The Annals of Mathematical Statistics
1970, Vol. 41, No. 6, 1942–1951

ON SELECTION PROCEDURES BASED ON RANKS: COUNTEREXAMPLES CONCERNING LEAST FAVORABLE CONFIGURATIONS

BY M. HASEEB RIZVI[1] AND GEORGE G. WOODWORTH[2]

Kansas State University and Stanford University

1. Introduction. Let $\pi_1, \pi_2, \cdots, \pi_k$ denote $k > 2$ univariate populations differing only in location; that is, an observation X_i drawn from π_i has cumulative distribution function (cdf) $F(x-\theta_i)$ where F is a known continuous cdf with square integrable density f but the location parameter vector $\boldsymbol{\theta} = (\theta_1, \cdots, \theta_k)$ is unknown. Let the ordered values of the location parameters be denoted by $\theta_{[1]} \leqq \theta_{[2]} \leqq \cdots \leqq \theta_{[k]}$.

Selecting the t best populations. The decision problem here is to select the populations corresponding to the $t < k$ largest θ-values. The goal of the decision maker is to find a procedure, say R, and a sample size n such that the probability of a correct selection using rule R, $P[CS \mid R, \boldsymbol{\theta}]$, has the property that

$$\inf_{\boldsymbol{\theta} \in D(\delta^*)} P[CS \mid R, \boldsymbol{\theta}] \geqq P^*, \tag{1.1}$$

where

$$D(\delta^*) = \{\boldsymbol{\theta} : \theta_{[k-t+1]} - \theta_{[k-t]} \geqq \delta^*\}, \tag{1.2}$$

and $\binom{k}{t}^{-1} < P^* < 1$ and $\delta^* > 0$ are preassigned constants.

Selecting a subset containing the best population. The decision problem here is to select a subset of the k populations containing the population associated with $\theta_{[k]}$. The goal of the decision maker is to find for fixed n and preassigned $P^* < 1$ a procedure, say R', such that

$$\inf_{\boldsymbol{\theta}} P[CS \mid R', \boldsymbol{\theta}] \geqq P^*.$$

We consider two procedures (proposed elsewhere) based on rank sums and show by counterexamples in Section 2 and Section 3 that they do not satisfy (1.1) (or (1.3)).

2. A procedure based on rank sums for selecting the t best populations. Let $\{X_{ij} : i = 1, \cdots, k, j = 1, \cdots, n\}$ be k samples each of size n (n is to be determined by (1.1)), X_{ij} being the jth observation from π_i, and let R_{ij} be the rank of X_{ij} among all the observations.

Received October 18, 1968; revised September 25, 1969.

[1] This author's work was supported by Federal Highway Administration Contract FH11-6667 with Stanford University and NASA Grant NGR 36-008-040 with the Ohio State University.

[2] This author's work was supported by the Office of Naval Research under Contract Nonr-225(53) with Stanford University.

Define the rank sums

(2.1) $$T_{in} = n^{-2}\sum_{j=1}^{n} R_{ij}, \qquad i = 1, \cdots, k$$

(2.2) $$= n^{-2}\sum_{j=1}^{n}\sum_{s=1}^{n}\sum_{r=1}^{k} I(X_{ij} > X_{rs}) + n^{-1},$$

where $I(\cdot)$ is the indicator of the event in parentheses.

The proposed selection rule, call it $R(n)$, is as follows:

(i) Draw samples of size n from each population and compute T_{in} for $i = 1, \cdots, k$.

(ii) Select the t populations having the largest T_{in}-values, resolving ties by the obvious randomization.

The problem now is to find a value $n = n(\delta^*, P^*; k, t, F)$ such that $R(n)$ satisfies (1.1).

In solving this problem a crucial role is played by the slippage configuration $\boldsymbol{\theta}_0$:

(2.3) $$\theta_{[1]} = \cdots = \theta_{[k-t]} = \theta_{[k-t+1]} - \delta^* = \cdots = \theta_{[k]} - \delta^*.$$

Many selection rules, for example the rule based on the sample means, have the property that the infimum in (1.1) is attained when $\boldsymbol{\theta}$ is in the slippage configuration; in other words for many rules the slippage configuration is the least favorable configuration. For such rules it is a relatively easy task to find the appropriate value of n (see, for instance, Example 1 of [1]). The following counterexample, kindly communicated to the authors by E. L. Lehmann,[3] shows that for the rank-sum rule $R(n)$ the slippage configuration is not least favorable.

COUNTEREXAMPLE 1 (E. L. Lehmann). Let $k = 3$, $t = 1$ and let F be a continuous cdf which places probability q and $p = 1-q$ respectively on the intervals $(0, \varepsilon)$ and $(1, 1+\varepsilon)$; $\varepsilon < \frac{1}{3}$ is a constant. Let $\delta^* = \varepsilon$ and consider two parameter values:

$$\boldsymbol{\theta}_0 = (0, 0, \delta^*), \qquad \boldsymbol{\theta}_1 = (0, \delta^*, 2\delta^*).$$

For $n = 2$, we show that

(2.4) $$P[CS \mid R(2), \boldsymbol{\theta}_0] > P[CS \mid R(2), \boldsymbol{\theta}_1].$$

Since $\boldsymbol{\theta}_0$ is in the slippage configuration and $\boldsymbol{\theta}_0, \boldsymbol{\theta}_1 \in D(\delta^*)$, defined by (1.2), this provides the required counterexample.

PROOF. The supports of the distributions of the populations under the two parameter configurations can be depicted as shown in Figure 1.

Let B_i be 0, 1 or 2 according as 0, 1 or 2 observations from π_i are in the upper interval of the support of its distribution, $\mathbf{B} = (B_1, B_2, B_3)$ and $\mathbf{b} = (b_1, b_2, b_3)$ is a realization of $\mathbf{B}$. Clearly $P[\mathbf{B} = \mathbf{b} \mid \boldsymbol{\theta}] = \prod_{i=1}^{3} \binom{2}{b_i} p^{b_i} q^{2-b_i}$ for $\boldsymbol{\theta} = \boldsymbol{\theta}_0$ or $\boldsymbol{\theta}_1$.

$\mathbf{R} = \{R_{ij}: i = 1, 2, 3, j = 1, 2\}$ is the vector of ranks and $\mathbf{r} = \{r_{ij}\}$ is a realization of $\mathbf{R}$. Given $\mathbf{R} = \mathbf{r}$ a correct selection (selection of π_3) occurs with probability 1 if

[3] Professor Lehmann informs us that P. S. Puri was the first to express doubts that the slippage configuration (2.3) was least favorable for procedures based on ranks.

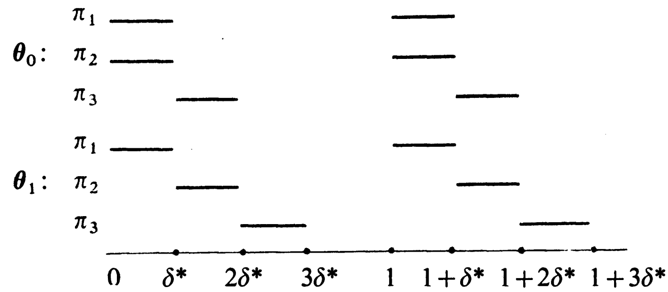

FIG. 1. Supports of distributions.

$r_{31}+r_{32} > \max(r_{21}+r_{22}, r_{11}+r_{12})$, with probability $\frac{1}{2}$ if $r_{31}+r_{32} = r_{21}+r_{22} > r_{11}+r_{12}$ or $r_{31}+r_{32} = r_{11}+r_{12} > r_{21}+r_{22}$, and with probability $\frac{1}{3}$ if $r_{31}+r_{32} = r_{21}+r_{22} = r_{11}+r_{12}$. The conditional probability that $\mathbf{R} = \mathbf{r}$ given $\mathbf{B} = \mathbf{b}$ is easy to compute, for example, given that $B = (0, 0, 0)$ and $\boldsymbol{\theta} = \boldsymbol{\theta}_0$, $X_{11}, X_{12}, X_{21}, X_{22}$ are independent and uniformly distributed (IUD) on the interval $(0, \delta^*)$ and X_{31}, X_{32} are IUD $(\delta^*, 2\delta^*)$. Thus $(R_{11}, R_{12}, R_{21}, R_{22})$ is equally likely to be any permutation (ELAP) of (1, 2, 3, 4), (R_{31}, R_{32}) is ELAP (5, 6) and these two vectors are independent. If $\boldsymbol{\theta} = \boldsymbol{\theta}_1$, then X_{11}, X_{12} are IUD $(0, \delta^*)$, X_{21}, X_{22} are IUD $(\delta^*, 2\delta^*)$, and X_{31}, X_{32} are IUD $(2\delta^*, 3\delta^*)$. Thus (R_{11}, R_{12}) is ELAP (1, 2), (R_{21}, R_{22}) is ELAP (3, 4), (R_{31}, R_{32}) is ELAP (5, 6) and these three vectors are independent. In particular, the probability that $\mathbf{R} = (1, 2; 3, 4; 5, 6)$ given $\mathbf{B} = (0, 0, 0)$ is $1/4!\cdot 2! = 1/48$ for $\boldsymbol{\theta} = \boldsymbol{\theta}_0$ and $1/2\cdot 2\cdot 2 = 1/8$ for $\boldsymbol{\theta} = \boldsymbol{\theta}_1$. Thus, for each of the 27 values of $\mathbf{b}$ one can determine the conditional probability of a correct selection given $\mathbf{B} = \mathbf{b}$ under $\boldsymbol{\theta}_0$ and $\boldsymbol{\theta}_1$. For most of the $\mathbf{b}$ the probability is the same under $\boldsymbol{\theta}_0$ and $\boldsymbol{\theta}_1$ but in the six cases listed in Table 1 there is a difference.

TABLE 1

$\mathbf{b}$	$P[\mathbf{B} = \mathbf{b}]$	$P[CS \mid \mathbf{B} = \mathbf{b}, \boldsymbol{\theta}]$ $\boldsymbol{\theta}_0$	$\boldsymbol{\theta}_1$
(0, 1, 0)	$2pq^5$	5/6	1/2
(1, 0, 0)	$2pq^5$	5/6	1
(1, 1, 0)	$4p^2q^4$	1/6	0
(1, 2, 1)	$4p^4q^2$	1/2	0
(2, 1, 1)	$4p^4q^2$	1/2	1
(2, 2, 1)	$2p^5q$	1/9	0

Thus

$$P[CS \mid R(2), \boldsymbol{\theta}_0] - P[CS \mid R(2), \boldsymbol{\theta}_1] = \tfrac{1}{3}pq^5 + \tfrac{2}{3}p^2q^4 + \tfrac{2}{9}p^5q > 0,$$

which establishes counterexample 1.

The possibility still remains that the slippage configuration is asymptotically $(\delta^* \to 0)$ least favorable; an asymptotic solution based on this assumption has been claimed by various authors ([4], [7] and [8]). This solution is as follows:

Let $A(P^*; k, t)$ be the solution of

$$\int \Phi^{k-t}(x+A)\, d\Phi^t(x) = P^* \tag{2.5}$$

where Φ is the standard normal cdf, and define $n(\delta^*, P^*; k, t, F)$ to be the smallest integer larger than

$$A^2(P^*; k, t)/12[\delta^* \int f^2(x)\, dx]^2, \tag{2.6}$$

where f is the derivative of F. The selection rule $R(\delta^*, P^*; k, t, F) = R(\delta^*, P^*)$ is the rule $R(n)$ with n set equal to $n(\delta^*, P^*; k, t, F)$. The natural inclination to call $R(\delta^*, P^*)$ "distribution-free" must be resisted; obviously one needs to know F to carry out this procedure.

If $\boldsymbol{\theta}$ is in the slippage configuration (2.3), then it can be shown ([7] or [8]) that

$$\lim_{\delta^* \to 0} P[CS \mid R(\delta^*, P^*), \boldsymbol{\theta}_0] = P^*.$$

The authors of [4] and [8] have incorrectly asserted that the slippage configuration is least favorable from which it would follow that $R(\delta^*, P^*)$ satisfies (1.1) asymptotically as $\delta^* \to 0$; i.e. for fixed P^*, it has been claimed that

$$\lim_{\delta^* \to 0} \inf_{\boldsymbol{\theta} \in D(\delta^*)} P[CS \mid R(\delta^*, P^*), \boldsymbol{\theta}] = P^*. \tag{2.7}$$

The next counterexample shows that (2.7) is false; and it seems to us that this invalidates $R(\delta^*, P^*)$ as a reasonable procedure since the infimum of $P[CS]$ is not controlled even asymptotically. The expedient of the authors of [7] of considering only that part of the parameter space where $\theta_{[k]} - \theta_{[1]} = O(n^{-\frac{1}{2}})$ is difficult to translate into practice. Does it mean that one should use $R(\delta^*, P^*)$ only when one is convinced that $\theta_{[k]} - \theta_{[1]} = O(n^{-\frac{1}{2}})$?

COUNTEREXAMPLE 2. Consider the logistic cdf $F(x) = (1 + e^{-x})^{-1}$ and let $\boldsymbol{\theta}(\delta^*) \in D(\delta^*)$ be a sequence of $\boldsymbol{\theta}$-values depending on δ^* as follows:

$$\theta_1 = \cdots = \theta_{k-t-1} = -\theta_0, \qquad \theta_{k-t} = 0, \qquad \theta_{k-t+1} = \delta^*, \tag{2.8}$$
$$\theta_{k-t+2} = \cdots = \theta_k = \theta_0,$$

where θ_0 is a fixed positive constant and $\delta^* < \theta_0$.

We now prove the following assertion: For each $k \geqq 3$ and each $t < k$, there exists a value of P^*, say P_0^*, $\binom{k}{t}^{-1} < P_0^* < 1$, such that

$$\lim_{\delta^* \to 0} P[CS \mid R(\delta^*, P_0^*), \boldsymbol{\theta}(\delta^*)] < P_0^*, \tag{2.9}$$

which clearly contradicts (2.7).

LEMMA 1.

$$\lim_{\delta^* \to 0} P[CS \mid R(\delta^*, P^*), \boldsymbol{\theta}(\delta^*)] \leqq \Phi(2^{-\frac{1}{2}} A^* \rho(\theta_0)), \tag{2.10}$$

where

$$A^* = A(P^*; k, t), \tag{2.11}$$

$$\rho(\theta_0) = 3^{\frac{1}{2}}\int H_0(2F-1)\,dF/[\int H_0^2\,dF-(\int H_0\,dF)^2]^{\frac{1}{2}} \quad \textit{and} \tag{2.12}$$

$$H_0(x) = k^{-1}[(k-t-1)F(x+\theta_0)+2F(x)+(t-1)F(x-\theta_0)]. \tag{2.13}$$

PROOF. Notice first that if $\theta_1 \leqq \theta_2 \leqq \cdots \leqq \theta_k$, then

$$\begin{aligned} &P[CS \mid R(\delta^*, P^*), \boldsymbol{\theta}] \\ &\leqq P[\max_{1\leqq i \leqq k-t} T_{in} \leqq \min_{k-t<j\leqq k} T_{jn} \mid \boldsymbol{\theta}] \\ &\leqq P[T_{k-t+1,n}-T_{k-t,n} \geqq 0 \mid \boldsymbol{\theta}], \end{aligned} \tag{2.14}$$

where n is the smallest integer greater than (2.6). From (2.2) one has, with probability one,

$$\begin{aligned} T_{k-t+1,n}-T_{k-t,n} = n^{-2}\textstyle\sum_{j=1}^n\sum_{s=1}^n \{&2I(X_{k-t+1,j} > X_{k-t,s})-1 \\ &+\textstyle\sum_{i\neq k-t \text{ or } k-t+1}[I(X_{k-t+1,j} > X_{is})-I(X_{k-t,j} > X_{is})]\}. \end{aligned}$$

Notice that this is a sum of several two-sample U-statistics.

In working out the details of Problem 8, page 257 of [3] one finds that

$$\begin{aligned} n^{-2}\textstyle\sum_{j=1}^n\sum_{s=1}^n I(X_{2j} > X_{1s}) = n^{-1}\sum_{j=1}^n F_1(X_{2j})-n^{-1}\sum_{j=1}^n F_2(X_{1j}) \\ +1-\int F_1\,dF_2+\varepsilon_n, \end{aligned}$$

where $E\varepsilon_n^2 \leqq (n-1)^{-2}$. Thus

$$\begin{aligned} &T_{k-t+1,n}-T_{k-t,n} \\ &= n^{-1}\textstyle\sum_{j=1}^n\sum_{i\neq k-t,k-t+1}\{F(X_{ij})-F(X_{ij}-\delta^*)\} \\ &\quad -n^{-1}\textstyle\sum_{j=1}^n\{2F(X_{k-t,j}-\delta^*)+(k-t-1)F(X_{k-t,j}+\theta_0) \\ &\qquad\qquad +(t-1)F(X_{k-t,j}-\theta_0)\} \\ &\quad +n^{-1}\textstyle\sum_{j=1}^n\{2F(X_{k-t+1,j})+(k-t-1)F(X_{k-t+1,j}+\theta_0) \\ &\qquad\qquad +(t-1)F(X_{k-t+1,j}-\theta_0)\} \\ &\quad +1-2\int F(x+\delta^*)\,dF(x)-(k-t-1)\int F(x+\theta_0)\,d(F(x-\delta^*)-F(x)) \\ &\qquad -(t-1)\int F(x-\theta_0)\,d(F(x-\delta^*)-F(x))+\varepsilon_n(\theta_0, \delta^*), \end{aligned} \tag{2.15}$$

where $E\varepsilon_n^2(\theta_0, \delta^*) \leqq C/n^2$ and C is an absolute constant.

Let

$$W_n = n^{\frac{1}{2}}(T_{k-t+1,n}-T_{k-t,n}); \tag{2.16}$$

routine calculation yields

$$\begin{aligned} EW_n = n^{\frac{1}{2}}\{2\int F(x+\delta^*)\,dF(x)-1+(k-t-1)\int(F(x-\theta_0)-F(x-\theta_0-\delta^*))\,dF(x) \\ +(t-1)\int(F(x+\theta_0)-F(x+\theta_0-\delta^*))\,dF(x)\}. \end{aligned}$$

By (2.6) and (2.11) one has $n^{\frac{1}{2}}\delta^* \to A^*/12^{\frac{1}{2}}\int f^2$ as $\delta^* \to 0$; thus, by Olshen's Lemma (page 1766 of [5])

$$\text{(2.17)} \quad \lim_{\delta^*\to 0} EW_n = (A^*/12^{\frac{1}{2}}\int f^2)\{2\int f^2(x)\,dx+(k-t-1)\int f(x-\theta_0)f(x)\,dx + (t-1)\int f(x+\theta_0)f(x)\,dx\}.$$

If we let $X_i = X_{i1}$, $i = 1, \cdots, k$, define $\Delta F(x) = F(x)-F(x-\delta^*)$, and recall (2.13), then it follows from (2.15) that

$$\begin{aligned}\operatorname{Var}(W_n) = (k-t-1)&\operatorname{Var}[\Delta F(X_1)]+(t-1)\operatorname{Var}[\Delta F(X_k)]\\ &+\operatorname{Var}[H_0(X_{k-t})-2\Delta F(X_{k-t})]\\ &+\operatorname{Var}[H_0(X_{k-t+1})]+O(n^{-\frac{1}{2}}).\end{aligned}$$

Since $E[\Delta F(X)]^2 \to 0$ as $\delta^* \to 0$ for any random variable X and $n \to \infty$ as $\delta^* \to 0$ it follows that,

$$\text{(2.18)} \qquad \lim_{\delta^*\to 0} \operatorname{Var}(W_n) = 2k^2\{\int H_0{}^2\,dF-(\int H_0\,dF)^2\}.$$

If we set $F(x) = (1+e^{-x})^{-1}$, then $f(x) = F(x)(1-F(x))$ and $\int f^2 = \frac{1}{6}$, so that (2.17) becomes, after integrating by parts,

$$\lim_{\delta^*\to 0} EW_n = 3^{\frac{1}{2}}A^*k\int H_0(2F-1)\,dF.$$

Since (2.15) is asymptotically normal by Liapunov's theorem, it follows that

$$\begin{aligned}&\lim_{\delta^*\to 0} P[CS \mid R(\delta^*, P^*), \theta(\delta^*)]\\ &\quad\leqq \lim_{\delta^*\to 0} P[T_{k-t+1,n}-T_{k-t,n} \geqq 0 \mid \theta(\delta^*)]\\ &\quad= \lim_{\delta^*\to 0} P[(W_n-EW_n)/(\operatorname{Var}(W_n))^{\frac{1}{2}} \geqq -EW_n/(\operatorname{Var}(W_n))^{\frac{1}{2}} \mid \theta(\delta^*)]\\ &\quad= \Phi(2^{-\frac{1}{2}}A^*\rho(\theta_0)),\end{aligned}$$

which proves Lemma 1.

REMARK. For $\theta_0 > 0$, H_0 is clearly not a linear function of F and, since H_0 and F are both monotone increasing, we have

$$\text{(2.19)} \qquad 0 \leqq \rho(\theta_0) < 1.$$

LEMMA 2. *For any k and t*

$$\text{(2.20)} \qquad \lim_{P^*\to 1} 2^{\frac{1}{2}}\Phi^{-1}(P^*)/A^* = 1,$$

where $A^* = A(P^*; k, t)$ and A is defined by (2.5).

PROOF. Let $Z_1, \cdots, Z_k$ be independent normal (0, 1) random variables. Then,

$$\begin{aligned}1-P^* &= 1-\int \Phi^{k-t}(x+A^*)\,d\Phi^t(x)\\ &= P[\max_{1\leqq i\leqq k-t} Z_i > \min_{k-t<j\leqq k} Z_j+A^*]\\ &= P[\textstyle\bigcup_{1\leqq i\leqq k-t<j\leqq k}\{Z_i > Z_j+A^*\}]\\ &\leqq t(k-t)P[Z_1 > Z_k+A^*]\\ &= t(k-t)[1-\Phi(2^{-\frac{1}{2}}A^*)].\end{aligned}$$

Also clearly

$$1-P^* \geqq [1-\Phi(2^{-\frac{1}{2}}A^*)].$$

Thus,

$$1 \leqq \frac{2^{-\frac{1}{2}}A^*}{\Phi^{-1}(P^*)} \leqq \frac{\Phi^{-1}((1-P^*)/t(k-t))}{\Phi^{-1}(1-P^*)}.$$

Lemma 2 now is a consequence of the following fact:

$$\lim_{u\to 0}\Phi^{-1}(u)/[-2\log(u)]^{\frac{1}{2}} = -1$$

which can be proved with the help of the well-known approximation to Mills' ratio.

Counterexample 2 now follows from (2.10), (2.19) and (2.20) by selecting $P_0{}^*$ large enough so that

$$2^{-\frac{1}{2}}A(P_0{}^*; k, t)/\Phi^{-1}(P_0{}^*) < 1/\rho(\theta_0).$$

A remark on the scale parameter case. Suppose π_i has cdf $F(x/\sigma_i)$ where $F(x) = 0$ for $x < 0$, F is known, and $\boldsymbol{\sigma} = (\sigma_1, \cdots, \sigma_k)$ is unknown (if $F(x) \neq 0$ for $x < 0$ then replace x by $|x|$). $R(n)$, with X_{ij} replaced by $-X_{ij}$, is proposed in [6] as a procedure to select the t smallest σ-values, subject to the requirement that the probability of a correct selection should be at least P^* when $\sigma_{[t+1]}^2/\sigma_{[t]}^2 \geqq \theta^*$, θ^* being a constant bigger than one. The slippage configuration,

$$\theta^*\sigma_{[1]}^2 = \cdots = \theta^*\sigma_{[t]}^2 = \sigma_{[t+1]}^2 = \cdots = \sigma_{[k]}^2,$$

is not least favorable, even asymptotically as $\theta^* \to 1$. This follows from Counterexample 2 by considering the random variable $Y = -\log(X)$, since if X has cdf $F(x/\sigma)$ then Y has cdf $1-F(\exp(\mu-y))$, where $\mu = -\log\sigma$, and Y_{ij} has the same rank as $-X_{ij}$.

The authors of [6] avoid this problem by confining $\boldsymbol{\sigma}$ to that part of the parameter space where $\sigma_{[k]}^2/\sigma_{[1]}^2 = 1+O(n^{-\frac{1}{2}})$; but one could criticize this as in the remarks just before Counterexample 2.

3. A procedure based on rank sums for selecting a subset containing the best population. The authors of [2] propose two kinds of rank procedures for this problem: randomized and non-randomized. The non-randomized procedure, call it $R'(n)$, puts π_i in the selected subset iff

$$T_{in} \geqq \max_j T_{jn} - c_n,$$

where

$$c_n = (12n)^{-\frac{1}{2}}kA^* + o(n^{-\frac{1}{2}}) \tag{3.1}$$

and $A^* = A(P^*; k, 1)$, defined by (2.5).

Their randomized procedure is of the same form as $R'(n)$ with T_{in} replaced by the randomized rank sum

$$T_{in}' = (N+1)/n^2\textstyle\sum_{j=1}^n Z(R_{ij}), \qquad i = 1, \cdots, k,$$

where $Z(1), \cdots, Z(N)$ are the order statistics of a sample of $N = k \cdot n$ uniform (0, 1) random variables, independent of the X's. The symbol $R'(n)$ denotes either procedure.

We shall show that the slippage configuration: $\theta_{[1]} = \theta_{[2]} = \cdots = \theta_{[k]}$ is not least favorable (the sentence containing (3.6) of [2] is false).

COUNTEREXAMPLE 3. Let $\boldsymbol{\theta}_1$ denote the configuration

$$\theta_1 = \cdots = \theta_{k-2} = -1, \qquad \theta_{k-1} = \theta_k = 0$$

and let $\boldsymbol{\theta}_0$ denote the slippage configuration for this problem: $\theta_1 = \theta_2 = \cdots = \theta_k$. If $F(x)$ is as in (3.7) and $k \geqq 3$, then

$$\lim_{n\to\infty} P[CS \mid R'(n), \boldsymbol{\theta}_1] < P^* = \lim_{n\to\infty} P[CS \mid R'(n), \boldsymbol{\theta}_0]. \tag{3.2}$$

PROOF. The equality is established in [2] and the inequality below. Clearly, first for non-randomized form of $R'(n)$,

$$P[CS \mid R'(n), \boldsymbol{\theta}_1] \leqq P[T_{kn} - T_{k-1,n} \geqq -c_n \mid \boldsymbol{\theta}_1]. \tag{3.3}$$

It follows as in the proof of Lemma 1 that $W_n = n^{\frac{1}{2}}(T_{kn} - T_{k-1,n})$ has a limiting normal distribution with zero mean and variance

$$\sigma^2(H) = 2k^2\{\int H^2\,dF - (\int H\,dF)^2\},$$

where

$$H(x) = k^{-1}[(k-2)F(x+1) + 2F(x)]. \tag{3.4}$$

Thus by (3.1) and (3.3)

$$\lim_{n\to\infty} P[CS \mid R'(n), \boldsymbol{\theta}_1] = \Phi(k(12)^{-\frac{1}{2}}A^*/\sigma(H)).$$

It follows from (2.20) that for any $\varepsilon > 0$ there exists $\frac{1}{2} < P_\varepsilon^* < 1$ such that

$$A^* = A(P_\varepsilon^*; k, 1) \leqq (1+\varepsilon)2^{\frac{1}{2}}\Phi^{-1}(P_\varepsilon^*).$$

Thus the counterexample will be proved if it can be shown that

$$\sigma^2(H) > k^2/6. \tag{3.5}$$

From (3.4)

$$\sigma^2(H)/2 = 4/12 + 4(k-2)\operatorname{Cov}(F(X), F(X+1)) + (k-2)^2 \operatorname{Var}(F(X+1)), \tag{3.6}$$

where X has cdf F.

Now let

$$\begin{aligned} F(x) &= 1/2 + x/2b & -b < x \leqq 0 \\ &= 1/2 & 0 < x \leqq 1 \\ &= 1/2 + (x-1)/2a & 1 < x \leqq 1+a, \end{aligned} \tag{3.7}$$

where $0 < a < 1 < b$ are constants to be determined below.

Thus,

$$\begin{aligned} F(x+1) &= 1/2+(x+1)/2b & -(b+1) < x \leqq -1 \\ &= 1/2 & -1 < x \leqq 0 \\ &= 1/2+x/2a & 0 < x \leqq a \\ &= 1 & a < x \end{aligned}$$

or, except for a set having zero $F(x)$-measure,

$$\begin{aligned} F(x+1) &= F(x)+1/2b & 0 < F(x) \leqq 1/2-1/2b \\ &= 1/2 & 1/2-1/2b < F(x) \leqq 1/2 \\ &= 1 & 1/2 < F(x) \leqq 1. \end{aligned} \tag{3.8}$$

If X has cdf F then $F(X)$ is a uniform random variable and it follows from (3.6) and (3.8) that

$$\begin{aligned} \sigma^2(H)/2 = k^2/12+(13k-10)(k-2)/192-\beta(3k^2-8k+4)/8 \\ +3\beta^2(k-2)^2/8+\beta^3(k^2-4)/6-\beta^4(k-2)^2/4 \end{aligned} \tag{3.9}$$

where $\beta = (2b)^{-1}$. It is clear that for sufficiently small β (large b) the right side of (3.9) can be made larger than $k^2/12$ so that (3.5) is satisfied.

Let $V_n = n^{\frac{1}{2}}[(T'_{kn}-T'_{k-1,n})-(T_{kn}-T_{k-1,n})]$. We now prove that the inequality in (3.2) holds for the randomized rank sum procedure by showing that $EV_n^2 \to 0$.

If we let a_{ij} equal one or zero according as the jth smallest of all N observations is or is not a member of the ith sample, then

$$\frac{n}{N+1} V_n = n^{-\frac{1}{2}} \sum_{j=1}^{N} \left[Z(j) - \frac{j}{N+1} \right] (a_{kj} - a_{k-1,j}).$$

Since $\theta_k = \theta_{k-1}$ and $a_{ij} = 1$ if and only if $R_{il} = j$ for some l, it follows that

$$\begin{aligned} Ea_{kj} &= Ea_{k-1,j} = nP[R_{k1} = j] \\ Ea_{kj} \cdot a_{kj'} &= n(n-1)P[R_{k1} = j, R_{k2} = j'] \qquad \text{and} \\ Ea_{kj} \cdot a_{k-1,j'} &= n^2 P[R_{k1} = j, R_{k-1,1} = j'] \\ &= n^2 P[R_{k1} = j, R_{k2} = j']. \end{aligned}$$

Routine calculations yield

$$\begin{aligned} \left(\frac{n}{N+1}\right)^2 EV_n^2 &= 2\textstyle\sum_{j=1}^{N} \operatorname{Var}[Z(j)] \cdot P[R_{k1} = j] \\ &\quad -4\textstyle\sum\sum_{j<j'} \operatorname{Cov}[Z(j), Z(j')] \cdot P[R_{k1} = j, R_{k2} = j'] \\ &\leqq \max_j \operatorname{Var}[Z(j)] \leqq [4(N+2)]^{-1}, \end{aligned}$$

so that $EV_n^2 \to 0$ and the proof is complete.

4. Concluding remarks. Procedures $R(n)$ and $R'(n)$ are special cases of the scores procedures proposed in [2], [4], [6], [7] and [8]. The second counterexample probably works for any scores procedure when F (instead of being logistic) is the cdf against which the scores are locally most powerful, but it is not clear that a counterexample like 3 can be constructed if non-uniform scores are used.

REFERENCES

[1] BARR, D. R. and RIZVI, M. H. (1966). An introduction to ranking and selection procedures. *J. Amer. Statist. Assoc.* **61** 640–646.

[2] BARTLETT, N. S. and GOVINDARAJULU, Z. (1968). Some distribution-free statistics and their application to the selection problem. *Ann. Inst. Statist. Math.* **20** 79–97.

[3] FRASER, D. A. S. (1957). *Nonparametric Methods in Statistics.* Wiley, New York.

[4] LEHMANN, E. L. (1963). A class of selection procedures based on ranks. *Math. Ann.* **150** 268–275.

[5] OLSHEN, RICHARD A. (1967). Sign and Wilcoxon tests for linearity. *Ann. Math. Statist.* **38** 1759–1769.

[6] PURI, M. L. and PURI, P. S. (1968). Selection procedures based on ranks: scale parameter case. *Sankhyā Ser. A* **30** 291–302.

[7] PURI, M. L. and PURI, P. S. (1969). Multiple decision procedures based on ranks for certain problems in analysis of variance. *Ann. Math. Statist.* **40** 619–632.

[8] WOODWORTH, G. G. (1965). An extension of a result of Lehmann on the asymptotic efficiency of selection procedures based on ranks. Technical Report No. 66, Department of Statistics, Univ. of Minnesota.