NOTE ON A TUKEY TEST FOR ORDERED ALTERNATIVES

WALTER R. PIRIE AND MYLES HOLLANDER

(Received Oct. 30, 1972; revised Aug. 15, 1973)

Summary

In this note we consider a method proposed by Tukey [6], for detecting ordered alternatives among k treatments in a randomized block design. A rank test obtained by this method is shown to be equivalent to a generalized sign test. The test is easily motivated and is quick and simple to compute, however its efficiency properties make it unattractive except for relatively small k.

1. Introduction

Consider a randomized block model where X_{ij}, $i=1,\cdots,n$, $j=1,\cdots,k$ are independent and $P\{X_{ij}\leqq x\}=F_j(x-a_i)$, with F_j continuous and a_i represents the nuisance effect (fixed or random) of block i. To test $H_0: F_1=F_2=\cdots=F_k$ against ordered alternatives of the form $F_1\geqq F_2\geqq\cdots\geqq F_k$ (where at least one of the inequalities is strict), Tukey [6] suggested the following approach: Compute a measure S_i (say) of the differences between treatment responses for each block; Let R_i be the rank of S_i in the ranking from least to greatest of $\{S_1,\cdots,S_n\}$ and reject H_0 for large values of $J_n=\sum_{i=1}^n R_i\eta_i$, where $\eta_i=1$ if $X_{i1}\leqq X_{i2}\leqq\cdots\leqq X_{ik}$, and 0 otherwise. For the measure S_i, Tukey settles on the "*least differences between responses to adjacent treatments*" within each block. If one interprets the italicized phrase to mean the least *absolute* differences, then for $k=2$, J_n is the Wilcoxon signed rank statistic. For $k>2$, however, the properties of this test are intractable. If the word *absolute* is not intended, but instead S_i is defined as

$$S_i=\min\{X_{i,j+1}-X_{ij};\ 1\leqq j\leqq k-1\} \tag{1}$$

then we have

THEOREM. *With S_i defined by* (1), *the J_n test is equivalent to the test which rejects for large values of $K_n=\sum_{i=1}^n \eta_i$.*

PROOF. Let d_i be the index of the block with rank i (i.e., $R_{d_i}=i$). Then $J_n=\sum_{i=1}^{n} i\eta_{d_i}$. If $K_n=m$, then $\eta_{d_1}=\eta_{d_2}=\cdots=\eta_{d_{n-m}}=0$, $\eta_{d_{n-m+1}}=\cdots=\eta_{d_n}=1$, and $J_n=\sum_{i=n-m+1}^{n} i=K_n(2n+1-K_n)/2$. It follows easily that within the range of possible values, J_n is a strictly increasing function of K_n and the theorem follows.

Thus, with S_i defined by (1), J_n can be viewed as a generalized sign test. Note that K_n has the binomial distribution with parameters n and $p=\mathrm{P}\{X_{11}\leqq X_{12}\leqq\cdots\leqq X_{1k}\}$ and, under H_0, $p=(k!)^{-1}$. Although this generalized sign test is distribution-free under H_0, and is particularly easy to compute, we show in Section 2 that (as one would expect) as k increases its asymptotic efficiency ($n\to\infty$) decreases drastically, making it unattractive for large k.

2. Asymptotic relative efficiency of the generalized sign test

Consider the ordered location alternatives

$$X_{ij}=a_i+c\{j-[(k+1)/2]\}\theta+e_{ij}\ , \qquad c>0,\ \theta>0\ , \tag{2}$$

where the e's are independent and identically distributed according to F which is assumed to have a square integrable density f. For this model we compute the efficacy, $e_n(K)=\{\mu'_{K_n}(0)\}^2/\mathrm{var}_0\, K_n$, of the generalized sign test, where $\mu_{K_n}(\theta)=\mathrm{E}_\theta\, K_n$ and $\mu'_{K_n}(0)=(d/d\theta)\mu_{K_n}(\theta)|_{\theta=0}$. We have $\mathrm{E}_\theta\, K_n=n\,\mathrm{P}_\theta(A)$ where $A=\{e_{11}<e_{12}+c\theta<e_{13}+2c\theta<\cdots<e_{1k}+(k-1)c\theta\}$. It follows that

$$\mu'_{K_n}(0)=nc\sum_{i=1}^{k-1}\int\cdots\int_{A^*} f(x_i)\prod_{j=1}^{k-1} dF(x_j)=nc\{(k-2)!\}^{-1}\int f^2\ , \tag{3}$$

where $A^*=\{x_1<x_2<\cdots<x_{k-1}\}$. Since $\mathrm{var}_0\, K_n=n(k!-1)(k!)^{-2}$, the efficacy is

$$e_n(K)=nc^2\{(k!-1)^{-1}\}k^2(k-1)^2\left[\int f^2\right]^2 . \tag{4}$$

For comparison, we consider Page's [2] test, a standard distribution-free test for the problem, which rejects H_0 for large values of $L_n=\sum_{i=1}^{n}\sum_{j=1}^{k} jR_{ij}$, where R_{ij} is the rank of X_{ij} in the ranking least to greatest of $\{X_{i1},\cdots,X_{ik}\}$. (This test is equivalent to rejection for large values of $\sum_{i=1}^{n}\rho_i$ where ρ_i is Spearman's rank correlation coefficient between postulated order and observed order in block i.) From Hollander [1], the efficacy $e_n(L)$ of the Page test for model 2, is found to be

$$e_n(L)=nc^2k^2(k-1)\left[\int f^2\right]^2. \tag{5}$$

The definition of the Pitman asymptotic relative efficiency e_{KL} of the generalized sign test with respect to Page's test for model (2) with θ replaced by $\theta n^{-1/2}$ is $e_{KL}=\lim_{n\to\infty} e_n(K)/e_n(L)$ and we obtain from (4) and (5)

$$e_{KL}=(k-1)\{k!-1\}^{-1}. \tag{6}$$

Note that e_{KL} does not depend on f for these particular alternatives. (This is not true in general.)

For $k=2$, $e_{KL}=1$ as both procedures are equivalent to the sign test; for $k=3$, $e_{KL}=.400$, for $k=4$, $e_{KL}=.130$, and $\lim_{k\to\infty} e_{KL}=0$. From these efficiency values, it is clear that K_n is useful only for small k, as an easily motivated test which is very quick and easy to implement.

Hollander's paper [1] contains references to several earlier proposals of parametric and nonparametric tests for this problem; for more recent developments, consult [3], [4] and [5].

Pitman asymptotic relative efficiencies of K_n with respect to other competing tests are readily obtained by the usual relationship $e_{T_1,T_3}=e_{T_1,T_2}\cdot e_{T_2,T_3}$ where e_{T_i,T_j} is the relative efficiency of two competing sequences of tests, $\{T_{in},\ n=1,\cdots\}$ and $\{T_{jn},\ n=1,\cdots\}$. Efficiency formulas for other tests are contained in the cited papers.

Acknowledgement

This work was supported in part by the U.S. Office of Naval Research (contract number NONR-988(08)NR 042-004) and in part by a training grant (5 TOI GM 00913) of the U.S. National Institute of General Medical Sciences.

VIRGINIA POLYTECHNIC INSTITUTE AND STATE UNIVERSITY
FLORIDA STATE UNIVERSITY

REFERENCES

[1] Hollander, M. (1967). Rank tests for randomized blocks when the alternatives have an *a priori* ordering, *Ann. Math. Statist.*, **38**, 867-877.

[2] Page, E. B. (1963). Ordered hypotheses for multiple treatments: a significance test for linear ranks, *J. Amer. Statist. Ass.*, **58**, 216-230.

[3] Pirie, W. R. (1974). Comparing rank tests for ordered alternatives in randomized blocks, *Ann. Statist.*, **2**, 374-382.

[4] Pirie, W. R. and Hollander, M. (1972). A distribution-free normal scores test for ordered alternatives in the randomized block design, *J. Amer. Statist. Ass.*, **67**, 855-857.

[5] Puri, M. L. and Sen, P. K. (1968). On Chernoff-Savage tests for ordered alternatives in randomized blocks, *Ann. Math. Statist.*, **39**, 967-972.

[6] Tukey, J. W. (1957). Sums of random partitions of ranks, *Ann. Math. Statist.*, **28**, 987-992.