1.00	.999	1.00	1.00
S_3^*	.043	.161	.277	.124	.352	.533	.270	.566	.736	.434	.754	.883	.740	.945	.986	.909	.992	.999	.977	.999	1.00	.999	1.00	1.00

GAMMA

α	.01	.05	.10	.01	.05	.10	.01	.05	.10	.01	.05	.10	.01	.05	.10	.01	.05	.10	.01	.05	.10	.01	.05	.10
$\theta + 1$		1.50			2.00			2.50			3.00			3.50			4.00			4.50			5.50	
W_1	.036	.162	.288	.068	.275	.455	.110	.375	.583	.159	.447	.665	.195	.553	.733	.214	.576	.775	.232	.603	.797	.264	.628	.810
W_2	.036	.144	.249	.055	.225	.386	.084	.287	.482	.116	.356	.555	.142	.410	.600	.157	.449	.635	.158	.454	.664	.184	.498	.678
W_3	.023	.178	.312	.055	.323	.497	.093	.452	.659	.138	.548	.747	.176	.635	.814	.204	.691	.876	.221	.720	.879	.258	.746	.896
S_1^*	.045	.175	.298	.096	.328	.492	.203	.502	.680	.321	.647	.809	.434	.775	.893	.541	.864	.953	.651	.905	.960	.803	.971	.993
S_2^*	.033	.161	.251	.063	.283	.415	.131	.434	.577	.226	.570	.718	.300	.695	.803	.395	.782	.886	.470	.841	.915	.640	.922	.962
S_3^*	.052	.190	.323	.121	.365	.553	.238	.559	.748	.364	.707	.859	.509	.834	.929	.622	.906	.974	.719	.934	.983	.864	.987	.998

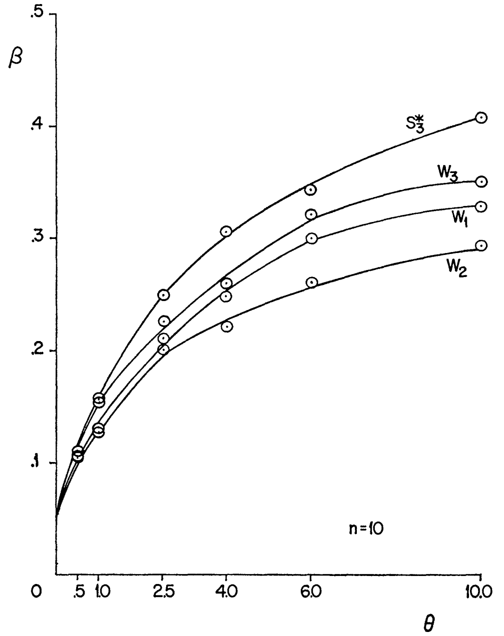

FIG. 1. Power curves for linear F.R. alternatives

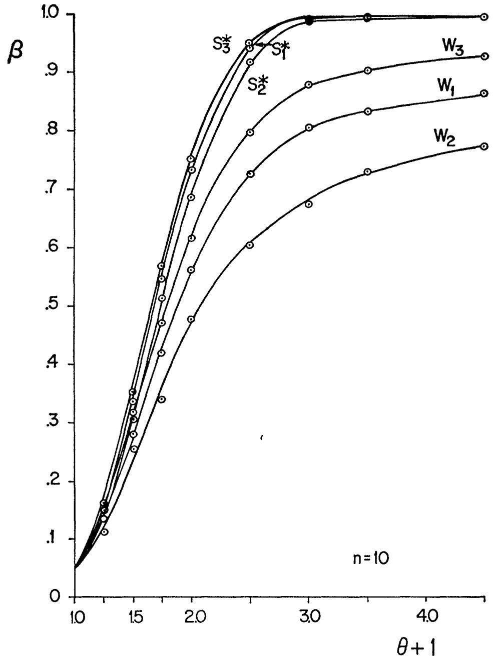

FIG. 2. Power curves for Weibull alternatives

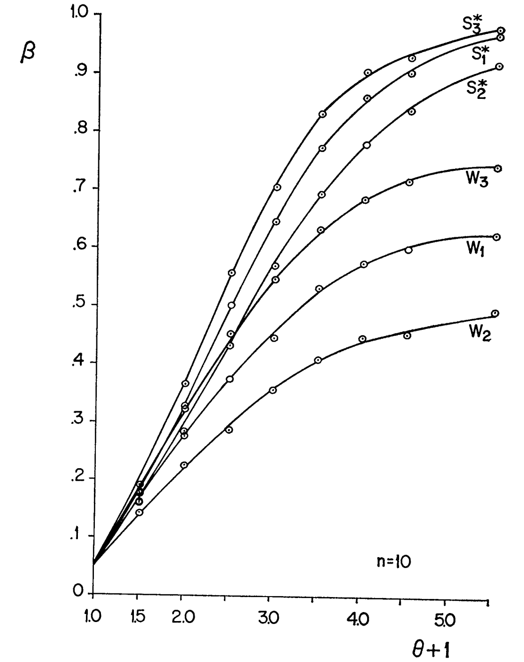

FIG. 3. Power curves for gamma alternatives

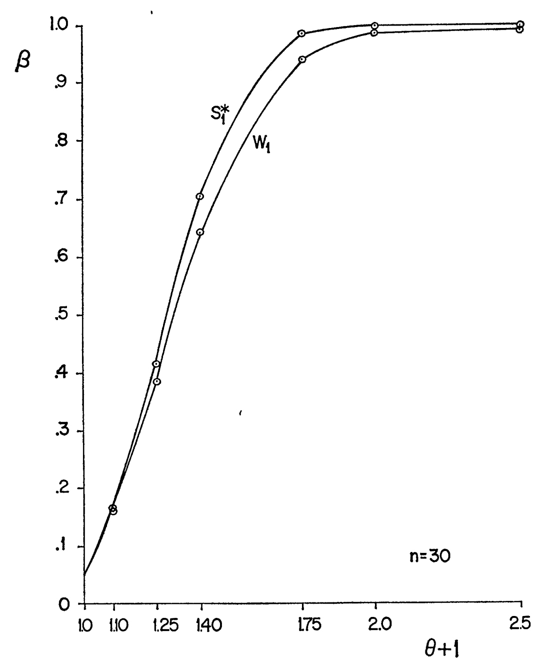

FIG. 4. Power curves for Weibull alternatives

with $n = 10$ is poor. This led us to compute Table 6.3 in which $n = 30$ and the Weibull distribution is considered. This table shows that for small θ, there is no difference in the powers between the total time on test statistic S_1^* and its rank counterpart W_1, while for larger θ, the total time on test statistic is again better. However, as expected from the asymptotic theory, the difference in power seems to be decreasing.

The results of the Monte-Carlo study are shown graphically for $\alpha = .05$ in Figures 1–4. On comparing Figures 2 and 3 with Figure 1 of Barlow (1967), it is seen that the power of S_3^* is about the same as that of his best test, namely the "likelihood ratio statistic for IFRA." Note that in Figure 1, the graphs for S_1^* and S_2^* are omitted since they would essentially coincide with the graph for S_3^*.

7. Appendix: Proof of Theorem 4.1: We use the notation of Section 4. Heuristically our argument is very simple and is essentially the same as that used in Chernoff, Gastwirth, and Johns (1967).

$$\begin{aligned} T_n &= n^{-\frac{1}{2}}\{\textstyle\sum_{i=1}^n [h(X_{(i)}) - E[h(X_{(i)})]\} \\ (7.1) \quad &\doteq n^{-\frac{1}{2}}\{\textstyle\sum_{i=1}^n [h(X_{(i)}) - h(E(X_{(i)}))]\} \\ &\doteq n^{-\frac{1}{2}}\{\textstyle\sum_{i=1}^n h'(E(X_{(i)}))[X_{(i)} - E(X_{(i)})]\} \\ &\doteq n^{-\frac{1}{2}}\textstyle\sum_{i=1}^n h'(-\log(1 - i/(n+1)))\sum_{j=1}^i (D_j - 1)(n - j + 1)^{-1} \end{aligned}$$

under H_0. From the last approximate identity our result follows. The justification of these approximations poses some minor technical difficulties. We proceed with the proof of the theorem. We show that i, ii and iii b) suffice. Let,

$$(7.2) \qquad \begin{aligned} J_\delta(t) &= 1, \quad \delta \leqq t \leqq 1 - \delta \\ &= 0, \quad \text{otherwise.} \end{aligned}$$

From Theorem 3 of [6] it follows that, if $T_n^\delta = n^{-\frac{1}{2}}\sum_{i=1}^n J_\delta(i/(n+1))h(X_{(i)})$, then

$$(7.3) \qquad T_n^\delta = n^{\frac{1}{2}}\mu_n^\delta + n^{-\frac{1}{2}}\textstyle\sum_{j=1}^n a^\delta(j(n+1)^{-1})(D_j - 1) + o_p(1),$$

where

$$(7.4) \qquad a^\delta(s) = (1 - s)^{-1}\textstyle\int_s^1 J_\delta(t)h'(-\log(1 - t))\,dt,$$

$$(7.5) \qquad \mu_n^\delta = n^{-1}\textstyle\sum_{j=1}^n J_\delta(i/(n+1))h(-\log(1 - i(n+1)^{-1}))$$

and $o_p(1)$ as usual denotes a remainder converging to 0 in probability.

Now let

$$(7.6) \qquad \gamma_n^\delta = -n^{-\frac{1}{2}}\textstyle\sum_{j=1}^n J_\delta(i/(n+1))E(h(X_{(j)})),$$

$$(7.7) \qquad R_{n1}^\delta = n^{-\frac{1}{2}}\textstyle\sum_{j<(n+1)\delta}\{h(X_{(j)}) - E(h(X_{(j)}))\},$$

and

$$(7.8) \qquad R_{n2} = n^{-\frac{1}{2}}\textstyle\sum_{j<(n+1)(1-\delta)}\{h(X_{(j)}) - E(h(X_{(j)}))\}.$$

Then,

(7.9) $$T_n = T_n^{\delta} + R_{n1}^{\delta} + R_{n2}^{\delta} + \gamma_n^{\delta}.$$

We begin with,

LEMMA 7.1 *If* $\delta_n \to 0$ *as* $n \to \infty$, *then*

(7.10) $$n^{-\frac{1}{2}} \sum_{i=1}^{n} [a^{\delta_n}(i/(n+1)) - a(i/(n+1))](D_i - 1) \to 0$$

in probability.

PROOF.

(7.11) $$E\{n^{-\frac{1}{2}} \sum_{i=1}^{n} [a^{\delta_n}(i/(n+1)) - a(i/(n+1))](D_i - 1)\}^2 = n^{-1} \sum_{j=1}^{n} [a^{\delta_n}(i/(n+1)) - a(i/(n+1))]^2,$$

(7.12) $$|a^{\delta}(s)|^2 \leqq (1-s)^{-2}[\int_s^1 |h'(-\log(1-t))|\, dt]^2 \quad \text{for} \quad \delta \geqq 0.$$

Moreover,

(7.13) $$\int_0^1 (1-s)^{-2}(\int_s^1 |h'(-\log(1-t))|\, dt)^2\, ds < \infty.$$

To see (7.13) note that the left-hand side equals

(7.14) $$\int_0^1 (1-s)^{-2} \int_s^1 |h'(-\log(1-t))|\, dt \int_s^1 |h'(-\log(1-v))|\, dv\, ds = 2\int_0^\infty \int_0^u |h'(u)| \cdot |h'(r)| e^{-u}(1-e^{-r})\, dr\, du$$

after some standard arguments. Now by (c) there exists $\delta > 0$ such that $h'(x)$ has constant sign for $x < \delta$, $x > \delta^{-1}$ and $\sup_{\delta \leqq x \leqq \delta^{-1}} |h'(x)| < \infty$. Then

$$\int_\delta^{\delta^{-1}} \int_\delta^u |h'(u)| \cdot |h'(r)| e^{-u}(1-e^{-r})\, dr\, du < \infty.$$

Suppose for simplicity $h'(x) \geqq 0$, $x < \delta$; $h'(x) \leqq 0$, $x > \delta^{-1}$. Then

(7.15) $$\begin{aligned} &\int_\delta^{\delta^{-1}} \int_0^\delta |h'(u)|\, |h'(r)| e^{-u}(1-e^{-r})\, du\, dr \\ &= \int_\delta^{\delta^{-1}} |h'(u)| e^{-u} \int_0^\delta h'(r)(1-e^{-r})\, dr\, du \\ &= \int_\delta^{\delta^{-1}} |h'(u)| \{\int_0^\delta h(r) e^{-r}\, dr + h(\delta)(1-e^{-\delta})\}\, du < \infty \end{aligned}$$

since $\int_0^\infty e^{-r}|h(r)|\, dr < \infty$. Continuing,

(7.16) $$\begin{aligned} &2\int_0^\delta \int_0^u |h'(u)| \cdot |h'(r)| e^{-u}(1-e^{-r})\, du\, dr \\ &= \lim_{\lambda \to 0} [2\int_\lambda^\delta h'(u) e^{-u}\{h(u) - h(\lambda)\}\, du - \int_\lambda^\delta \{d[\int_\lambda^u e^{-v} h'(v)\, dv]^2\, du\}\, du] \\ &= \lim_{\lambda \to 0} [\int_\lambda^\delta e^{-u}\, dh^2(u) - 2h(\lambda)\int_\lambda^\delta h'(u) e^{-u}\, du - [\int_\lambda^\delta h'(u) e^{-u}\, du]^2] \\ &= \int_0^\delta h^2(u) e^{-u}\, du - h^2(\delta) e^{-\delta} - 2[\int_0^\delta e^{-u} h(u)\, du] h(\delta) e^{-\delta} \\ &\quad + \lim_{\lambda \to 0} \{h^2(\lambda)(1-e^{-\lambda}) + 2h(\lambda)(1-e^{-\lambda})\int_\lambda^\delta h(u) e^{-u}\, du \\ &\quad + 2h(\delta) e^{-\delta} h(\lambda)(1-e^{-\lambda})\} < \infty \end{aligned}$$

as long as $h^2(\lambda)(1-e^{-\lambda})$ is bounded as $\lambda \to 0$. But this, of course, holds if

$$\int_0^\epsilon h^2(u) e^{-u}\, du < \infty, \text{ and } h \uparrow \text{ in } (0, \delta).$$

One can dispose of the other pieces of the integral by similar arguments. From (7.11) and (7.12) it readily follows that

$$a^{\delta_n}(s) \to a(s) \tag{7.17}$$

for every $0 < s < 1$ and again by (7.12), (7.16) and the dominated convergence theorem the lemma follows. Suppose we can show

$$\lim_{\delta\to 0} \sup_n \operatorname{Var} R_{ni}^{\delta} = 0 \tag{7.18}$$

for $i = 1, 2$. We claim the theorem follows under our assumptions. To see this note that (7.9) and (7.18) imply that for every $\epsilon > 0$, there exists a δ such that

$$\limsup_n d(T_n, T_n^{\delta} + \gamma_n^{\delta}) \leqq \epsilon \tag{7.19}$$

where

$$d(X, Y) = E\{|X - Y|(1 + |X - Y|)^{-1}\} \tag{7.20}$$

is the usual metric for convergence in probability. Now, (7.3) yields for δ as above,

$$\limsup_n d(T_n, n^{-\frac{1}{2}}\textstyle\sum_{i=1}^n a^{\delta}(i(n+1)^{-1})(D_i - 1) + n^{\frac{1}{2}}\mu_n^{\delta} + \gamma_n^{\delta}) \leqq \epsilon, \tag{7.21}$$

and by Lemma 7.1, for δ sufficiently small

$$\limsup_n d(T_n, S_n + n^{\frac{1}{2}}\mu_n^{\delta} + \gamma_n^{\delta}) \leqq \epsilon. \tag{7.22}$$

This, of course, implies that there exists a sequence of constants K_n such that,

$$\lim_n d(T_n, S_n + K_n) = 0. \tag{7.23}$$

Since S_n and T_n are both asymptotically normal with mean 0 by Lemma 4.1 of Bickel (1967) and the central limit theorem, it follows that $K_n \to 0$. We prove (7.18) for $i = 1$; for $i = 2$ it is proved analogously. Let δ be defined as in the discussion preceding (7.15). Define,

$$\begin{aligned} Z_i &= h(X_i) && \text{if } 0 \leqq X_i < -\log(1-\delta) \\ &= U_i && \text{otherwise,} \end{aligned} \tag{7.24}$$

where the U_i are uniformly distributed on $(-\log(1-\delta), B)$ and independent of each other and of the X_i. B is so chosen that the density of Z_1 at $h(-\log(1-\delta))$ is the same as the density of $h(X_i)'$ at $h(-\log(1-\delta))$.

We argue as in [4]. Let $Z_{(1)} < \cdots < Z_{(n)}$ be the order statistics of the Z_i's. Then,

$$\begin{aligned} &E(\textstyle\sum_{i\leqq\delta n/2}(h(X_{(i)}) - Z_{(i)}))^2 \\ &\quad= E(\textstyle\sum_{i\leqq\delta n/2}(h(X_{(i)}) - Z_{(i)})I_{[X(i)\geqq-\log(1-\delta)]})^2 \\ &\quad\leqq 2E(\textstyle\sum_{i\leqq\delta n/2} Z_{(i)}I_{[X(i)\geqq-\log(1-\delta)]})^2 \\ &\quad\quad+ 2\textstyle\sum_{i\leqq\delta n/2}[E\{h^2(X_{(i)})I_{[X(i)\geqq-\log(1-\delta)]}\}]^{\frac{1}{2}}. \end{aligned} \tag{7.25}$$

The identity follows by definition of the Z_i, the inequality from the c_r and Minkowski inequalities. Now,

$$(7.26)\quad E\Big(\sum_{i\leq \delta n/2} Z_{(i)} I_{[X(i)\geqq -\log(1-\delta)]}\Big)^2 \leqq B^2 \tfrac{1}{4}\delta^2 n^2 P[X_{[\delta n/2]} > -\log(1-\delta)] \to 0$$

since by Lemma 2.2 of [4], $P[X_{[\delta n/2]} \geqq -\log(1-\delta)] \to 0$ exponentially.

On the other hand,

$$(7.27)\quad \begin{aligned} &E(h^2(X_{(i)}) I_{[X(i)\geqq -\log(1-\delta)]}) \\ &= \int_{v\geqq -\log(1-\delta)} h^2(v) n![(i-1)!(n-i)!]^{-1} e^{-(n-i+1)v} \cdot (1-e^{-v})^{i-1}\, dv \\ &\leqq \sup_{v\geqq -\log(1-\delta)} e^{-(n-i)v}(1-e^{-v})^{i-1} n\tbinom{n-1}{i-1} E(h^2(X_1)). \end{aligned}$$

Now,

$$(7.28)\quad \sup_{v\geqq -\log(1-\delta)} e^{-(n-i)v}(1-e^{-v})^{i-1} n\tbinom{n-1}{i-1} = \delta^{i-1}(1-\delta)^{n-i} n\tbinom{n-1}{i-1} \quad \text{for} \quad i \leqq \tfrac{1}{2}\delta n.$$

Finally,

$$(7.29)\quad \sup_{i\leqq \delta n/2} \delta^{i-1}(1-\delta)^{n-i} n\tbinom{n-1}{i-1} = \delta^{[\delta n/2]-1}(1-\delta)^{n-[\delta n/2]} n\tbinom{n-1}{i-1}$$

by an easy induction argument on i. By Lemma 2.2 of [4] the right-hand side of (7.29) $\to 0$ exponentially. Then, (7.27), (7.28) and (7.29) imply that

$$\lim_{\delta'\to 0} \limsup_n \operatorname{Var} R_{n1}^{\delta'} = 0$$

if and only if

$$(7.30)\quad \lim_{\delta'\to 0} \limsup_n \operatorname{Var} \Big[\sum_{i\leqq \delta' n} Z_{(i)}\Big] = 0,$$

where Z_i are defined for fixed δ. Since the Z_i's are independently and indentically distributed with a density positive on its convex support and $E(Z_1^2) < \infty$, (7.30) follows from [4]. The sufficiency of (i), (ii) and (iii) (b) for the conclusion of the theorem follows. The sufficiency of (i), (ii), and (iii) (a) is an easy consequence of Theorems of [6].

Using the methods of [4] on T_n^δ one can show that $E(S_n - T_n)^2 \to 0$.

As in [6] the smoothness conditions on h may be relaxed.

Acknowledgment. We would like to thank Mr. Vasant Ubhaya who obtained the figures in Tables 6.2 and 6.3 for us, and the referee for his conscientious reading and contributions to Section 2.

REFERENCES

[1] Barlow, R. E. (1967) Likelihood ratio tests for restricted families. *Ann. Math. Statist* **39** 547–560.

[2] Barlow, R. E. and Proschan, F. (1966). Inequalities for linear combinations of order statistics from restricted families. *Ann. Math. Statist.* **37** 1574–1591.

[3] BELL, C. B. and DOKSUM, K. (1967). Distribution-free tests of independence. *Ann. Math. Statist.* **38** 429–446.

[4] BICKEL, P. J. (1967). Some contributions to the theory of order statistics. *Proc. Fifth Berkeley Symp. Math. Statist. Prob.* **1** Univ. of California Press.

[5] BICKEL, P. J. (1969). Tests for monotone failure rate II. *Ann. Math. Statist.* **40**

[6] CHERNOFF, H., GASTWIRTH, J. and JOHNS, M. V. (1967). Asymptotic distribution of linear combinations of functions of order statistics with application to estimation. *Ann. Math. Statist.* **38** 52–73.

[7] HÁJEK, J. (1961). Some extensions of the Wald-Wolfowitz-Noether theorems. *Ann. Math. Statist.* **32** 506–523.

[8] HÁJEK, J. (1969). *A Course in Nonparametric Statistics*. Holden-Day, San Francisco.

[9] HÁJEK, J. and ŠIDÁK, Z. (1967). *Theory of Rank Tests*. Academic Press, New York.

[10] HODGES, J. L. and LEHMANN, E. L. (1961). Comparison of the normal scores and Wilcoxon tests. *Proc. Fourth Berkeley Symp. Math. Statist. Prob.* Univ. of California Press.

[11] HOEFFDING, W. (1951). "Optimum" nonparametric tests. *Proc. Second Berkeley Symp. Math. Statist. Prob.* Univ. of California Press.

[12] LECAM, L. (1966). Likelihood functions for large numbers of independent observations. *Festschrift for J. Neyman*. Wiley New York.

[13] LEHMANN, E. L. (1955). Ordered families of distributions. *Ann. Math. Statist.* **26** 399–419.

[14] LEHMANN, E. L. (1966). Some concepts of dependence. *Ann. Math. Statist.* **37** 1137–1154.

[15] LEWIS, P. A. W. (1965). Some results on tests for Poisson processes. *Biometrika* **52** 67–77.

[16] NADLER, J. and EILBOTT, J. (1967). Testing for monotone failure rates. (Unpublished).

[17] PROSCHAN, R. and PYKE, R. (1967). Tests for monotone failure rate. *Proc. Fifth Berkeley Symp. Math. Statist. Prob.* **3**. Univ. of California Press.

[18] VAN ZWET, W. (1964). *Convex Transformations of Random Variables*. Math Centrum, Amsterdam.

[19] WALD, A. (1941). Asymptotically most powerful tests of statistical hypotheses. *Ann. Math. Statist.* **12 1–19.**