*The Annals of Mathematical Statistics*
1969, Vol. 40, No. 2, 619–632

# MULTIPLE DECISION PROCEDURES BASED ON RANKS FOR CERTAIN PROBLEMS IN ANALYSIS OF VARIANCE

BY MADAN L. PURI[1] AND PREM S. PURI[2]

*Indiana University, Bloomington and University of California, Berkeley*

**1. Summary.** This paper is concerned with single-sample multiple-decision procedures based on the ranks of the observations for selecting from $c$ continuous populations (a) the "best $t$" populations *without* regard to order, (b) the "best $t$" populations *with* regard to order, and (c) a subset which contains all populations "as good or better than a standard one." The "bestness" of a population is characterised by its location parameter; the best population being the one having the largest location parameter; the second best being the one having the second largest location parameter, etc. Large-sample methods are provided for computing the sample sizes necessary to guarantee a preassigned probability of correct grouping (or ranking) under specified conditions on location parameters. It is shown that the asymptotic efficiency of these procedures relative to the normal theory procedures (see, for example, Bechhofer [1] and Gupta and Sobel [3]) is the same as that of the associated tests in one-way analysis of variance model I problem. If the ratio of the sample sizes is equal to this efficiency, the two procedures being compared are shown to have the same asymptotic performance characteristic. Finally, in the case of problem (c) two alternative rank-score procedures are proposed which are asymptotically equi-efficient.

**2. Introduction.** It is well known that in most of the practical situations to which the analysis of variance tests are applied, they do not supply the information that the experimenter aims at. If, for example, the hypothesis is rejected in actual application of the $F$-test, the resulting conclusion that the true means $\theta_1, \theta_2, \cdots, \theta_c$ are not all equal, would by itself usually be insufficient to satisfy the experimenter. In fact his problems would begin at this stage. He may desire to select the "best" population or a group of the "best" population; he may like to rank the populations in order of "goodness" or he may like to draw some other inferences about the parameters of interest to him.

In parametric theory most of the work along these lines has been done by Bechhofer [1], [2], Gupta [4], Gupta and Sobel [3], Lehmann [5], and Paulson [7], among others. (See Lehmann [5] and Bechhofer [1] for references.) In nonparametric theory, attempts to meet the need of decision procedures relevant to such problems has been initiated by Lehmann [6] only very recently. In [6], Lehmann

Received 3 January 1966; revised 10 June 1968.

[1] This research was supported by U. S. Navy Contract N00014-67-A-0467-0004 and National Institute of Mental Health, Grant No. MH 14934-02.

[2] This investigation was supported (in part) by a research grant (No. GM-10525) from the National Institutes of Health, Public Health Service.

Now at Purdue University.

considered the problem of selecting the "best" population, that is, the one having the largest location parameter. He proposed a class of decision procedures which may be regarded as the most direct predecessor of the procedures considered here. In this paper we extend Lehmann's procedures so as to cover the following problems:

(a) The problem of selecting out of $c$ populations, the "$t$ best" ones *without* regard to order.

(b) The problem of selecting out of $c$ populations, the "$t$ best" ones *with* regard to order.

(c) The problem of finding a subset which contains all populations "as good or better than a standard one."

The procedures developed here are compared with the normal theory procedures studied by Bechhofer [1] and Gupta and Sobel [3]. Following Lehmann [6], the asymptotic relative efficiency of two procedures is defined as the ratio of the sample sizes required to achieve the same minimum probability of selecting the desired group. The same asymptotic relative efficiencies are obtained as for the problem of testing the equality of means in one-way analysis of variance model (see Puri [8]). The analogue of the efficiency statements for the Kruskal-Wallis and normal score tests [8] applies. It is also shown that the procedures based on rank scores are robust for selecting appropriate sample sizes.

**3. The mathematical model and related definitions.** Let $X_{ij}$ ($j = 1, 2, \cdots, n$; $i = 1, 2, \cdots, c$) be independent samples from populations $\Pi_1, \Pi_2, \cdots, \Pi_c$, with continuous cumulative distribution functions $F(x - \theta_i)$, $i = 1, 2, \cdots, c$. Let $\theta_{[1]} \leqq \theta_{[2]} \leqq \cdots \leqq \theta_{[c]}$ be the ranked $\theta$'s. We assume that it is not known which population is associated with $\theta_{[i]}$. We further assume that a population is characterised by its parameter value. Thus the "best" population is the one which has the largest parameter value; the "second best" being the one which has the second largest parameter value, and so on. Our aim is to develop some procedures based on the ranks of the observations for the problems (a), (b) and (c) mentioned in Section 2.

We shall denote the sample mean from $i$th population by $\bar{X}_i$; the sample mean associated with the population having the mean $\theta_{[i]}$ by $\bar{X}_{(i)}$ and the ranked $\bar{X}_i$, $i = 1, 2, \cdots, c$, by $\bar{X}_{[1]} < \bar{X}_{[2]} < \cdots < \bar{X}_{[c]}$.

**4. Problem (a).**

4A. *Bechhofer procedure.* When $F$ is normal, Bechhofer [1] has proposed the following procedure: Select the $t$ populations associated with

$$\bar{X}_{[c-t+1]}, \cdots, \bar{X}_{[c]}. \tag{4A.1}$$

Most of the literature on selection problems is concerned with the determination of sample sizes required to guarantee a preassigned probability of a correct grouping, say $\gamma$, so that

$$P(\text{correct selection of } t \text{ best populations}) \geqq \gamma, \tag{4A.2}$$

where $\theta_{[c-t]}$ and $\theta_{[c-t+1]}$ are subject to the condition that

$$\theta_{[c-t+1]} - \theta_{[c-t]} \geqq \Delta^*. \tag{4A.3}$$

Here $\Delta^*$ is a given constant, and denotes the smallest value of the difference $\theta_{[c-t+1]} - \theta_{[c-t]}$ which is "worth detecting."

If the assumption of normality is considered unreasonable, we can find a large sample solution of this problem, which as will be seen below, depends only on the variance $\sigma^2$ of $F$. To this end, as in [6] consider a sequence of situations for increasing $n$ and define the condition (4A.3) as

$$\theta_{[c-t+1]} - \theta_{[c-t]} \geqq \Delta^{(n)}, \tag{4A.4}$$

where $\Delta^{(n)}$ is given by (4A.7). Then for all $F$ the left hand side of (4A.2) takes on its minimum value when the following holds:

$$\begin{aligned} \theta_{[1]} = \theta_{[2]} = \cdots = \theta_{[c-t]} &= \theta_{[c-t+1]} - \Delta^{(n)}; \\ \theta_{[c]} - \theta_{[c-t+1]} &= 0. \end{aligned} \tag{4A.5}$$

This follows from the stochastic increasing property of the family of distributions $F(x - \theta_i)$. We refer to the condition (4A.5) as the least favorable configuration of $\theta$'s. The sample size $n$ is determined by the relation

$$P[\max(\bar{X}_{(1)}, \cdots, \bar{X}_{(c-t)}) < \min(\bar{X}_{(c-t+1)}, \cdots, \bar{X}_{(c)})] = \gamma, \tag{4A.6}$$

under the assumption that (4A.5) holds.

The following lemma gives the large sample solution of the sample size problem.

LEMMA 4A1. *For fixed $\gamma$, and under the condition* (4A.4), *let $n$ be determined so that* (4A.5) *and* (4A.6) *hold. Then as $n \to \infty$,*

$$\Delta^{(n)} = \Delta\sigma n^{-\frac{1}{2}} + o(n^{-\frac{1}{2}}). \tag{4A.7}$$

*Here $\sigma^2$ is the variance of $F$ and $\Delta$ is determined by the condition*

$$\gamma = tQ_{c-1}(\underbrace{\Delta 2^{-\frac{1}{2}}, \Delta 2^{-\frac{1}{2}}, \cdots, \Delta 2^{-\frac{1}{2}},}_{(c-t) \text{ times}} \underbrace{0, 0, \cdots, 0)}_{(t-1) \text{ times}} \tag{4A.8}$$

*where $Q_{c-1}$ is the cumulative distribution function of a normally distributed vector $(U_1, \cdots, U_{c-t}, W_{c-t+2}, \cdots, W_c)$ with*

$$\begin{aligned} E(U_i) = E(W_j) = 0; \qquad &\operatorname{Cov}(U_i, U_{i'}) = \tfrac{1}{2}(\delta_{ii'} + 1); \\ \operatorname{Cov}(W_j, W_{j'}) = \tfrac{1}{2}(\delta_{jj'} + 1), \qquad &\operatorname{Cov}(U_i, W_j) = -\tfrac{1}{2}; \\ i, i' = 1, 2, \cdots, c - t; \; &j, j' = c - t + 2, \cdots, c; \end{aligned} \tag{4A.9}$$

*where $\delta$'s are the Kronecker deltas.*

The proof of this lemma, being an immediate extension of Lehmann's Lemma 1 [6], is omitted. A similar expression for (4A.8) which can actually be used for calculation is either of the following:

$$\begin{aligned} \gamma &= \int_{-\infty}^{\infty} \Phi^{c-t}(x + \Delta)[1 - \Phi(x)]^{t-1} \, d\Phi(x), \\ &= (c - t) \int_{-\infty}^{\infty} \Phi^{c-t-1}(x)[1 - \Phi(x - \Delta)]^{t} \, d\Phi(x), \end{aligned} \tag{4A.10}$$

where $\Phi$ is standard normal distribution function. The solution in terms of $\Delta$ for fixed $c$, $t$ and $\gamma$ has been tabulated by Bechhofer [1].

Suppose now that we are given a value $\Delta^*$ and we wish to find the smallest sample size $n$ for which (4A.2) holds subject to (4A.3). Then, from the above lemma, a large sample size solution is given by

$$n = (\Delta\sigma/\Delta^*)^2. \tag{4A.11}$$

4B. *Procedures based on ranks.* We shall now consider the procedures based on the ranks of the observations. Let $Z_{N,j}^{(i)} = 1$, if the $j$th smallest of $N = cn$ observations $X_{ij}$; $j = 1, 2, \cdots, n$; $i = 1, 2, \cdots, c$ is from the $i$th sample and otherwise $Z_{N,j}^{(i)} = 0$. Denote

$$T_i = n^{-1} \textstyle\sum_{j=1}^{N} E(V^{(j)}) Z_{N,j}^{(i)}; \qquad i = 1, 2, \cdots, c, \tag{4B.1}$$

where $V^{(1)} < \cdots < V^{(N)}$ is an ordered sample from a given continuous distribution $F_0$ and $E$ denotes the expectation. Further as in Section 3, denote the statistic associated with the population having the mean $\theta_{[i]}$ by $T_{(i)}$ and let the ranked $T_i$'s be denoted as $T_{[1]} < T_{[2]} < \cdots < T_{[c]}$. Then the rank scores procedure proposed is as follows: Select the $t$ populations associated with

$$T_{[c-t+1]}, \cdots, T_{[c]}. \tag{4B.2}$$

The procedures based on the statistic $T_i$'s will be referred as $F_0$-*scores procedure* $T(F_0)$.

In what follows, we shall judge the relative merits of the procedures (4A.1) and (4B.2) on the basis of the sample sizes required to guarantee (4A.2) subject to the condition that the differences of any two $\theta$'s is of order $m^{-\frac{1}{2}}$ where for clarity the sample size for procedure $T(F_0)$ is denoted by $m$. Specifically, consider any sequence of parameter points satisfying

$$\theta_{[c-t+1]} - \theta_{[i]}^{(m)} \equiv \Delta_i^{(m)} = \Delta_i \cdot m^{-\frac{1}{2}} + o(m^{-\frac{1}{2}}), \tag{4B.3}$$
$$i = 1, 2, \cdots, c;\ i \neq c - t + 1,$$

where the $\Delta$'s are some constants, positive for $i = 1, 2, \cdots, c - t$ and negative for $i \geqq c - t + 2$. Suppose now that the scores-procedure (4B.2) requires the smallest sample size $m$ if it is to satisfy (4A.2) subject to the conditions (4B.3) and

$$\theta_{[c-t+1]}^{(m)} - \theta_{[c-t]}^{(m)} \geqq \tilde{\Delta}^{(m)}, \tag{4B.4}$$

where $\tilde{\Delta}^{(m)}$ is given by (4B.15). Then to obtain the least favorable configuration of $\theta$'s (subject to (4B.3)) for the procedure $T(F_0)$, we have the following theorem, the proof of which being an immediate consequence of Theorem 6.1 of [8], is omitted.

THEOREM 4B.1. *For* $m = 1, 2, \cdots$, *let* $X_{ij}$ $(j = 1, 2, \cdots, m; i = 1, 2, \cdots, c)$ *be independently distributed according to* $F_i(x) = F(x - \theta_i^{(m)})$ *with the sequence of parameter points* $\theta^{(m)} = (\theta_1^{(m)}, \cdots, \theta_c^{(m)})$ *and suppose that the assumptions of*

*Theorem* 6.1 *and Lemma* 7.2 *of* [8] *are satisfied. Then the limiting distribution of the random vector*

$$(m/2A^2)^{\frac{1}{2}}[T_{(i)} - T_{(j)} - \mu_i(\theta^{(m)}) + \mu_j(\theta^{(m)}),$$
$$i = 1, 2, \cdots, c - t; j = c - t + 1, \cdots, c],$$

*where*

(4B.5) $\mu_i(\theta^{(m)}) = \int J[H(x)]\, dF_i(x), \qquad H(x) = \sum_{i=1}^{c} F_i(x)/c, \qquad J = F_0^{-1},$

*is the distribution of a* $t(c - t)$ *dimensional normal vector* $(U_{ij}; i = 1, 2, \cdots, c - t; j = c - t + 1, \cdots, c)$ *with*

(4B.6) $$E(U_{ij}) = 0, \quad \mathrm{Var}\,(U_{ij}) = 1, \quad \mathrm{Cov}\,(U_{ij}, U_{ij'}) = \mathrm{Cov}\,(U_{ij}, U_{i'j}) = \tfrac{1}{2},$$
$$\mathrm{Cov}\,(U_{ij}, U_{i'j'}) = 0,$$
$$i \neq i', j \neq j'; i, i' = 1, 2, \cdots, c - t; j, j' = c - t + 1, \cdots, c,$$

*where*

(4B.7) $$A^2 = \int_0^1 J^2(x)\, dx - \left(\int_0^1 J(x)\, dx\right)^2.$$

Now the probability of correct selection of $t$ best populations is given by

(4B.8) $$P[\max\,(T_{(1)}, \cdots, T_{(c-t)}) < \min\,(T_{(c-t+1)}, \cdots, T_{(c)})]$$
$$= P[T_{(i)} - T_{(j)} < 0, i = 1, 2, \cdots, c - t; j = c - t + 1, \cdots, c]$$
$$= P[(m/2A^2)^{\frac{1}{2}}(T_{(i)} - T_{(j)} - \mu_{(i)}(\theta^{(m)}) + \mu_{(j)}(\theta^{(m)}))$$
$$< (m/2A^2)^{\frac{1}{2}}(\mu_{(j)}(\theta^{(m)}) - \mu_{(i)}(\theta^{(m)})),$$
$$i = 1, 2, \cdots, c - t; j = c - t + 1, \cdots, c].$$

Since, for large $m$ (c.f. Lemma 7.2 [8])

(4B.9) $$m^{\frac{1}{2}}(\mu_{(j)}(\theta^{(m)}) - \mu_{(i)}(\theta^{(m)})) \sim m^{\frac{1}{2}}B(\theta_{[j]}^{(m)}) - \theta_{[i]}^{(m)})$$

where

(4B.10) $$B = \int (d/dx)J[F(x)]\, dF(x),$$

the right hand side of (4B.8) by virtue of Theorem 4B.1 is asymptotically equivalent to

(4B.11) $$P[U_{ij} < m^{\frac{1}{2}}B(\theta_{[j]}^{(m)} - \theta_{[i]}^{(m)})(2A^2)^{-\frac{1}{2}};$$
$$i = 1, 2, \cdots, c - t; j = c - t + 1, \cdots, c].$$

Now in view of (4B.4), $\theta_{[i]}^{(m)}$'s satisfy

(4B.12) $$\theta_{[1]}^{(m)} < \theta_{[2]}^{(m)} < \cdots < \theta_{[c-t]}^{(m)} < \theta_{[c-t+1]}^{(m)} - \tilde{\Delta}^{(m)} < \theta_{[c-t+1]}^{(m)} < \cdots < \theta_{[c]}^{(m)}.$$

Among the values of $\theta$'s which satisfy (4B.3) and (4B.12), the least favorable configuration of $\theta$'s for which (4B.8) takes on the minimum value is easily seen from (4B.11) to be

$$\theta_{[1]}^{(m)} = \cdots = \theta_{[c-t]}^{(m)} = \theta_{[c-t+1]}^{(m)} - \tilde{\Delta}^{(m)}; \qquad \theta_{[c]}^{(m)} - \theta_{[c-t+1]}^{(m)} = 0. \tag{4B.13}$$

Hence for large samples, the sample size $m$ is determined by the condition

$$P[\max(T_{(1)}, \cdots, T_{(c-t)}) < \min(T_{(c-t+1)}, \cdots, T_{(c)})] = \gamma, \tag{4B.14}$$

where the left hand side is derived subject to (4B.13). For the case of $T(F_0)$-procedure, the following lemma is the analogue of Lemma 4A.1.

LEMMA 4B.1. *For fixed* $\gamma$, *let* $m$ *be determined so that* (4B.14) *holds subject to* (4B.13), *and suppose that* $F$ *and* $J = F_0^{-1}$ *satisfy the regularity conditions of Theorem* 6.1 *and Lemma* 7.2 *of* [8]. *Then as* $m \to \infty$

$$\tilde{\Delta}(m) = \Delta m^{-\frac{1}{2}} A / \int (d/dx)\{J[F(x)]\}\, dF(x) + o(m^{-\frac{1}{2}}), \tag{4B.15}$$

*where* $A^2$ *is given by* (4B.7) *and* $\Delta$ *satisfies* (4A.8).

PROOF. Let

$$F_i(x) = F(x - \theta_i^{(m)}) \tag{4B.16}$$

where $\theta_i^{(m)} = \theta_{[c-t+1]} - \Delta_i^{(m)}$; $\Delta_i^{(m)}$ are as defined in (4B.3) with $\Delta_{c-t+1}^{(m)} = 0$. Denote

$$\mu_i = \int J[H(x)]\, dF_i(x), \qquad J = F_0^{-1}, \tag{4B.17}$$

where $H(x) = \sum_{i=1}^{c} F_i(x)/c$.

Then from [8], the random variables $m^{\frac{1}{2}}(T_i - \mu_i(\theta^{(m)}))$, $i = 1, 2, \cdots, c$, have asymptotically a joint normal distribution with zero means and covariance matrix

$$\sigma_{ii'} = (\delta_{ii'} - c^{-1})A^2, \tag{4B.18}$$

where the $\delta_{ii'}$ are the Kronecker deltas, $A^2$ is given by (4B.7) and where $\theta^{(m)} = (\theta_1^{(m)}, \cdots, \theta_c^{(m)})$.

Now the equation (4B.14) in the limit is equivalent to

$$\begin{aligned} \gamma &= \lim_{m\to\infty} \textstyle\sum_{r=c-t+1}^{c} P[\max(T_{(1)}, \cdots, T_{(c-t)}) \\ &\qquad < \min(T_{(l)}; l \neq r, l = c-t+1, \cdots, c)] \\ &= \lim_{m\to\infty} \textstyle\sum_{r=c-t+1}^{c} P[(T_{(k)} - T_{(r)} - \eta_{k,r}(\theta^{(m)}))/A2^{\frac{1}{2}} \\ &\qquad \leqq \eta_{r,k}(\theta^{(m)})/A\, 2^{\frac{1}{2}}; (T_{(r)} - T_{(l)} - \eta_{r,l}(\theta^{(m)}))/A2^{\frac{1}{2}} \leqq \eta_{l,r}(\theta^{(m)})/A2^{\frac{1}{2}}; \\ &\qquad k = 1, 2, \cdots, c-t;\ l \neq r;\ l = c-t+1, \cdots, c], \end{aligned} \tag{4B.19}$$

where

$$\eta_{a,b}(\theta^{(m)}) = \mu_{(a)}(\theta^{(m)}) - \mu_{(b)}(\theta^{(m)}). \tag{4B.20}$$

From Theorem 6.1 and Lemma 7.2 of [8], it follows that the random variables $m^{\frac{1}{2}}(T_{(k)} - T_{(r)} - \eta_{k,r}(\theta^{(m)}))/A2^{\frac{1}{2}}$ and $m^{\frac{1}{2}}(T_{(r)} - T_{(l)} - \eta_{r,l}(\theta^{(m)}))/A2^{\frac{1}{2}}$; $k = 1, 2, \cdots, c - t; l \neq r; l = c - t + 1, \cdots, c$, have the limiting joint distribution of a $(c - 1)$-dimensional normal vector $(U_1, \cdots, U_{c-t}, W_{c-t+1}, \cdots, W_{r-1}, W_{r+1}, \cdots, W_c)$ satisfying (4A.9) with $i, i' = 1, 2, \cdots, c - t; j, j' = c - t + 1, \cdots, c; j, j' \neq r$.

Now under the least favorable configuration (4A.5) with $\Delta^{(n)}$ equated to $\tilde{\Delta}^{(m)}$, we have $\eta_{l,r}(\theta^{(m)}) = 0$ for $l \neq r$, $l = c - t + 1, \cdots, c$, and $\eta_{r,k}(\theta^{(m)})$ is independent of $r$ and $k$ for $r = c - t + 1, \cdots, c$ and $k = 1, 2, \cdots, c - t$. Thus using (4A.8) and (4B.9), we obtain $\lim_{m\to\infty} \eta_{r,k}(\theta^{(m)})/A2^{\frac{1}{2}} = \Delta/2^{\frac{1}{2}}$, that is, omitting details

$$\lim_{m\to\infty} m^{\frac{1}{2}}\tilde{\Delta}^{(m)} \int (d/dx)\{J[F(x)]\}\, dF(x)/A2^{\frac{1}{2}} = \Delta/2^{\frac{1}{2}}.$$

The lemma follows.

4C. *Asymptotic relative efficiency.* We are now in a position to make large sample comparison between the score-procedure $T(F_0)$ and Bechhofer procedure (hereafter referred as $B$-procedure). We shall adopt a method developed by Lehmann [6] who defined the asymptotic relative efficiency of the two procedures as the limiting ratio of the sample sizes to attain the same minimum probability of correct selection subject to the same condition (4A.3) in both the cases.

THEOREM 4C.1. *The asymptotic efficiency of the $T(F_0)$ procedure relative to the $B$-procedure is*

$$e_{T(F_0),B}(F) = \sigma^2 A^{-2}\left(\int (d/dx)\{J[F(x)]\}\, dF(x)\right)^2. \tag{4C.1}$$

The proof follows by equating $\tilde{\Delta}^{(m)}$ defined by (4B.15) to $\Delta^{(n)}$ defined by (4A.7), as both are set equal to $\Delta^*$.

The relative efficiency $e_{T(F_0),B}$ of the $F_0$-scores procedure $T(F_0)$ to the Bechhofer procedure is the same as that found by Puri [8] for the corresponding tests in the $c$-sample problem and also shown by Lehmann [6] to be valid for the problem of selecting the best population. For the ease of reference we give below the efficiency comparisons of the $T(F_0)$ and $B$-procedures for different $F$'s and two $F_0$'s.

| $F_0$ | $F$ | | |
|---|---|---|---|
| | Normal | Uniform | Double exponential |
| Normal $N(0, 1)$ | 1 | $\infty$ | $4/\pi \backsim 1.273$ |
| Uniform $R(0, 1)$ | .955 | 1 | 0.927 |

More generally, if $F_0$ is a uniform distribution over $(0, 1)$, $e_{T(F_0),B} \geqq 0.864$ for all $F$, and if $F_0$ is a standard $N(0, 1)$ distribution, $e_{T(F_0),B} > 1$ for all nonnormal $F$.

4D. *Comparison of performance characteristis.* We have shown above that the $T(F_0)$ and $B$-procedures have approximately the same performance characteristic if the sample sizes $m$ and $n$ are determined such that (4C.1) holds and the parameter point satisfies (4A.5) or equivalently (4B.13) since $\Delta^{(n)} = \tilde{\Delta}^{(m)}$. We shall now study the performance characteristics of the two procedures when the parameter point does not satisfy (4A.5).

To discuss this, consider any sequence of parameter points satisfying

$$\theta_{[c-t+1]} - \theta_{[i]}^{(n)} = \Delta_i^{(n)} = \Delta_i\sigma n^{-\frac{1}{2}} + o(n^{-\frac{1}{2}}), \quad i = 1, 2, \cdots, c; \quad i \neq c - t + 1, \tag{4D.1}$$

where not all $\Delta_i = \Delta$ for $i = 1, 2, \cdots, c - t$ and/or not all $\Delta_i = 0$ for $i = c - t + 1, \cdots, c$. However, the condition (4A.4) still holds, so that

$$\theta_{[1]}^{(n)} < \theta_{[2]}^{(n)} < \cdots < \theta_{[c-t]}^{(n)} < \theta_{[c-t+1]}^{(n)} - \Delta^{(n)} < \theta_{[c-t+1]}^{(n)} < \cdots < \theta_{[c]}^{(n)}. \tag{4D.2}$$

Then for the $B$-procedure

$$\begin{aligned}
&\lim_{n\to\infty} P(\text{correct selection of } t \text{ best populations}) \\
&= \lim_{n\to\infty} P[\max(\bar{X}_{(1)}, \cdots, \bar{X}_{(c-t)}) < \min(\bar{X}_{(c-t+1)}, \cdots, \bar{X}_{(c)})] \\
&= \lim_{n\to\infty} \textstyle\sum_{l=c-t+1}^{c} P[\max(\bar{X}_{(1)}, \cdots, \bar{X}_{(c-t)}) < \bar{X}_{(l)} \\
&\qquad < \min(\bar{X}_{(r)}; r = c - t + 1, \cdots, c; r \neq l)] \\
&= \textstyle\sum_{l=c-t+1}^{c} \lim_{n\to\infty} P[\bar{X}_{(s)} - \bar{X}_{(l)} < 0; s = 1, 2, \cdots, c - t, \\
&\qquad \bar{X}_{(l)} - \bar{X}_{(r)} < 0; r = c - t + 1, \cdots, c, r \neq l] \\
&= \textstyle\sum_{l=c-t+1}^{c} \lim_{n\to\infty} P[U_s < \xi_{l,s}^{(n)}; W_r < \xi_{r,l}^{(n)}; s = 1, 2, \cdots, c - t; \\
&\qquad r = c - t + 1, \cdots, c; r \neq l]
\end{aligned} \tag{4D.3}$$

where $\xi_{ij}^{(n)} = (n/2\sigma^2)^{\frac{1}{2}}(\theta_{[i]}^{(n)} - \theta_{[j]}^{(n)})$, and where the last equality follows from the central limit theorem and the fact that the convergence is uniform with respect to the argument of the distribution function, and where the vector $(U_1, \cdots, U_{c-t}, W_{c-t+1}, \cdots, W_{l-1}, W_{l+1}, \cdots, W_c)$ is normally distributed with the cdf given by $Q_{c-1}$ and with the variance covariance matrix given by (4A.9). Furthermore, using (4D.1)

$$\begin{aligned}
\xi_{l,s}^{(n)} &= (1/2^{\frac{1}{2}})\rho_{s,l} + o(1); & s &= 1, 2, \cdots, c - t, \\
\xi_{r,s}^{(n)} &= (1/2^{\frac{1}{2}})\rho_{l,r} + o(1); & r &= c - t + 1, \cdots, c; \quad r \neq l,
\end{aligned} \tag{4D.4}$$

with $\rho_{ij} = (\Delta_i - \Delta_j)$; whence (4D.3) is given by

$$\textstyle\sum_{l=c-t+1}^{c} Q_{c-1}(\rho_{1,l}/2^{\frac{1}{2}}, \cdots, \rho_{c-t,l}/2^{\frac{1}{2}}, \rho_{l,c-t+1}/2^{\frac{1}{2}}, \cdots, \rho_{l,l-1}/2^{\frac{1}{2}}, \rho_{l,l+1}/2^{\frac{1}{2}}, \cdots, \rho_{l,c}/2^{\frac{1}{2}}). \tag{4D.5}$$

Consider now the $T(F_0)$ procedure based on the samples of size $m = g(n)$ satisfying (4C.1). Then the following theorem gives the limiting behavior of this procedure.

THEOREM 4D.1. *For* $n = 1, 2, \cdots$, *let* $X_{ij}$ $(j = 1, 2, \cdots, n; i = 1, 2, \cdots, c)$ *be independently distributed according to* $F(x - \theta_i^{(n)})$ *with the sequence of parameter point* $\theta^{(n)} = (\theta_1^{(n)}, \cdots, \theta_c^{(n)})$ *and suppose that the assumptions of Theorem* 6.1

*and Lemma* 7.2 *of* [8] *are satisfied. Furthermore, let* $m = g(n)$ *be determined so that* (4C.1) *holds. Then if, for any fixed* $l$ *between* $c - t + 1$ *and* $c$,

$$V_s^{(n)} = (\tfrac{1}{2}nA^{-2})^{\frac{1}{2}}[T_{(s)} - T_{(l)} - \eta_{s,l}(\theta^{(m)})]; \qquad s = 1, 2, \cdots, c - t,$$

$$V_r^{(n)} = (\tfrac{1}{2}nA^{-2})^{\frac{1}{2}}[T_{(l)} - T_{(r)} - \eta_{l,r}(\theta^{(m)})]; \quad r = c - t + 1, \cdots, c; \quad r \neq l,$$

*where* $T$*'s and* $\eta$*'s are defined by* (4B.1) *and* (4B.20) *respectively, the joint limiting distribution of the random variables* $(V_1^{(n)}, \cdots, V_{c-t}^{(n)}, V_{c-t+1}^{(n)}, \cdots, V_{l-1}^{(n)}, V_{l+1}^{(n)}, \cdots, V_c^{(n)})$ *is the distribution of a* $(c - 1)$ *dimensional normal vector* $(U_1, U_2, \cdots, U_{c-t}, W_{c-t+1}, \cdots, W_{l-1}, W_{l+1}, \cdots, )$ *satisfying* (4A.9).

The proof of this theorem is analogous to that of Theorem 1 [6] and is therefore omitted.

Finally for the parameter points satisfying (4D.1) and for $l$ between $c - t + 1$ and $c$,

$$\lim_{n\to\infty} (\tfrac{1}{2}nA^{-2})^{\frac{1}{2}}\eta_{l,s}(\theta^{(m)}) = \rho_{s,l}/2^{\frac{1}{2}}; \qquad s = 1, 2, \cdots, c - t,$$

and

$$\lim_{n\to\infty} (\tfrac{1}{2}nA^{-2})^{\frac{1}{2}}\eta_{r,l}(\theta^{(m)}) = \rho_{l,r}2^{\frac{1}{2}}; \qquad r = c - t + 1, \cdots, c; \quad r \neq l,$$

and therefore in view of (4D.4) using the above theorem it can be easily seen that if the ratio of the sample sizes is equal to (4C.1), the $T(F_0)$ and $B$-procedures have the same asymptotic characteristics, given by (4D.5). The reader may refer [6] for analogous details.

**5. Problem (b).** We shall now consider the problem of selecting out of $c$ populations the "$t$ best" ones *with* regard to order. Our aim as in Problem (a) is to make large sample comparisons between the procedures based on the ranks of the observations and the means procedure proposed by Bechhofer [1]. The discussion runs parallel to the Problem (a) and is therefore indicated briefly.

5A. *Bechhofer procedure.* When $F$ is normal, the "$t$ best" populations *with* regard to order are the ones associated with $\bar{X}_{[c]}, \bar{X}_{[c-1]}, \cdots, \bar{X}_{[c-t+1]}$ respectively. As before, our object is to choose the sample sizes in such a way that under specified conditions, the proportion of correct statements associated with the decision procedure is at least some predetermined value $\gamma$; that is

(5A.1) $P$ (correct selection of $t$ best populations with regard to order) $\geqq \gamma$,

where $\theta$'s are subject to the conditions

$$\theta_{[i+1]} - \theta_{[i]} \geqq \delta^{(n)}; \qquad i = c - t, \cdots, c - 1, \tag{5A.2}$$

and where $\delta^{(n)}$ is given by (5A.5). Then for all $F$, the least favorable configuration of $\theta$'s is given by

$$\theta_{[1]} = \theta_{[2]} = \cdots = \theta_{[c-t]} = \theta_{[c-t+1]} - \delta^{(n)}; \tag{5A.3}$$

$$\theta_{[i]} - \theta_{[i+1]} = \delta^{(n)}; \quad i = c - t + 1, \cdots, c - 1;$$

and the sample size is therefore determined by the condition

$$P[\max(\bar{X}_{(1)}, \cdots, \bar{X}_{(c-t)}) < \bar{X}_{(c-t+1)} < \cdots < X_{(c)}] = \gamma \tag{5A.4}$$

where $(\theta_{[1]}, \cdots, \theta_{[c]})$ satisfies (5A.3).

The large-sample solution of the sample size problem then follows from the following lemma, the proof of which is omitted.

LEMMA 5A.1. *For fixed* $\gamma$, *and under the condition* (5A.2), *let* $n$ *be determined so that* (5A.3) *and* (5A.4) *hold. Then as* $n \to \infty$

$$\delta^{(n)} = \delta\sigma n^{-\frac{1}{2}} + o(n^{-\frac{1}{2}}). \tag{5A.5}$$

*Here* $\sigma^2$ *is the variance of* $F$ *and* $\delta$ *is determined by the condition*

$$(c - t)Q_{c-1}(\underbrace{0, 0, \cdots, 0,}_{c-t-1 \text{ times}} \underbrace{\delta 2^{-\frac{1}{2}}, \delta 2^{-\frac{1}{2}}, \delta 2^{-\frac{1}{2}})}_{t \text{ times}} = \gamma; \quad 1 \leqq t \leqq c - 1; \tag{5A.6}$$

*where* $Q$ *is the cumulative distribution function of a normally distributed vector* $(U_1, \cdots, U_{c-t-1}, W_{c-t}, \cdots, W_{c-1})$ *satisfying*

$$\begin{aligned} &E(U_i) = E(W_j) = 0; \quad \operatorname{Cov}(U_i, U_{i'}) = \tfrac{1}{2}(1 + \delta_{ii'}); \\ &\operatorname{Cov}(W_j, W_{j'}) = 1 \quad \text{or} \quad -\tfrac{1}{2} \quad \text{according as} \quad j = j' \quad \text{or} \quad |j - j'| = 1 \\ &\quad \text{and} \quad 0 \quad \text{otherwise}; \\ &\operatorname{Cov}(U_i, W_j) = -\tfrac{1}{2} \quad \text{if} \quad j = c - t \quad \text{and zero otherwise}; \\ &i, i' = 1, 2, \cdots, c - t - 1; \qquad j, j' = c - t, \cdots, c - 1. \end{aligned} \tag{5A.7}$$

5B. *Rank procedures and their performance characteristic.* Let $T_i$, $i = 1, 2, \cdots, c$, be defined as in (4B.1). Then the $T(F_0)$ procedure for selecting "$t$ best" populations *with* regard to order calls for the selection of the populations associated with $T_{[c]}, T_{[c-1]}, \cdots, T_{[c-t+1]}$.

As before, we shall confine ourselves to the set of $\theta$'s such that the difference of two of them is of order $m^{-\frac{1}{2}}$. More specifically, we shall consider sequences of parameter points satisfying

$$\begin{aligned} \theta_{[c-t+1]} - \theta_{[i]}^{(m)} &= \delta_i^{(m)} = \delta_i m^{-\frac{1}{2}} + o(m^{-\frac{1}{2}}); \qquad i = 1, 2, \cdots, c - t, \\ \theta_{[i+1]}^{(m)} - \theta_{[i]}^{(m)} &= \delta_{i+1,i}^{(m)} = \delta_{i+1,i} m^{-\frac{1}{2}} + o(m^{-\frac{1}{2}}); \\ &\qquad i = c - t + 1, \cdots, c - 1. \end{aligned} \tag{5B.1}$$

Then arguing as in Section 4B, it can be shown that the left side of (5A.1) subject to

$$\theta_{[i+1]}^{(m)} - \theta_{[i]}^{(m)} \geqq \bar{\delta}^{(m)}; \qquad i = c - t, \cdots, c - 1, \tag{5B.2}$$

takes on its minimum value when the least favorable configuration of $\theta$'s is given by

$$\begin{aligned} \theta_{[1]}^{(m)} &= \theta_{[2]}^{(m)} = \cdots = \theta_{[c-t]}^{(m)} = \theta_{[c-t+1]} - \bar{\delta}^{(m)} \\ \theta_{[i]}^{(m)} &= \theta_{[i+1]}^{(m)} - \bar{\delta}^{(m)}; \qquad i = c - t + 1, \cdots, c - 1, \end{aligned} \tag{5B.3}$$

where $\tilde{\delta}^{(m)}$ is given by (5B.5) below. Thus for large samples, the sample $m$ is determined by the condition

(5B.4) $$\Pr[\max(T_{(1)}, \cdots, T_{(c-t)}) < T_{(c-t+1)} < \cdots < T_{(c)}] = \gamma$$

where $(\theta_1^{(m)}, \cdots, \theta_c^{(m)})$ satisfies (5B.3). Furthermore proceeding as in Section 4B, it can be shown under the conditions of Lemma 4B.1, that

(5B.5) $$\tilde{\delta}^{(m)} = \delta m^{-\frac{1}{2}} A / \int (d/dx)\{J[F(x)]\}\, dF(x) + o(m^{-\frac{1}{2}}),$$

where $A$ and $\delta$ are as defined in (4B.7) and (5A.6) respectively.

Now if $m$ and $n$ for the two procedures are determined by (5A.1) and if $\delta^{(n)}$ of (5A.2) and $\tilde{\delta}^{(m)}$ of (5B.2) are equated to a common prefixed $\delta^*$, then proceeding as in Section 4C, the asymptotic efficiency of the scores procedure $T(F_0)$ with regard to order is the same as in (4C.1), and hence the concluding remarks of that section may be carried along to the present case.

Consider now any sequence of parameter points satisfying (5B.1) and not necessarily (5B.3). Then, proceeding along the lines of Section 4.D it can be shown that if the ratio of the sample sizes is equal to the efficiency (4C.1) the two procedures $T(F_0)$ and $B$ have the same asymptotic performance characteristic even with regard to order. The details are omitted to avoid repetition.

**6. Problem (c).** We shall now consider the problem of selecting a subset of $c$ populations such that the probability, that all populations "as good or better than the standard one" are included in the subset, is at least a predetermined number $\gamma$.

Let $\{X_{ij};\, i = 0, 1, 2, \cdots, c;\, j = 1, 2, \cdots, n\}$ be $c + 1$ independent samples from populations $\Pi_0, \Pi_1, \cdots, \Pi_c$ having continuous cumulative distribution functions $F(x - \theta_0), F(x - \theta_1), \cdots, F(x - \theta_c)$ respectively. We further assume that $\Pi_0$ represents the standard population. We shall call a population $\Pi_i$ "as good or better than $\Pi_0$", if

(6.1) $$\theta_i \geqq \theta_0 + \Delta^{(n)}$$

where $\Delta^{(n)}$ is given by (6A.5). Hereafter for convenience, we shall use the word "good" in the sense "as good or better than the standard one."

6A. *The means procedure.* For each $i = 1, 2, \cdots, c$ select $\Pi_i$ if and only if $\bar{X}_i \geqq \bar{X}_0$, where $\bar{X}_0$ and $\bar{X}_i$, $i = 1, 2, \cdots, c$, are the sample means corresponding to $\Pi_0, \Pi_i;\, i = 1, 2, \cdots, c$.

Suppose without loss of generality that the only good populations are $\Pi_1, \Pi_2, \cdots, \Pi_s$ $(s \leqq c)$, where $s$ is unknown. Then $\theta_i \geqq \theta_0 + \Delta^{(n)}$ for $i = 1, 2, \cdots, s$ and $\theta_i < \theta_0 + \Delta^{(n)}$ for $i = s + 1, \cdots, c$. As before, we wish to determine the minimum sample size $n$ such that

(6A.1) $$P\,[\text{selected subset includes } \Pi_1, \cdots, \Pi_s] \geqq \gamma.$$

We may remark that the purpose here is not to avoid the non-good populations for inclusion into the subset, but mainly to aim at the selection of good populations, subject to (6A.1). The fact that some of the non-good populations may get included into the subset is of no consequence to us. Again, when $s = 0$, the condition (6A.1) is trivially satisfied by any subset. Thus we restrict our attention to the case with $1 \leqq s \leqq c$. Since $s$ is unknown, we find a least favorable configuration jointly of both $\theta$'s as well as $s$, for which the left hand side of (6A.1) is minimum. This we attain in two stages. Firstly, for any fixed $s$, the least favorable configuration of $\theta$'s is

$$\theta_1 = \theta_2 = \cdots = \theta_s = \theta_0 + \Delta^{(n)}, \tag{6A.2}$$

the left hand side of (6A.1) being independent of $\theta_{s+1}, \cdots, \theta_c$; with the result

$$\begin{aligned}&\min_{\theta_i \geqq \theta_0 + \theta^{(n)}, i=1,2,\cdots,s} P\,[\text{selected subset includes} \\ &\qquad \Pi_1, \cdots, \Pi_s \mid \Pi_1, \cdots, \Pi_s \text{ are good}] \\ &\quad = P[\bar{X}_i > \bar{X}_0\,;\, i = 1, 2, \cdots, s \mid \theta_i = \theta_0 + \Delta^{(n)};\, i = 1, 2, \cdots, s].\end{aligned} \tag{6A.3}$$

Secondly, since right hand side of (6A.3) is a decreasing function of $s$, the least favorable value of $s$ is $s = c$. The sample size $n$ is therefore determined by the condition

$$\begin{aligned}\gamma &= \min_s \min_{\theta_i \geqq \theta_0 + \Delta^{(n)}, i=1,2,\cdots,s} [P \text{ selected subset includes} \\ &\qquad \Pi_1, \cdots, \Pi_s \mid \Pi_1, \cdots, \Pi_s \text{ are good}] \\ &= P[\bar{X}_i > \bar{X}_0\,;\, i = 1, 2, \cdots, c \mid \theta_i = \theta_0 + \Delta^{(n)};\, i = 1, 2, \cdots, c].\end{aligned} \tag{6A.4}$$

The large-sample solution of the sample size problem then follows from the following lemma.

Lemma 6A.1. *For fixed $\gamma$ and with "goodness" of a population defined by* (6.1) *let $n$ be determined so that* (6A.4) *holds. Then as $n \to \infty$,*

$$\Delta^{(n)} = \Delta\sigma n^{-\frac{1}{2}} + o(n^{-\frac{1}{2}}). \tag{6A.5}$$

*Here $\sigma^2$ is the variance of $F$ and $\Delta$ is determined by the condition*

$$Q_c(\Delta 2^{-\frac{1}{2}}, \cdots, \Delta 2^{-\frac{1}{2}}) = \gamma, \tag{6A.6}$$

*where $Q_c$ is the* cdf *of a normally distributed vector $(U_1, \cdots, U_c)$ satisfying*

$$EU_i = 0, \qquad \operatorname{Cov}(U_i, U_{i'}) = \tfrac{1}{2}(1 + \delta_{ii'}); \qquad i, i' = 1, 2, \cdots, c. \tag{6A.7}$$

The proof is straightforward.

6B. *Procedures based on ranks.* Let $Z_{M,r}^{(i)} = 1$ if the $r$th smallest of $M = m(c + 1)$ observations $X_{ij}$; $i = 0, 1, \cdots, c$; $j = 1, 2, \cdots, m$ is from the $i$th sample and otherwise let $Z_{M,r}^{(i)} = 0$. Denote

$$T_i = m^{-1} \sum_{r=1}^{M} E(V^{(r)}) Z_{M,r}^{(i)}\,; \qquad i = 0, 1, 2, \cdots, c; \tag{6B.1}$$

where $V^{(1)} < \cdots < V^{(M)}$ is an ordered sample from a given distribution $F_0$ and $E$ denotes the expectation. Then the proposed procedure is:

$$\text{Select } \Pi_i \text{ if and only if } T_i \geqq T_0\,; \qquad i = 1, 2, \cdots, c. \tag{6B.2}$$

Later, we shall propose an alternative procedure which is asymptotically equiefficient to the procedure (6B.2). To find the desired sample size $m$, as before we restrict ourselves to the $\theta$'s satisfying

$$\text{(6B.3)} \qquad \theta_i^{(m)} - \theta_0 = \Delta_i^{(m)} = \Delta_i m^{-\frac{1}{2}} + o(m^{-\frac{1}{2}}); \qquad i = 1, 2, \cdots, c.$$

Then subject to the condition

$$\text{(6B.4)} \qquad \theta_i^{(m)} \geqq \theta_0 + \tilde{\Delta}^{(m)},$$

the least favorable configuration of $\theta$'s (amongst those satisfying (6B.3)) and $s$ subject to (6A.1) turns out to be

$$\text{(6B.5)} \qquad s = c; \qquad \theta_i^{(m)} = \theta_0 + \tilde{\Delta}^{(m)}; \qquad i = 1, 2, \cdots, c,$$

where $\tilde{\Delta}^{(m)}$ under the conditions of Lemma 4B.1 is given by

$$\text{(6B.6)} \qquad \tilde{\Delta}^{(m)} = \Delta m^{-\frac{1}{2}} A / \int (d/dx)\{J[F(x)]\}\, dF(x) + o(m^{-\frac{1}{2}}).$$

Here $A$ and $\Delta$ are given by (4B.7) and (6A.6) respectively. The sample size $m$ is determined by the condition

$$\text{(6B.7)} \qquad P[T_i \geqq T_0\,; i = 1, 2, \cdots, c] = \gamma;$$

where the left hand side is derived under the condition that $(\theta_0, \theta_1^{(m)}, \cdots, \theta_c^{(m)})$ satisfies (6B.5).

The asymptotic efficiency of the means procedure relative to the procedure (6B.2) is again given by (4C.1) and finally, if we consider a sequence of parameter points $(\theta_0^{(m)}, \theta_1^{(m)}, \cdots, \theta_c^{(m)})$ not satisfying (6B.5), it can be shown as in the previous sections that if the ratio of the sample sizes $n/m$ is equal to the efficiency (4C.1), the two procedures have the same asymptotic performance. The details are omitted to avoid repetition of the argument.

6C. *An alternative rank procedure.* As an alternative to the procedure (6B.2), one might consider a procedure based on combining separately each of the $c$ samples, $X_{ij}$, $i = 1, 2, \cdots, c; j = 1, 2, \ldots, m$, with the sample $X_{0j}$, $j = 1, 2, \cdots, m$, from the standard population, instead of combining them all together. Let for $i = 1, 2, \cdots, c$, $Z_{m,r}^{(i,0)} = 1$, if the $r$th smallest of $2m$ observations

$$\{X_{ij}, X_{0j}, j = 1, 2, \cdots, m\},$$

is from the $i$th sample and zero otherwise. Denote for $i = 1, 2, \cdots, c$,

$$\tilde{T}_i = m^{-1} \textstyle\sum_{r=1}^{2m} E(V^{(r)}) Z_{m,r}^{(i,0)} \quad \text{and} \quad \tilde{T}_{i,0} = m^{-1} \sum_{r=1}^{2m} E(V^{(r)})(1 - Z_{m,r}^{(i,0)});$$

where $V^{(1)} < \cdots < V^{(2m)}$ is an ordered sample from a given distribution $F_0$. The alternative procedure is then given by

$$\text{(6C.1)} \qquad \text{Select } \Pi_i \text{ if and only if } \tilde{T}_i \geqq \tilde{T}_{i,0}\,; \qquad i = 1, 2, \cdots, c;$$

which is equivalent to $\tilde{T}_i \geqq$ constant.

Again following Puri [8], it can be easily shown that under the normal regularity conditions, the joint limiting distribution of the random variables

$(2m)^{\frac{1}{2}}[\tilde{T}_1 - \mu_1, \cdots, \tilde{T}_c - \mu_c]/A$ is the distribution of a $c$-dimensional normal vector $(U_1, U_2, \cdots, U_c)$, satisfying (6A.7) where $A$ is given by (4B.7) and

$$\mu_i = \int J[H_{i,0}(x)]\, dF(x - \theta_i)$$

where

$$H_{i,0}(x) = \tfrac{1}{2}[F(x - \theta_0) + F(x - \theta_i)]; \qquad J = F_0^{-1}.$$

Using this, one obtains results parallel to those of the previous section with regard to the asymptotic performance of this procedure. Omitting the details however, it is sufficient to state that the present procedure and the one considered in the previous section are asymptotically equi-efficient.

**Acknowledgment.** The authors would like to express sincere thanks to Professor Milton Sobel for making some valuable suggestions, and to the referee for pointing out an error in the original version of the paper.

## REFERENCES

[1] BECHHOFER, R. E. (1954). A single sample multiple decision procedure for ranking means of normal populations with known variances. *Ann. Math. Statist.* **25** 273–289.

[2] BECHHOFER, R. E. (1958). A sequential multiple-decision procedure for selecting the best one of several normal populations with a common unknown variance, and its use with various experimental designs. *Biometrics* **14** 408–429.

[3] GUPTA, SHANTI S. and MILTON SOBEL (1958). On selecting a subset which contains all populations better than a standard. *Ann. Math. Statist.* **29** 235–244.

[4] GUPTA, SHANTI S. (1956). On a decision rule for a problem in ranking means. Mimeograph series No. 150, Institute of Statistics, Univ. of North Carolina, Chapel Hill.

[5] LEHMANN, E. L. (1961). Some model I problems of selection. *Ann. Math. Statist.* **32** 990–1012.

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

[7] PAULSON, EDWARD (1952). On the comparison of several experimental categories with a control. *Ann. Math. Statist.* **23** 239–246.

[8] PURI, M. L. (1964). Asymptotic efficiency of a class of $c$-sample tests. *Ann. Math. Statist.* **35** 102–121.