RANK TESTS FOR RANDOMIZED BLOCKS WHEN THE ALTERNATIVES HAVE AN A PRIORI ORDERING[1]

By Myles Hollander

The Florida State University

0. Summary. Let X_{ij}, $i = 1, \cdots, n$, $j = 1, \cdots, k$, be independent with X_{ij} having the continuous distribution function $P(X_{ij} \leqq x) = F_j(x - b_i)$ where b_i is the nuisance parameter corresponding to block i. (These assumptions shall be called the H_A assumptions.) This paper is concerned with procedures for testing the null hypothesis

$$H_0 : F_j = F \text{ (unknown)}, \qquad j = 1, \cdots, k, \tag{0.1}$$

which are sensitive to the ordered alternatives

$$H_a : F_1 \geqq F_2 \geqq \cdots \geqq F_k, \tag{0.2}$$

where at least one of the inequalities is strict.

In particular, we introduce a test statistic (Y) based on a sum of Wilcoxon signed-rank statistics. In Section 2 we develop the asymptotic distribution of Y and find that, under H_0, Y is neither distribution-free for finite n, nor asymptotically distribution-free. However, a consistent estimate of the null variance of Y is used to define a procedure which is asymptotically distribution-free.

In Section 3 we derive, under the H_A assumptions, necessary and sufficient conditions for the consistency of Y and two of its nonparametric competitors, viz., (1) Jonckheere's τ test [11] based on Kendall's rank correlation coefficient between observed order and postulated order in each block; (2) Page's ρ test [17] based on Spearman's rank correlation coefficient between observed order and postulated order in each block. We find that (i) Y is consistent if and only if $\sum_{u<v} \int H_u \, dH_v/k(k-1) > \frac{1}{4}$ where $H_u = F_u{}^*F_u$, $u = 1, \cdots, k$, (ii) Jonckheere's test is consistent if and only if $\sum_{u<v} \int F_u \, dF_v/k(k-1) > \frac{1}{4}$, and (iii) Page's test is consistent if and only if $\sum_{u<v} (v-u) \int F_u \, dF_v > k(k-1) \cdot (k+1)/12$.

Section 4 is devoted to efficiency comparisons of the rank tests with respect to a normal theory t-test defined in Section 1. For a class of shift alternatives we show that the Pitman efficiency of Y with respect to t $(E(Y, t))$ is greater than .864 for every F and every k. When F is normal, $E(Y, t) = .963$ for $k = 3$ and $\rightarrow .989$ as $k \rightarrow \infty$. These values compare favorably with the corresponding ones of Page's test (.716, .955) and Jonckheere's procedure (.694, .955). For

Received 15 October 1965; revised 28 November 1966.

[1] This paper is based on the author's doctoral dissertation which was written at Stanfrod University with the support of Public Health Service Grant USPHS-5T1 GM 25-07. Additional support was provided by NASA Grant NGR-10-004-029 at F.S.U.

these shift alternatives we also show that $.576 \leqq E(\rho, t) \leqq \infty$ and $.576 \leqq E(\tau, t) \leqq \infty$.

1. Introduction and definitions. To test (0.1) versus (0.2), the following statistics, among others, have been proposed:

(i) (Jonckheere [11]) Let τ_i denote Kendall's rank correlation coefficient between postulated order and observation order in the ith block. Jonckheere's procedure is to reject H_0 for large values of

$$\tau = \sum_{i=1}^{n} \tau_i . \tag{1.1}$$

(ii) (Page [17]) This test is similar to Jonckheere's as Page suggests a rejection region consisting of large values of

$$\rho = \sum_{i=1}^{n} \rho_i \tag{1.2}$$

where ρ_i is Spearman's rank correlation coefficient between postulated order and observation order in block i.

(iii) Let $X_{ij} = b_i + (j-1)\theta + \epsilon_{ij}$ where the ϵ_{ij} are independent and identically distributed according to $N(0, \sigma^2)$. The likelihood ratio statistic for testing $\theta = 0$ is

$$t = \hat{\theta}/\hat{\sigma}_{\hat{\theta}} \tag{1.3}$$

where $\hat{\theta}$ is the least squares estimate of θ and $\hat{\sigma}_{\hat{\theta}}$ is the appropriate estimate of the standard deviation of $\hat{\theta}$. Specifically,

$$\hat{\theta} = 6\sum_{i=1}^{n}\sum_{j=1}^{k} (2j - k - 1)X_{ij}/nk(k-1)(k+1) \quad \text{and}$$
$$\hat{\sigma}_{\hat{\theta}}^2 = 12s^2/k(k-1)(k+1)n$$

where

$$s^2 = \sum_{i=1}^{n}\sum_{j=1}^{k} (X_{ij} - \hat{b}_i - (j-1)\hat{\theta})^2/n(k-1) - 1 \quad \text{and}$$
$$\hat{b}_i = (\sum_{j=1}^{k} X_{ij}/k) - (k-1)\hat{\theta}/2.$$

The analogue of (1.3) for the one-way layout has been considered by Hogg [8].

We do not list standard tests of H_0 such as Friedman's rank test [5] and the usual normal theory F test as they guard against more general alternatives and do not take the prior ordering into account.

Now we proceed to define the Y statistic. Let $Y_{uv}^{(i)} = |X_{iu} - X_{iv}|$ and $R_{uv}^{(i)} =$ rank of $Y_{uv}^{(i)}$ in the ranking from least to greatest of $[Y_{uv}^{(i)}]_{i=1}^{n}$. Furthermore, let

$$T_{uv} = \sum_{i=1}^{n} R_{uv}^{(i)}\psi_{uv}^{(i)} \tag{1.4}$$

where

$$\psi_{uv}^{(i)} = 1 \quad \text{if } X_{iu} < X_{iv}$$
$$= 0 \quad \text{otherwise.} \tag{1.5}$$

The statistic

$$Y =_{\text{def}} \sum_{u<v} T_{uv} \tag{1.6}$$

is proposed as one which will be sensitive to the ordered alternatives H_a. Here, T_{uv} is a measure of the difference between the uth and vth treatments, and the particular summation $\sum_{u<v}$ takes into account the prior ordering of the treatments. It should be mentioned that in this respect, τ, ρ, and t are similar in character to Y. In fact, if we let $Z_\tau = \sum_{u<v} N_{uv}$ where $N_{uv} = \sum_{i=1}^{n} \psi_{uv}^{(i)}$, and $Z_\rho = \sum_{u<v} (R_v - R_u)$ where $R_u = \sum_{i=1}^{n} r_{iu}$ and r_{iu} is the rank of X_{iu} in the joint ranking of $[X_{i\alpha}]_{\alpha=1}^{k}$, then it is easily seen that $\tau = (4Z_\tau/k(k-1)) - n$ and $\rho = 6Z_\rho/(k^3 - k)$. Hence, the tests τ, ρ, t and Y all employ the summation $\sum_{u<v}$ with N_{uv}, $(R_v - R_u)$, $(X_{\cdot v} - X_{\cdot u})$, and T_{uv}, respectively, being plausible measures of the difference between the uth and vth treatments.

2. The asymptotic distribution of Y.

LEMMA 2.1. *Assume* $P(X_{ij} \leqq x) = F_j(x - b_i)$ *and* $0 < \int F_u\, dF_v < 1$ *for each* (u, v) *pair. Then the* $[T_{uv}]$, $1 \leqq u < v \leqq k$ *have an asymptotic joint* $k(k-1)/2$-*variate normal distribution.*

PROOF. Consider the n vectors $X_\alpha' = (X_{\alpha 1}, X_{\alpha 2}, \cdots, X_{\alpha k})$, $\alpha = 1, \cdots, n$, and write T_{uv} in the form of (3.3). (This representation is due to Tukey [19].) Then

$$\binom{n}{2}^{-1} T_{uv} = \binom{n}{2}^{-1} \sum_{i<j}^{n} \psi_{uv}(X_i', X_j') + \binom{n}{2}^{-1} \sum_{i=1}^{n} \psi_{uv}(X_i', X_i').$$

The joint asymptotic normality of the $[\binom{n}{2}^{-1} \sum_{i<j}^{n} \psi_{uv}(X_i', X_j')]$, $1 \leqq u < v \leqq k$, is a consequence of Hoeffding's U-statistic theorem [7]. The vectors $[X_\alpha']_{\alpha=1}^{n}$ are not identically distributed as in Hoeffding's theorem but the block parameters $[b_\alpha]$ do not affect the result due to the nature of the ψ_{uv} functions. The proof is completed by noting that $p\text{-}\lim n^{\frac{1}{2}} \binom{n}{2}^{-1} \sum_{i=1}^{n} \psi_{uv}(X_i', X_i') = 0$.

Since Y is a linear combination of the $[T_{uv}]$ we may state

THEOREM 2.1. *If* $0 < \int F_u\, dF_v < 1$ *for at least one* (u, v) *pair, then* Y, *suitably normed, has an asymptotic normal distribution.*

Unlike the rank statistics τ and ρ, Y is not distribution-free under H_0 despite the fact that each T_{uv} enjoys this property. This can be seen by verifying that the null correlation coefficient $\rho_0^n(F)$ between T_{uv} and T_{uw} ($u \neq v$, $u \neq w$, $v \neq w$) depends on F (except for $n = 1$), and hence so does the null variance of Y. In [9] it is shown that

$$(2.1) \quad \rho_0^n(F) = [(24\lambda(F) - 6)n^2 + (48\mu(F) - 72\lambda(F) + 7)n + (48\lambda(F) - 48\mu(F) + 1)][(n+1)(2n+1)]^{-1}$$

and

$$(2.1') \qquad \rho^*(F) = \lim_n \rho_0^n(F) = 12\lambda(F) - 3$$

where $\mu(F) = P(X_1 < X_2 ; X_1 < X_5 + X_6 - X_7)$ and $\lambda(F) = P(X_1 < X_2 + X_3 - X_4 ; X_1 < X_5 + X_6 - X_7)$ when $X_1, X_2, \cdots, X_7$ are independent and identically distributed according to F.

The well known expressions $E_0(T_{uv}) = n(n+1)/4$ and $\sigma_0^2(T_{uv}) = n(n+1) \cdot (2n+1)/24$ are readily obtained from (1.4), and it follows that

$$(2.2) \qquad E_0(Y) = k(k-1)n(n+1)/8$$

and

$$(2.3)\quad \sigma_0^2(Y) = n(n+1)(2n+1)k(k-1)(3+2(k-2)\rho_0^n(F))/144.$$

Equation (2.3) may be derived by writing

$$(2.4)\quad \sigma_0^2(Y) = \sum_{u<v} \sigma_0^2(T_{uv}) + 2\sum_{u<v<w} \operatorname{Cov}_0(T_{uv}, T_{uw}) + 2\sum_{u<v<w} \operatorname{Cov}_0(T_{uv}, T_{vw}) + 2\sum_{u<v<w} \operatorname{Cov}_0(T_{uw}, T_{vw}).$$

Terms of the form $\operatorname{Cov}_0(T_{uv}, T_{wx})$ where all subscripts are different do not appear in (2.4) since they are equal to zero due to the independence of T_{uv} and T_{wx}. Equation (2.3) is then obtained from (2.4) by employing the obvious symmetries and noting that $\operatorname{Cov}_0(T_{uv}, T_{uw}) = -\operatorname{Cov}_0(T_{uv}, T_{vw})$.

In view of (2.1′) and (2.3), in order that the test based on Y be asymptotically distribution-free, we require a consistent estimate of $\rho^*(F)$, or equivalently $\lambda(F)$. Lehmann [14] proposed estimating $\lambda(F)$ by the proportion (over all sextuples $(\alpha, \beta, \gamma; u, v, w)$) of cases in which the event $(X_{\alpha u} < X_{\alpha v} + X_{\beta u} - X_{\beta v};\ X_{\alpha u} < X_{\alpha w} + X_{\gamma u} - X_{\gamma w})$ occurs. As Lehmann mentions, this estimate is computationally tedious and in practice an estimate based on a small subset of the original number of sextuples should be used. In our situation we can utilize the prior ordering in deciding what subset of inequalities should be checked to estimate $\lambda(F)$. Specifically, consider an estimate of a slightly different form, namely let $\hat{\lambda}_1$ denote the relative frequency of the event $(X_{\alpha k} < X_{\alpha 1} + X_{\beta 1} - X_{\beta k};\ X_{\alpha k} < X_{\alpha 2} + X_{\gamma 1} - X_{\gamma k})$ over a subset of the total number of (α, β, γ) 3-tuples. Under H_a, the above system of inequalities would tend to be satisfied less frequently than a set which is symmetric in the column subscript, and hence we would be increasing the power of the Y test against H_a. We can also increase the power of the Y test by using an estimate of the form $\hat{\rho}_n = \min[\rho_u^n; \hat{\rho}]$ where $\hat{\rho}$ is a consistent estimate of $\rho^*(F)$ and ρ_u^n is an upper bound for $\rho_0^n(F)$. Upper bounds for $\rho_0^n(F)$ can be obtained by replacing $\mu(F)$ and $\lambda(F)$ in (2.1) with corresponding upper bounds. Lehmann proved $\lambda(F) \leqq \frac{7}{24}$ and we now establish

LEMMA 2.2. *For all* F, $\mu(F) \leqq ((2^{\frac{1}{2}} + 6)/24) = .3089$.

PROOF. Let $X_1, X_2, \cdots, X_n$ be a random sample with distribution F and let $Y_1, Y_2, \cdots, Y_n$ also be a random sample with distribution F. Let U_1 denote the Mann-Whitney-Wilcoxon statistic, $U_1 = \sum_{i=1}^n \sum_{j=1}^n \phi(X_i, Y_j)$ where $\phi(a, b) = 1$ if $a < b$, 0 otherwise, and let U_2 denote Wilcoxon's signed rank statistic applied to a random pairing of the X's with the Y's. Again using Tukey's representation (3.3), we write $U_2 = \sum_{i<j}^n \phi(X_i + X_j, Y_i + Y_j) + \sum_{i=1}^n \phi(X_i, Y_i)$. A direct calculation then shows that the correlation between U_1 and U_2 is given by

$$(2.5)\quad r^n(F) = [n^2(24\mu(F) - 6) + n(23 - 72\mu(F)) + (48\mu(F) - 14)] \{(2n+1)[n(n+1)/2]^{\frac{1}{2}}\}^{-1}$$

and

$$(2.6)\quad r^*(F) = \lim_n r^n(F) = (24\mu(F) - 6)/(2)^{\frac{1}{2}}.$$

The result follows since $r^*(F) \leqq 1$.

Table 2.1 contains values of ρ_u^n for various values of n.

A comparison of these values with those of $\rho_0^n(F)$ when F is rectangular (given in [9]) indicates that the upper bounds are quite good.

Thus, the proposed test is to reject H_0 at the α-level if $Y > E_0(Y) + z^{1-\alpha}\hat{\sigma}_0(Y)$ where $\hat{\sigma}_0^2(Y)$ is obtained by replacing $\rho_0^n(F)$ by $\hat{\rho}_n = \min [\rho_u^n; 12\hat{\lambda}_1 - 3]$ in (2.3), and $z^{1-\alpha}$ is the $1 - \alpha$ percentile point of a standardized normal random variable. If $n > 15$, defining $\hat{\rho}_n = \min [\frac{1}{2}; 12\hat{\lambda}_1 - 3]$ will suffice; if $\hat{\rho}_n < 0$, redefine it equal to 0.

3. Consistency of the rank tests. In this section we derive necessary and sufficient conditions for the consistency of the rank tests under the H_A assumptions. We note that the asymptotic normality of ρ and τ under H_A is a consequence of the central limit theorem. (The case where $\int F_u \, dF_v$ equals zero or one for every (u, v) pair is excluded, for in this case both τ and ρ (and Y) are constants with probability one.)

THEOREM 3.1. *A necessary and sufficient condition for the consistency of the test based on* τ *is* $\sum_{u<v} \int F_u \, dF_v/k(k-1) > \frac{1}{4}$.

PROOF. Using the representation $Z_\tau = \sum_{u<v} N_{uv}$, we have

$$E_A(Z_\tau) = n \sum_{u<v} \int F_u \, dF_v \tag{3.1}$$

and, in particular, $E_0(Z_\tau) = nk(k-1)/4$. It is obvious that $\sigma_A^2(Z_\tau) = O(n)$ and hence it follows from Chebychev's inequality that under H_0,

$$p\text{-lim } Z_\tau/nk(k-1) = \tfrac{1}{4} \tag{3.1$'$}$$

while under H_A,

$$p\text{-lim } Z_\tau/nk(k-1) = \sum_{u<v} \int F_u \, dF_v/k(k-1). \tag{3.1$''$}$$

The sufficiency is a consequence of (3.1′) and (3.1″) and the asymptotic normality of τ under H_0 and H_A insures that the condition is also necessary.

THEOREM 3.2. *A necessary and sufficient condition for the consistency of the test based on* ρ *is* $\sum_{j=1}^{k} (2j - k - 1)\cdot(\sum_{\alpha=1}^{k} \int F_\alpha \, dF_j) > 0$, *or equivalently,* $\sum_{u<v} (v - u) \int F_u \, dF_v > k(k-1)(k+1)/12$.

PROOF. Using the representation $Z_\rho = \sum_{u<v} (R_v - R_u)$, we have,

$$E_A(Z_\rho) = \sum_{i=1}^{n} \sum_{j=1}^{k} (2j - k - 1)E_A(r_{ij}).$$

TABLE 2.1

n	1	2	3	4	5	6	7	8	9	10
ρ_u^n	.333	.389	.416	.433	.444	.452	.458	.463	.467	.470
n	11	12	13	14	15	20	25	40	50	∞
ρ_u^n	.472	.474	.476	.478	.479	.484	.487	.492	.493	.500

We easily find,

$$E_A(r_{ij}) = (k+1)/2 + \sum_{\alpha=1}^{k} (2 \int F_\alpha \, dF_j - 1)/2$$

from which it follows that,

$$E_A(Z_\rho) = n \sum_{j=1}^{k} (2j - k - 1) \cdot (\sum_{\alpha=1}^{k} \int F_\alpha \, dF_j). \tag{3.2}$$

In particular, $E_0(Z_\rho) = 0$. It is also easy to verify that $\sigma_A^2(Z_\rho) = O(n)$. Hence, by Chebychev's inequality, under H_0

$$p\text{-lim}\, Z_\rho/n = 0 \tag{3.2$'$}$$

while under H_A,

$$p\text{-lim}\, Z_\rho/n = \sum_{j=1}^{k} (2j - k - 1) \cdot (\sum_{\alpha=1}^{k} \int F_\alpha \, dF_j). \tag{3.2$''$}$$

Sufficiency is implied by (3.2′) and (3.2″) and necessity follows from the asymptotic normality of ρ under H_0 and H_A.

We should mention that there exist alternatives in H_A such that τ is consistent and ρ is not consistent and other alternatives in H_A where ρ is consistent and τ is not consistent. To illustrate this we need only exhibit orderings of the integers $1, \cdots, k$, for which Kendall's and Spearman's correlation coefficients (with respect to the natural ordering) have opposite signs. Then, if we choose the distribution functions $[F_j]$ so that the corresponding random variables achieve these orderings with sufficiently high probability, we will have produced alternatives where $\rho(\tau)$ is consistent and $\tau(\rho)$ is not. As a simple example, consider the ordering (6 3 1 2 4 5). The τ correlation between (6 3 1 2 4 5) and (1 2 3 4 5 6) is $\frac{1}{15}$ while the corresponding ρ measure is $-\frac{1}{35}$. If we then set (with $\theta > 0$) $F_6 = N(\theta, \sigma^2)$, $F_3 = N(2\theta, \sigma^2)$, $F_1 = N(3\theta, \sigma^2)$, $F_2 = N(4\theta, \sigma^2)$, $F_4 = N(5\theta, \sigma^2)$, $F_5 = N(6\theta, \sigma^2)$, and choose σ sufficiently small, the test based on τ will be consistent but the test based on ρ will not. If instead, we considered the ordering (3 2 6 5 1 4) for which $\tau = -\frac{1}{15}$ and $\rho = \frac{1}{35}$ we could, in the same manner, construct examples where the opposite consistency conclusion holds.

We now turn to the consistency condition for the Y test.

THEOREM 3.3. *A necessary and sufficient condition for the consistency of the test based on Y is $\sum_{u<v} \int H_u \, dH_v/k(k-1) > \frac{1}{4}$ where $H_u = F_u * F_u$ and $*$ denotes convolution.*

PROOF. Write T_{uv} as

$$T_{uv} = \sum_{i \leq j}^{n} \psi_{uv}(X_i', X_j') \tag{3.3}$$

where

$$\begin{aligned} \psi_{uv}(X_i', X_j') &= 1 \quad \text{if } X_{iu} - X_{iv} + X_{ju} - X_{jv} < 0 \\ &= 0 \quad \text{otherwise.} \end{aligned}$$

We then obtain

$$\begin{aligned} &(3.4) \quad E_A(T_{uv}) \\ &\qquad = (n(n-1)/2) P_A(X_{1u} + X_{2u} < X_{1v} + X_{2v}) + n P_A(X_{1u} < X_{1v}) \end{aligned}$$

where the $[X_{1u}, X_{2u}]_{u=1}^{k}$ are independent and identically distributed according to F_u. Hence,

$$\lim_n E_A(Y)/n(n-1) = \sum_{u<v} P_A(X_{1u} + X_{2u} < X_{1v} + X_{2v})/2.$$

From (3.3) it is easily seen that $\sigma_A^2(T_{uv}) = O(n^3)$ and hence $\sigma_A^2(Y) = O(n^3)$. Thus, by Chebychev's inequality, we have under H_0

$$(3.4') \qquad p\text{-}\lim 2Y/n(n-1) = k(k-1)/4$$

whereas under H_A,

$$(3.4'') \qquad p\text{-}\lim 2Y/n(n-1) = \sum_{u<v} \int H_u dH_v .$$

The sufficiency now follows from (3.4′) and (3.4″) and the necessity is a consequence of the asymptotic normality of Y under H_0 and H_A.

Theorems 3.1 and 3.3 show that the consistency parameters of the τ and Y tests are quite similar. Again, it is easy to produce examples that show that there are alternatives against which one of Y and τ is consistent and the other is not.

For instance, take $k = 2$ and let $f_1(x) = 1$ if $4 \leqq x \leqq 5$, and 0 otherwise, and $f_2(x) = .6$ if $1 \leqq x \leqq 2$, .4 if $10 \leqq x \leqq 11$, and 0 otherwise. Then $\int F_1\, dF_2 = .4$ but $\int H_1\, dH_2 = .64$, and Y will be consistent but τ will not. Of course, reversing the roles of f_1 and f_2 we get the opposite conclusion.

Theorems 3.1 and 3.2 show that the tests τ and ρ are consistent against large classes of alternatives which include the intersection of H_a and H_A while Theorem 3.3 shows that Y's consistency class includes $H_{a'} \cap H_A$ where $H_{a'}$ is defined as H_a with $F_i * F_i$ replacing F_i. (The consistency of τ and ρ against those alternatives in $H_a \cap H_A$ is a consequence of the fact that $F_1 > F_2$ implies $\int F_1\, dF_2 > \frac{1}{2}$.) Perhaps the most important conclusion to be drawn from the consistency conditions is the following: Although each statistic depends on the postulated order, we do not destroy the *desired* consistency of the tests if we postulate an order which is "close" to the correct one.

4. Asymptotic efficiencies. Let us begin by defining the S alternatives as

$$S: F_{ij}(x) = F(x - b_i - (j-1)\theta), \qquad i = 1, \cdots, n, j = 1, \cdots, k \quad \text{and} \quad \theta > 0.$$

From [15] it follows that if $T_{1(n)}$ and $T_{2(n)}$ are two test sequences (from the group of $\tau_{(n)}$, $\rho_{(n)}$, $t_{(n)}$ and $Y_{(n)}$), the Pitman efficiency of $T_{1(n)}$ with respect to $T_{2(n)}$ (hereafter called $E(T_1, T_2)$) is equal to

$$(4.1) \quad \lim_n \{(d/d\theta)E_\theta(T_{1(n)})|_{\theta=0}[(d/d\theta)E_\theta(T_{2(n)})|_{\theta=0}]^{-1}\}^2 \cdot \sigma_0^2(T_{2(n)})/[\sigma_0^2(T_{1(n)})]$$

Specifically, we are computing the efficiency for the alternatives

$$S^{(n)}: F_{ij}^{(n)}(x) = F(x - b_i - (j-1)cn^{-\frac{1}{2}})$$

but we simply say "$\theta \rightarrow 0$".

THEOREM 4.1. *For the S alternatives* $(\theta \rightarrow 0)$,

$$(4.2) \qquad E(\rho, t) = k(k+1)^{-1}(12\sigma^2[\textstyle\int f^2]^2),$$

where $\sigma^2 = \text{Var}(F)$ and f is the density (assumed to exist) corresponding to F.

PROOF. Using (3.2) we find that[2]

$$(d/d\theta)E_\theta(Z_\rho)|_{\theta=0} = (nk^2(k-1)(k+1)\int f^2)/6.$$

Also, $(d/d\theta)E_\theta(t)|_{\theta=0} \sim (12\sigma^2/nk(k-1)(k+1))^{-\frac{1}{2}}$ and $\sigma_0^2(t) \sim 1$. Since (see e.g. [12])$\sigma_0^2(\rho_i) = 1/(k-1)$, (4.2) follows from (4.1).

THEOREM 4.2. *For the S alternatives* $(\theta \to 0)$,

$$E(\tau, t) = 24(k+1)(2k+5)^{-1}\sigma^2[\int f^2]^2, \tag{4.3}$$

where $\sigma^2 = \mathrm{Var}(F)$ *and f is the density corresponding to F.*

PROOF. From (3.1) we see that

$$(d/d\theta)E_\theta(Z_\tau)|_{\theta=0} = (nk(k-1)(k+1)\int f^2)/6.$$

Since (see e.g. [12]) $\sigma_0^2(\tau_i) = 2(2k+5)/9k(k-1)$, the expression for $E(\tau, t)$ follows.

Hodges and Lehmann [6] have shown that $.864 \leqq 12\sigma^2[\int f^2]^2 \leqq \infty$ with the lower bound achieved for

$$f_1(x) = 3(5-x^2)/20\cdot 5^{\frac{1}{2}}, \qquad -5^{\frac{1}{2}} \leqq x \leqq 5^{\frac{1}{2}}, \tag{4.4}$$
$$= 0 \qquad \text{otherwise.}$$

We can thus state the following obvious corollary.

COROLLARY 4.3. *For the S alternatives* $(\theta \to 0)$,

$$.576 \leqq E(\rho, t) \leqq \infty, \qquad .576 \leqq E(\tau, t) \leqq \infty$$

with the lower bound achieved with f_1 *as in* (4.4) *and* $k = 2$.

COROLLARY 4.4. *For the S alternatives* $(\theta \to 0)$,

$$\mathrm{E}(\rho, \tau) = k(2k+5)/2(k+1)^2. \tag{4.5}$$

We note that $E(\rho, \tau) = 1$ when $k = 2$ since in this case the two procedures are equivalent to the paired sign test. Also, $E(\rho, \tau)$ increases until it reaches its maximum value of 1.042 at $k = 5$ and then decreases to its limiting value ($k \to \infty$) of 1. We also remark that Noether and van Elteren [4] found expression (4.2) to be the Pitman efficiency (translation alternatives) of Friedman's rank test with respect to the normal theory F-test.

The values of $E(\rho, t)$ and $E(\tau, t)$ for small k, especially when F is normal, are somewhat discouraging and one would like to be able to improve on these. The Y test, as we will now see, yields such an improvement.

THEOREM 4.5. *For the S alternatives* $(\theta \to 0)$,

$$E(Y, t) = 24(k+1)\sigma^2[\int g^2]^2/(3 + 2(k-2)\rho^*(F)) \tag{4.6}$$

where G is defined to be the distribution function of $X_1 - X_2$, *with corresponding density g, when* X_1, X_2 *are independent and identically distributed according to F, and* $\sigma^2 = \mathrm{Var}(F)$.

[2] Here, and in the sequel, when we differentiate under the integral we assume sufficient regularity.

PROOF. From (3.4) and the definition of Y it is easily seen that

$$\lim_n (d/d\theta)E_\theta(Y)/n(n-1) \mid_{\theta=0} = (k(k-1)(k+1) \int g^2)/6.$$

Expression (4.6) follows by recalling (2.3).

COROLLARY 4.6. *For the S alternatives* ($\theta \to 0$), $.864 < E(Y, t) \leqq \infty$.

PROOF. Since $\rho^*(F) \leqq \frac{1}{2}$, it follows from (4.6) that for each fixed F (σ^2 finite), $E(Y, t)$ is an increasing function of k, unless $\rho^*(F) = \frac{1}{2}$ in which case $E(Y, t)$ equals $24\sigma^2[\int g^2]^2$ for all k. For $k = 2$, $E(Y, t)$ is $24\sigma^2[\int g^2]^2$ and since $12\sigma^2[\int f^2]^2$ achieves its minimum for the f_1 given by (4.4), we conclude that $E(Y, t) > .864$. The last inequality is strict since there does not exist any density h such that $X_1 - X_2$ is distributed according to f_1 when X_1, X_2 are independent and identically distributed according to h. This can be seen as follows. For $f_1(x) = b(a^2 - x^2)$, $-a < x < a$, and 0 otherwise, the characteristic function is

$$\phi_{f_1}(t) = 4bt^{-3}(\sin(at) - (at)\cos(at)). \tag{4.7}$$

Thus, there does not exist a characteristic function ϕ_h such that

$$\phi_{f_1}(t) = \phi_h(t) \cdot \overline{\phi_h(t)}, \tag{4.8}$$

since a suitable choice of t makes $\phi_{f_1}(t)$ negative, contradicting (4.8).

While we do not establish an attainable lower bound for $E(Y, t)$ we note that for $k = 2$ and F uniform, $E(Y, t) = .889$. When $k = 2$, Y reduces to Wilcoxon's signed-rank test [21] and $E(Y, t)$ reduces to $24\sigma^2[\int g^2]^2$, the efficiency of Wilcoxon's signed-rank test with respect to the t-test.

In Table 4.1 we list some values of $E(\rho, t)$ and $E(Y, t)$ when F is normal. (From (4.2) and (4.6) it is easily seen that $E(Y, \rho) > 1$ for every k when F is normal.)

We should also point out that the Pitman efficiency results of this section are valid for the more general alternatives

$$S'^{(n)}: F_{ij}^{(n)}(x) = F(x - b_i - \alpha_j c n^{-\frac{1}{2}})$$

where we require that the sequence of constants $[\alpha_j]_{j=1}^k$ does not depend on n and the α's are not all equal. Of course the t-test is still to be interpreted as the one developed for the model specified by (iii), Section 1.

5. Related work. Doksum [3] has recently proposed a test that is very similar to the Y test. Doksum uses the random variables $U_{uv} = T_{uv} - \sum_{i=1}^n \psi_{uv}^{(i)}$, and considers the statistic $\sum_{u<v}^k (U_{u\cdot} - U_{v\cdot})$. An asymptotically equivalent statistic is $Y' = \sum_{u<v}^k (T_{u\cdot} - T_{v\cdot})$. Doksum shows that $E(Y', Y) \geqq 1$ for all F and k;

TABLE 4.1

	k									
	2	3	4	5	6	7	10	20	50	∞
$E(Y, t)$	.955	.963	.968	.972	.974	.976	.980	.984	.987	.989
$E(\rho, t)$	.637	.716	.764	.796	.819	.836	.868	.909	.936	.955

we note that $\lim_{k\to\infty} E(Y', Y) = 1$ for all F. How close $E(Y', Y)$ is to one depends on k and on how close $\lambda(F)$ is to Lehmann's upper bound of 7/24. All values of $\lambda(F)$ tabulated to this date are very near the upper bound and the values of $\max_k E(Y', Y)$ given below indicate that only a very slight increase in Pitman efficiency is obtained by using Y' instead of Y.

F	Normal	Rectangular	Exponential
$\text{Max}_k\, E(Y', Y)$	1.00356	1.00193	1.00575

In each case the maximum occurs at $k = 5$.

Doksum also considers the application of a Nüesch type test [16] to the multivariate vector $[T_{u\cdot} - T_{v\cdot}, u < v]$.

The analogous problem of testing homogeneity against ordered alternatives in the k-sample problem (where we have random variables X_{ij}, $i = 1, \cdots, n_j$, $j = 1, \cdots, k$, independent with distribution functions $P(X_{ij} \leqq x) = F_j(x)$) has received considerable attention. Procedures have been proposed by (among others) Bartholomew [1], Hogg [8], Kudō [13], in the parametric case and nonparametric tests have been proposed by Chacko [2], Jonckheere [10], Puri [18] and Whitney [20]. Puri also compares the performance of many of these procedures on the basis of asymptotic power and Pitman efficiency.

6. Acknowledgments. I wish to express my gratitude to my dissertation adviser, Professor Lincoln E. Moses. His instruction, suggestions, and criticisms, have been invaluable in the preparation of this paper, and, on a larger scale, in providing me with an exciting introduction to the field of nonparametric statistics. I also thank Professor Rupert G. Miller for helpful suggestions relating to the dissertation and Professor I. Richard Savage for his assistance in the preparation of this manuscript. Finally I thank Professor Kjell Doksum for permitting me to see his paper in advance of publication, and an associate editor and referees for their comments.

REFERENCES

[1] Bartholomew, D. J. (1961). Order tests in the analysis of variance. *Biometrika* **48** 325–332.

[2] Chacko, V. J. (1963). Testing homogeneity against ordered alternatives. *Ann. Math. Statist.* **34** 945–956.

[3] Doksum, K. A. (1967). Robust procedures for some linear models with one observation per cell. *Ann. Math. Statist.* **38** 878–883.

[4] Elteren, Ph. van and Noether, G. E. (1959). The asymptotic efficiency of the χ_n^2-test for a balanced incomplete block design. *Biometrika* **46** 475–477.

[5] Friedman, M. (1937). The use of ranks to avoid the assumption of normality implicit in the analysis of variance. *J. Amer. Statist. Assoc.* **32** 675–701.

[6] Hodges, J. L. and Lehmann, E. L. (1956). The efficiency of some nonparametric competitors of the t-test. *Ann. Math. Statist.* **27** 324–335.

[7] Hoeffding, W. A. (1948). A class of statistics with asymptotically normal distribution. *Ann. Math. Statist.* **19** 293–325.

[8] Hogg, R. V. (1965). On models and hypotheses with restricted alternatives. *J. Amer. Statist. Assoc.* **60** 1153–1162.

[9] HOLLANDER, M. (1966). An asymptotically distribution-free multiple comparison procedure-treatments vs. control. *Ann. Math. Statist.* **37** 735–738.

[10] JONCKHEERE, A. R. (1954). A distribution-free k-sample test against ordered alternatives. *Biometrika* **41** 135–145.

[11] JONCKHEERE, A. R. (1954). A test of significance for the relation between m rankings and k ranked categories. *Brit. J. Statist. Psych.* **7** 93–100.

[12] KENDALL, M. G. (1955). *Rank Correlation Methods*. Griffin, London.

[13] KUDŌ, A. (1963). A multivariate analogue of the one-sided test. *Biometrika* **50** 403–418.

[14] LEHMANN, E. L. (1964). Asymptotically nonparametric inference in some linear models with one observation per cell. *Ann. Math. Statist.* **35** 726–734.

[15] NOETHER, G. E. (1955). On a theorm of Pitman. *Ann. Math. Statist.* **26** 64–68.

[16] NÜESCH, P. E. (1966). On the problem of testing location in multivariate populations for restricted alternatives. *Ann. Math. Statist.* **37** 113–119.

[17] PAGE, E. B. (1963). Ordered hypotheses for multiple treatments: a significance test for linear ranks. *J. Amer. Statist. Assoc.* **58** 216–230.

[18] PURI, M. L. (1965). Some distribution-free k-sample rank tests of homogeneity against ordered alternatives. *Comm. Pure and Applied Math.* **18** 51–63.

[19] TUKEY, J. W. (1949). The simplest signed-rank tests. Princeton University Statistical Research Group Memo Report, No. 17.

[20] WHITNEY, D. R. (1951). A bivariate extension of the U statistic. *Ann. Math. Statist.* **22** 274–282.

[21] WILCOXON, F. (1945). Individual comparisons by ranking methods. *Biometrics* **1** 80–83.