Journal of the KSQC
Vol. 17, No. 2, 1989

# Nonlinear Rank Statistics for the Simple Tree Alternatives

Sang-Gue Park*
Tai-Kyoo Kim*

**ABSTRACT**

Nonlinear rank statistics for the simple tree alternatives problem are considered. Pitman efficiencies between several procedures are studied. A new maximin procedure is suggested and shown to have good efficiency properties. Additionally, it is desirable to terminate the experiment early comparing well known rank statistics or multiple comparison test statistics.

## I. Introduction

Let $X_{i1}$, $X_{i2}, \cdots$, $X_{in_i}$, $i=0, 1, \cdots, K$, be $(k+1)$ independent random samples, where $X_{ij}$, $j=1, 2, \cdots, n_i$, is a sample from the $i$th population with an absolutely continuous distribution function $F_i = F(x-\theta_i)$. In other words the distributions of the $(k+1)$ populations are the same except for a possible difference in their location parameters. Without any loss of generality let $F(0)=p$, so that $< i$ is the $p$th percentile of $F$. It is of interest to test the null hypothesis

$$H_0 : \theta_0 = \theta_i, \text{ for } i=1, 2, \cdots, k$$

against the partially ordered one-sided alternatives

*Department of applied Statistics, Hannam University, Daejeon, Korea

$H_1 : \theta_0 \leq \theta_i$, for $i=1, 2, \cdots, k$ and not $H$.

This is referred to "simple tree alternatives".
Abelson and Tukey(1963) studied various order restricted testing problems, but they assumed the parametric model, which is normal. And Shirahata(1980), Hettmansperger and Norton (1987) and Fairly and Fligner(1987) also discussed the other alternatives problem through the linear rank statistics. However, we propose a class of the nonlinear rank statistic.

We consider the following class of statistics

$$T = \sum_{i=1}^{k} a_i W_i \tag{1}$$

where a's are suitably chosen weights and $W_i$ is the number of observations in the $i$th sample not exceeding $M$, the $t$th order statistic in the sample from $F_0$, $t=[n_0 p]+1$, $[x]$ repreasents the largest integer not exceeding $x$ and $p$ is preassigned value.

The proposed class of test is reject $H_0$ in favor of the alternatives $H_1$ if

$$T < c \tag{2}$$

where $c$ is a constant such that $p(T<c|H_0)$.

Recently Chakraborti and Desu(1988a) proposed and studied a class of tests for the $k$-sample problems assuming the simle order alternatives. This class is a linear combination of statistics $W_j *$, $j=2, 3, \cdots, k$, where $W_j *$ is a generalization of Mathisen's median test statistic for comparing treatment $j$ with treatment 1.

The optimal test is compared with some of the previously proposed tests. Further Chakraborti and Desu(1988b) extended the Mathisen's two sample test to the problem of comparing several treatments with a control. The advantage of this test over the test of Slivka(1970), which is a multiple comparison test, is that new test requires an experiment with shorter average duration. This feature is quite an attractive one in the context of lifetime-testing experiments where testing time may be expensive. The weights a's are determined such that the Pitman efficacy of the test is maximized for a suitable sub-class of the local alternatives. Chakraborti and Desu (1988b)'s statistic is a member of our proposed statistics.

## II. Test Statistics

Consider the following sequences of Pitman alternatives

$$H_{1AN} = \theta_0 + c_i/\sqrt{N}, \quad i=1, 2\cdots, k \tag{3}$$

where $c_i \geq c_0 = 0$ with at least one strict inequality and $N = \sum_{i=0}^{k} n_i$. The weights a's are determined in each situation so that the Pitman efficacy of the test (2) based on $T$ is maximized. Since $T$ is not a linear ranlk statistic, many of the standard theorems can not be used to find the asymptotic distribution of $T$. However, we can use the following theorem which is an immediate corollary of theorem 5 of Chakraborti(1984).

**Theorem** 1 : Let $W$ be the $k$-dimensional vector whose ith element is

$$N(W_i/n_i - F_i(\theta_0)), \quad i = 1, 2, \cdots, k.$$

Suppose that, as $\min(n_0, n_1, \cdots, n_k) \to \infty$, $n_i/N \to \lambda_i$, $0 < \lambda_i < 1$, $i = 0, 1, \cdots, k$. Further, suppose that $F'(\theta_0) = f(\theta_0)$ exists and it is positive for $i = 0, 1, \cdots, k$. Then

$$\underline{W} \xrightarrow{D} N_k(0, S)$$

$$S = (\!( Q_i Q_j p_0 \lambda_0^{-1} + \delta_{ij} p_i \lambda_i^{-1} )\!)$$

where $p_0 = p(1-p)$, $p_i = F_i(\theta_0)\ (1 - F_i(\theta_0))$, $Q_i = f_i(\theta_0)/f_0(\theta_0)$, $i = 0, 1, \cdots, k$ and $\delta_{ij}$ is the kronecker delta.

From theorem 1, it can be shown that the Pitman efficacy of $T$ is

$$e = \frac{\sum_{i=1}^{k} a_i \lambda_i C_i}{\left( \sum_{i=1}^{k} a_i^2 \lambda_i + \frac{1}{\lambda_0 \left( \sum_{i=1}^{k} a_i \lambda_i \right)^2} \right)} \cdot \frac{f(0)}{p(1-p)} \tag{4}$$

Now if we reject $H_0$, when $T \leq \text{constant} \times Z_\alpha$, the test based on $T$ is approximately of size $\alpha$ and the approximate local power is $eZ_\alpha\theta$. The test with the largest value for $e$ is the most efficient and referred to as the optimal or the best test. The best test of this class is given in the following.

**Theorem** 2 : Given the pattern $\underline{c} = (c_0, c_1, \cdots, c_k')$, $e$ is maximized by $a_i = \lambda_0^{-1}(c_i - c_w)$ where

$$c_w = \sum_{i=1}^{k} \lambda_i c_i.$$

**Proof**

See Chakraborti and Desu(1988a).

Clearly the best test derived in theorem 2 has the maximum local power. The optimal test statistic is

$$T(opt)=\sum_{i=1}^{k}\lambda_0^{-1}(c_i-c_w)W_i \tag{5}$$

Now the asymptotic relative efficiency of any member of $T$ relative to the best is given by

$$\mathrm{ARE}(T,T(opt))=\Big(\sum_{i=1}^{k}a_i\lambda_i c_i\Big)^2\Big/\Big[\Big(\sum_{i=1}^{k}\lambda_0\lambda_i a_i^2+\Big(\sum_{i=1}^{k}a_i\lambda_i\Big)^2\Big)\Big(\lambda_0^{-1}\sum_{i=1}^{k}\lambda_i(c_i-c_w)^2+c_w^2\Big)\Big]$$

$$=[\underset{\sim}{a}'D_\lambda\underset{\sim}{c}^*]'/(\underset{\sim}{a}'H\underset{\sim}{a})(\underset{\sim}{c}^{*\prime}N\underset{\sim}{c}^*) \tag{6}$$

where $\underset{\sim}{c}^*=(c_0,c_1,\cdots,c_k)$, $H=(\!(h_{ij})\!)$ is given by $h_{ij}=\lambda_i(\lambda_i+\lambda_0)$ and $h_{ij}=\lambda_i\lambda_j)$ and $\mathrm{H}=(\!(\mathrm{n}_{ij})\!)$ is given by $n_{ij}=\lambda_0^{-1}\lambda_i(1-\lambda_j-\lambda_0\lambda_j)$ and $n_{ii}=\lambda_0^{-1}\lambda_i(1-\lambda_i-\lambda_0)$.

Specifying the location parameters vector $\underset{\sim}{c}$ may be called specifying the "pattern to be detected", according to Hettmansperger and Norton(1987). Now if the experimenter is not able to specify the pattern, $c$, it may be necessary to provide some guidelines concerning the choice of the constants a's. Thus we propose a new statistic which is enjoying a optimal property. It is referred to relative maximin ($RM$) statistic, a statistic which maximizes the minimum efficiency relative to the optimal test.

**Theorem** 3 :

$$T(RM)=\sum_{i=1}^{k}\sqrt{\frac{(1-\lambda_i)}{\lambda_i\lambda_0}}W_i \tag{7}$$

**Proof**

$$T(RM)=\max_{\underset{\sim}{a}}\min_{\underset{\sim}{c}\in H_{1AN}}\mathrm{ARE}(T,T(opt)) \tag{8}$$

From Park, Chakraborti and Desu(1989), $\min_{\underset{\sim}{c}\in H_{1AN}}\mathrm{ARE}(T,T(opt))$ occurred at the slippage alternatives, $c_0=c_1=\cdots=c_i-1=c_i+1=\cdots=c_k<c_i$, for some $i=1,2,\cdots,k$, Now we are maximizing $\min_{\underset{\sim}{c}\in H_{1AN}}\mathrm{ARE}(T,T(opt))$ with respect to $\underset{\sim}{a}$.

$$\min_{\underset{\sim}{c}\in H_{1AN}}\mathrm{ARE}(T,T(opt))$$

$$=\lambda_0\lambda_j a_j^2/(\underset{\sim}{a}'H\underset{\sim}{a})(1-\lambda_j)\ \text{ for some } j=1,2,\cdots,k.$$

$$=b_j^2/(\underset{\sim}{b}'D^{-1}HD^{-1}\underset{\sim}{b})$$

Letting $b_j=\sqrt{\frac{\lambda_j\lambda_0}{1-\lambda_j}}a_j$ and

$\underline{b} = D\,\underline{a}$ where $D = \text{diag}\left(\sqrt{\frac{\lambda_i \lambda_0}{1-\lambda_i}}\right)$ Then

$$MN = \max_{(\underline{b}\,:\, b_i \geq b_j)} b_j^2 / (\underline{b}' D^{-1} H D^{-1} \underline{b})$$

Since $b_j^2/\underline{b}' D^{-1} H D^{-1} \underline{b}$ does not dopend on $|b|$, but only on its direction we can take $b_j = 1$, so that

$$MN = \max_{\underline{b}\,:\, b_i \geq 1} 1/\underline{b}' V \underline{b}, \quad \text{where } V = D^{-1} H D^{-1}$$

Clearly, maximizing $1/\underline{b}' V \underline{b}$ is the same as minimizing $\underline{b}' V \underline{b}$. Since $V$ is nonnegative definite, $\underline{b} = \underline{1}$ makes $\underline{b}' V \underline{b}$ minimum.

Thus $\underline{1} = \underline{b}$

$= D\underline{a}$, and the proof is completed.

## III. Allocation of Observations.

We now turn to the problem of allocation of the experimental units to the treatments and the control is such a may that the approximate power of $T$ is maximized for the local alternatives. It is of interest to consider this problem when we plan to start the experiment. Let us assume that equal numbers of the experimental units are to be assigned to each of the treatments.

That is, we have $\lambda_0$ and $\lambda_1 = \lambda_2 = \cdots = \lambda_k = \lambda$.

**Theorem** 4 : With the restriction $\lambda_0 = 1/2$ and $\lambda = 1/k_2$ maximizes the efficacy of $T$, given $c, p$ and the density $f$.

**Proof**

When we have the restriction

$$T(RM) = \sum_{i=1}^{k} W_i \text{ ; that is, } a_i = 1, \quad i = 1, 2, \cdots, k$$

Thus the efficacy of $T(RM)$ with this restriction is

with $\lambda_0 + k\lambda = 1$. (9)

Since $\lambda = \frac{1-\lambda_0}{k}$, we are plugging this in (9).

$$\frac{\frac{(1-\lambda_0)}{k}\sum_{i=1}^{k} c_i}{\sqrt{(1-\lambda_0)+\frac{(1-\lambda_0)^2}{\lambda_0}}}\frac{f(0)}{P(1-p)}=\sqrt{\frac{\lambda_0(1-\lambda_0)}{k^2}}\sum c_i\frac{f(0)}{P(1-p)}.$$

Thus given $c, p$ and $f$, $\lambda_0=1/2$ maximizes the approximate local power of $T(RM)$ (efficacy of $T(RM)$).

## IV. Simulation

When we consider the oneway layout problem, it is natural to consider Kruscal-Wallis test statistic, so-called the omnibus test statistic. In this chapter, we are comparing our proposed statistic with the omnibus test statistic. The results are based on 10,000 iteration using Fortran 77. The study compares the estimated powers for various values of the location parameters under normal, Cauchy, uniform distribution. The maximum variance of these estimetes is under 0.01. The IMSL subroutines RANOR, RANCHY, RANUNI were used to generate the random numbers. The used $\alpha$-level was 0.05. Let denote the location parameter for population $i$ and $\theta=(\theta_0, \theta_1, \cdots, \theta_k)'$ denote the vector of location parameters, $\theta_0$ is the location parameters for the control population. Various configurations of $\theta$ were used for $k=2,3,4$ and $n_0=kn$, $ni=n$, $i=1,2,\cdots,k$. And we are using $p=1/2$. We are putting some of simulation results.

The overall recommendation is to use $T(RM)$. Especially it is suitable when the experimenter believes the underlying distribution is heavy tailed like Cauchy. However, if it is thought there whoud be a great deal of difference between treatments, Kruscal-Wallis statisitc is a bit working better than our proposed statistic. And Simulation studies show that we can get more powerful results if we increase the control sample size.
This one varifies the result of theorem 4.

Monte Carlo Power Estimate: $k=2$, $\alpha=0.05$

Underlying Distribution : Normal

| Location Parameter | | | $n=10$, $n_0=10$ | | : | $n=10$, $n_0=20$ | |
|---|---|---|---|---|---|---|---|
| $\theta_0$ | $\theta_1$ | $\theta_2$ | $K-W$ | $RM$ | : | $K-W$ | $RM$ |
| 0 | 0.0 | 0.0 | 0.05 | 0.05 | : | 0.05 | 0.05 |
| 0 | 0.0 | 0.5 | 0.095 | 0.12 | : | 0.09 | 0.16 |
| 0 | 0.5 | 0.5 | 0.093 | 0.18 | : | 0.18 | 0.24 |
| 0 | 0.5 | 1.0 | 0.26 | 0.28 | : | 0.44 | 0.35 |
| 0 | 1.0 | 1.0 | 0.42 | 0.40 | : | 0.70 | 0.48 |
| 0 | 1.0 | 1.5 | 0.685 | 0.55 | : | 0.91 | 0.61 |
| 0 | 1.5 | 1.5 | 0.84 | 0.67 | : | 0.98 | 0.75 |

Monte Carlo Power Estimate: $k=3$, $\alpha=0.05$

Underlying Distribution : Normal

| location Parameter | | | | $n=10$, $n_0=10$ | | : | $n=10$, $n_0=30$ | |
|---|---|---|---|---|---|---|---|---|
| $\theta_0$ | $\theta_1$ | $\theta_2$ | $\theta_3$ | $K-W$ | $RM$ | : | $K-W$ | $RM$ |
| 0 | 0.0 | 0.0 | 0.0 | 0.05 | 0.05 | : | 0.05 | 0.05 |
| 0 | 0.0 | 0.0 | 0.5 | 0.21 | 0.15 | : | 0.22 | 0.19 |
| 0 | 0.5 | 0.0 | 0.5 | 0.22 | 0.20 | : | 0.34 | 0.23 |
| 0 | 0.5 | 0.5 | 0.5 | 0.17 | 0.28 | : | 0.27 | 0.30 |
| 0 | 0.5 | 1.0 | 0.5 | 0.26 | 0.36 | : | 0.50 | 0.39 |
| 0 | 0.5 | 1.0 | 1.0 | 0.59 | 0.44 | : | 0.76 | 0.47 |
| 0 | 1.0 | 1.0 | 1.0 | 0.75 | 0.55 | : | 0.87 | 0.56 |
| 0 | 0.5 | 1.0 | 1.5 | 0.89 | 0.51 | : | 0.95 | 0.53 |
| 0 | 1.0 | 1.5 | 1.5 | 0.94 | 0.73 | : | 0.99 | 0.71 |
| 0 | 1.5 | 1.5 | 1.5 | 0.98 | 0.81 | : | 0.99 | 0.85 |

Monte Carlo Power Estimate: $k=2$, $\alpha=0.05$

Underlying Distribution : Cauchy

| Location Parameter | | | $n=10$, $n_0=10$ | | : | $n=10$, $n_0=20$ | |
|---|---|---|---|---|---|---|---|
| $\theta_0$ | $\theta_1$ | $\theta_2$ | $K-W$ | $RM$ | : | $K-W$ | $RM$ |
| 0 | 0.0 | 0.0 | 0.05 | 0.05 | : | 0.05 | 0.05 |
| 0 | 0.0 | 0.5 | 0.05 | 0.12 | : | 0.05 | 0.12 |
| 0 | 0.5 | 0.5 | 0.06 | 0.17 | : | 0.08 | 0.17 |
| 0 | 0.5 | 1.0 | 0.09 | 0.24 | : | 0.13 | 0.22 |
| 0 | 1.0 | 1.0 | 0.14 | 0.31 | : | 0.23 | 0.29 |
| 0 | 1.0 | 1.5 | 0.19 | 0.38 | : | 0.32 | 0.36 |
| 0 | 1.5 | 1.5 | 0.28 | 0.47 | : | 0.45 | 0.44 |

Monte Carlo Power Estimate: $k=3$, $\alpha=0.05$

Underlying Distribution : Cauchy

| Location Parameters | | | | $n=10$, $n_0=10$ | | : | $n=10$, $n_0=30$ | |
|---|---|---|---|---|---|---|---|---|
| | | | | $K-W$ | $RM$ | : | $K-W$ | $RM$ |
| $\theta_0$ | $\theta_1$ | $\theta_2$ | $\theta_3$ | | | | | |
| 0 | 0.0 | 0.0 | 0.0 | 0.05 | 0.05 | : | 0.05 | 0.05 |
| 0 | 0.0 | 0.0 | 0.5 | 0.10 | 0.13 | : | 0.11 | 0.16 |
| 0 | 0.5 | 0.0 | 0.5 | 0.12 | 0.16 | : | 0.15 | 0.19 |
| 0 | 0.5 | 0.5 | 0.5 | 0.07 | 0.20 | : | 0.11 | 0.24 |
| 0 | 0.5 | 1.0 | 0.5 | 0.08 | 0.25 | : | 0.16 | 0.28 |
| 0 | 0.5 | 1.0 | 1.0 | 0.13 | 0.30 | : | 0.26 | 0.34 |
| 0 | 1.0 | 1.0 | 1.0 | 0.15 | 0.36 | : | 0.34 | 0.40 |
| 0 | 0.5 | 1.0 | 1.5 | 0.25 | 0.34 | : | 0.42 | 0.38 |
| 0 | 1.0 | 1.5 | 1.5 | 0.25 | 0.48 | : | 0.53 | 0.51 |
| 0 | 1.5 | 1.5 | 1.5 | 0.28 | 0.55 | : | 0.62 | 0.56 |

## REFERENCES


1. Abelson, R. P. and Tukey, J. W. (1963) ; "*Efficient utilization of nonnumerical information in quantitative analysis,*" General theory and the case of simple order. Ann. Math. Statist., 34, 1347-69.
2. Chakraborti, S. (1984) ; "*A generalization of the control median test,*" Unpublished Ph. D dissertation, SUNY-Buffalo.
3. Chakraborti, S. and Desu, M. M. (1988 a) ; "*A class of distribution-free tests for testing homogeneity against ordered alternatives,*" Stat & Prob. letters, 6251-256.
4. Chakraborti, S. and Desu, M. M. (1988 b) ; "*Generalization of Mathisen's median test for comparing several treatments with a control,*" Comm in Statist.-Simul. and Comp.
5. Fairley, D. and Fligner, M. (1987) ; "*Linear rank statistics for the ordered alternatives problem,*" Comm. Statist.-theor. Meth., 16, 1-16.
6. Hettmansperger, T. P. and Norton, R. M. (1987) ; "*Tests for pattered alternatives in k-sample problems,*" JASA, 82, 292-299.
7. Park, S., Chakraborti, S. and Desu, M. M. (1989) ; "*Linear rank statistics for the simple tree alternatives problem,*" Submitted.
8. Shirahata, S. (1980) ; "*Rank tests for the k-sample problems with restricted alternatives,*" Comm. Statist.-Theor. Meth., A9. 1071-1086.
9. Slivka, J. (1970) ; "*A one-sided nonparametric multiple comparison control percentile test,*" treatments versus control, Biometrika 57, 431-38.