BULLETIN OF THE
INSTITUTE OF MATHEMATICS
ACADEMIA SINICA
Volume 13, Number 3, September 1985

NONPARAMETRIC RANDOMIZED SELECTION PROCEDURES FOR LOCATION PARAMETERS*

BY

WEN-TAO HUANG (黃文濤) AND
A. K. MD. EHSANES SALEH

Abstract. We consider k populations such that the sample from the ith population follows the distribution law $F(x-\theta_i-\beta_i t_{ij})$ where F is in some class $\mathcal{G}$, $\{t_{ij}\}\in T$, θ_i and β_i are unknown parameters. By sampling from each population (unequal sample sizes), we consider selection problem by both subset selection approach and indifference zone formulation. We propose randomized and non-randomized selection procedures for selecting the population(s) associated with the largest location parameter θ_i. For infifference zone formulation, some comparisons among proposed rules and some efficiency of rules are studied. For subset selection approach, we use Monte Carlo method to study the behaviors of some proposed rules for small sample sizes. The superiority of randomized rules is strongly supported.

1. **Introduction.** Let $\pi_1, \pi_2, \cdots, \pi_k$ denote k populations such that the jth the random sample $Y_{ij}=\theta_i+\beta_i t_{ij}+\varepsilon_{ij}$ where θ_i and β_i are unknown parameters and the error term ε_{ij} is iid with cdf $F(\cdot)$, F unspecified, belonging to certain class $\mathcal{G}$ (defined in section 2). In some cases a nonlinear model can be transformed to become a linear model, for instance, the well-known logistic model representing the stimulus-response relationship is in such case. When Y_{ij} belongs to such model, it follows then the distribution law $F(y-\theta_i-\beta_i t_{ij})$.

In a set of regression equations $Y_{ij}=\theta_i+\beta_i t_{ij}+\varepsilon_{ij}$, $i=1,2,\cdots,k$, $j=1,2,\cdots,n_i$, we consider ε_{ij} are iid with cdf $F\in\mathcal{G}$. Such general simultaneous regression models are important in econometrics. In some situations wnen the hypothesis of

Received by the editors January 4, 1984 and in revised form November 26, 1984.
* Research supported by NSERC grant N. A3088.
Key words: Nonparametric selection rules, randomized selection rules, selection of the largest location, preliminary test for parallelism of slopes, efficiency.

homogeneity $H_0 : \theta_1 = \cdots = \theta_k$ is rejected, it is often desirable to test $k!$ hypotheses $\{H_l : = \theta_{i_1} \leq \theta_{i_2} \leq \cdots \leq \theta_{i_k}$, $(i_1, i_2, \cdots, i_k)$ some permutation of $(1, 2, \cdots, k)\}$ with at least one inequality. However, instead, in some occasions, we are more desirable to select which θ_i has the largest (smallest) value. This situation occurs, for example, when we are interested in selecting "superior" drugs among k new drugs in which θ_i is associated with initial response of ith drug in certain measurement of chemical or physical reactions for certain sympton. This leads to the problem of selection of components of location parameter vector. For our convenience, let $\underset{\sim}{t}_{in} = (t_{i1}, t_{i2}, \cdots, t_{in})'$, $\underset{\sim}{\theta} = (\theta_1, \theta_2, \cdots, \theta_k)'$ and $\underset{\sim}{\beta} = (\beta_1, \beta_2, \cdots, \beta_k)'$.

In many occasions, a priori information that $H_0 : \underset{\sim}{\beta} = \underset{\sim}{\beta} 1_k$ is useful when H_0 is true. When the data support the hypotheses H_0 it is more desirable to consider some statistics depending on H_0 for our seletcion procedure and when H_0 is not tenable, we consider other statistics which are distribution-free from $\underset{\sim}{\beta}$. Usually, when treating $\underset{\sim}{\beta}$ as a nuisance parameter, the statistics for our selection procedures may not depend on $\underset{\sim}{\beta}$, but, the elimination of $\underset{\sim}{\beta}$, generally entails some loss of information. This loss may be recovered (to a certain extent) when $\underset{\sim}{\beta}$ is not specified but partial information is given on $\underset{\sim}{\beta}$, i.e. when we are restricted to a proper subset of original parameter space. In such a case, it is naturally appealing to have a preliminary decision on $\underset{\sim}{\beta}$. In selection and ranhing problems, a preliminary decision on nuisance parameter has not been considered so far in nonparametric case. The object of the present paper is to study the behaviors of some selection procedures for $\underset{\sim}{\theta}$ based on a preliminary tests on $\underset{\sim}{\beta}$ and some comparisons to other selection procedures without such tests.

The effect of tests following such preliminary test of significanc has been studied for some special cases in Bechhofer (1951), Bozvich, Bancroft and Hartly (1956) and others. Recently, Saleh and Sen (1978) and Sen and Saleh (1979) consider nonparametric case.

In this paper, we consider the selection of largest θ_i using subset selection approach (see Gupta and Panchapakesan (1979)), Four procedures are proposed and studied in section 3. In section 4, we consider indifference zone approach. Four procedures are studied and some comparisons are made. And in section 5, we make a Monte Carlo study for cases $k=2$ and $k=5$ for small sample size with normal model.

2. **Some definitions and preliminary tests.** Let $\mathcal{G}$ be the class of all symmetric absolutely continuous distribution functions with absolutely continuous pdf f and let $\mathcal{G}$ denote the class of sequences $\{t_{ij}\}$ satisfying following conditions:

$$(2.1) \qquad I(f) \equiv \int_{-\infty}^{\infty} \{f'(x)/f(x)\}^2 f(x)\, dx < \infty$$

Let

$$(2.2) \qquad \bar{t}_{in} = \sum_{j=1}^{n_i} t_{ij}/n_i \qquad u_{in} = \sum_{1}^{n_i} (t_{ij} - \bar{t}_{in})^2/n_i$$

$n = \sum_{i=1}^{k} n_i$. As $n_i \uparrow \infty$, $n_i/n \rightarrow \lambda_i (0 < \lambda_i < 1)$, $\bar{t}_{in} \rightarrow t_i$ and $u_{in} \rightarrow u_i$ with $|t_i| < \infty$ and $u_i < \infty$, $i = 1, 2, \cdots, k$. We consider our problem for model $F(x - \theta_i - \beta_i t_{ij})$ with $F \in \mathcal{G}$ and $\{t_{ij}\} \in \mathcal{G}$. We take n_i samples from π_i. For our convenience, we all use n as a symbol for second subscript to denote sample size.

Let $\phi(u)$ $(0 < u < 1)$ be a nondecreasing, skew symmetric $(\phi(u) + \phi(1-u) = 0,\ 0 < u < 1)$ and sequare integrable score function, $\phi^*(u) = \phi((1+u)/2)$ $(0 < u < 1)$. Define

$$(2.3) \qquad \psi_n(i) = \phi(1/(n+1))$$

$$(2.4) \qquad \psi_n^*(i) = \phi^*(i/(n+1)).$$

For a random sample $\underset{\sim}{Y}_{in} = (Y_{i1}, Y_{i2}, \cdots, Y_{in_i})'$ from π_i and for interval (a, b), define

$$(2.5) \qquad \underset{\sim}{Y}_{in}(a, b) = \underset{\sim}{Y}_{in} - a\underset{\sim}{1}_{j_i} - b\underset{\sim}{t}_{in} \text{ where}$$

$$\underset{\sim}{1}_{n_i} = (1, 1, \cdots, 1), \text{ and } \underset{\sim}{t}_{in} = (t_{i1}, t_{i2}, \cdots, t_{in_i})'.$$

$$(2.6) \qquad T_{in}(a, b) = \sum_{j=1}^{n_i} \text{sqn}(Y_{ij} - a - bt_{ij})\, \psi_{in}^*(R_{ij}^+(a, b))$$

$$(2.7) \qquad L_{in}(a, b) = \sum_{n=1}^{n_i} (t_{ij} - \bar{t}_{in})\, \psi_{in}(R_{ij}(a, b))$$

$$(2.8) \qquad L_n(b_1, b_2, \cdots, b_k) = \sum_{i=1}^{k} L_{in}(a, b_i)$$

where $R_{ij}(a, b)$ ($R^+_{ij}(a, b)$) is the rank of $Y_{ij}-a-bt_{ij}$ ($|Y_{ij}-a-bt_{ij}|$) among

$$Y_{i1}-a-bt_{i1},\ Y_{i2}-a-bt_{i2},\cdots,\ Y_{in_i}-a-bt_{in_i}$$
$$(|Y_{i1}-a-bt_{i1}|,|Y_{i2}-a-bt_{i2}|,\cdots,|Y_{in_i}-a-bt_{in_i}|)$$

for $i=1, 2,\cdots, k$. Note that $R_{ij}(a, b)$ is independent of a and hence we denote $L_{in}(b)$ instead of $L_{in}(a, b)$ when no confusion occurs. We use R_{ij}, T_{in}, L_{in} and L_n to denote $R_{ij}(0, 0)$, $T_{in}(0, 0)$, $L_{in}(0\ ,0)$ and $L_n(\underset{\sim}{0})$ respectively. We define the following estimators for our parameters β_i and θ_i, $i=1, 2,\cdots k$.

$$\hat{\beta}_{in}=\frac{1}{2}(\sup\{b : L_{in}(b)>\}+\inf\{b : L_{in}(b)<0\} \tag{2.9}$$

$$\tilde{\beta}_n=\frac{1}{2}(\sup\{b : L_n(b\underset{\sim}{1})>0\}+\inf\{b : L_n(b\underset{\sim}{1})<0\}) \tag{2.10}$$

$$\hat{\theta}_{in}=\frac{1}{2}(\sup\{a : T_{in}(a, \tilde{\beta}_n)>0\} + \inf\{a : T_{in}(a, \tilde{\beta}_n)<0\}) \tag{2.11}$$

$$\tilde{\theta}_{in}=\frac{1}{2}(\sup\{a : T_{in}(a, \hat{\beta}_{in})>0\} + \inf\{a : T_{in}(a, \hat{\beta}_{in})<0\}). \tag{2.12}$$

It is known that the estimator $\tilde{\theta}_n=(\tilde{\theta}_{1n}, \tilde{\theta}_{2n},\cdots, \tilde{\theta}_{kn})'$ for $\underset{\sim}{\theta}$ is translation invariant, robust and consistent (Adichie (1967), Saleh and Sen (1978)). Also, when $H_0 : \underset{\sim}{\beta}=\beta\underset{\sim}{1}_k$ is true, $\hat{\theta}_n=(\hat{\theta}_{1n}, \hat{\theta}_{2n},\cdots, \hat{\theta}_{kn})'$ is translation invariant, robust and consistent for $\underset{\sim}{\theta}$. For the least square estimate of $\underset{\sim}{\theta}$, we consider the following estimators: Let $\bar{Y}_{in}$ denote the usual sample mean from π_i of size n_i.

$$\hat{\hat{\beta}}_{in}=\sum_{j=1}^{n_i}(t_{ij}-\bar{t}_{in})(Y_{ij}-\bar{Y}_{in})\Big/\sum_{j=1}^{n_i}(t_{ij}-\bar{t}_{in})^2 \tag{2.13}$$

$$\tilde{\tilde{\beta}}_n=\sum_{i=1}^{k}\left(\sum_{j=1}^{n_i}(t_{ij}-\bar{t}_{in})^2\hat{\hat{\beta}}_{in}\right)\Big/\sum_{i=1}^{k}\sum_{j=1}^{n_i}(t_{ij}-\bar{t}_{in})^2 \tag{2.14}$$

$$\hat{\hat{\theta}}_{in}=\bar{Y}_{in}-\tilde{\tilde{\beta}}_n\bar{t}_{in} \tag{2.15}$$

$$\tilde{\tilde{\theta}}_{in}=\bar{Y}_{in}-\hat{\hat{\beta}}_{in}\bar{t}_{in}. \tag{2-16}$$

When H_0 is true, the common value of β is usually estimated by $\tilde{\tilde{\beta}}_n$ which is the weighted sum of each $\hat{\hat{\beta}}_{in}$. Thus, $\hat{\hat{\theta}}_{in}$ and $\tilde{\tilde{\theta}}_{in}$ are, respectively, codsidered as the estimate of θ_i when H_0 is true or not. We consider our selection procedures based on the rank estimates $\hat{\underset{\sim}{\theta}}_n$, $\tilde{\underset{\sim}{\theta}}_n$ and least square ettimates $\hat{\hat{\underset{\sim}{\theta}}}_n$ and $\tilde{\tilde{\underset{\sim}{\theta}}}_n$.

For the preliminary test on the parallelism of regression, Sen (1969) considered a test based on L_n defined by

$$L_n = A_\phi^{-2} \sum_{i=1}^{k} \left\{ L_{in}^2(\tilde{\beta}_n) \Big/ \sum_{j=1}^{n_i} (t_{ij} - \bar{t}_{in})^2 \right\} \tag{2.17}$$

where

$$A_\phi^2 = \int_0^1 \phi^2(u)\,du - \left(\int_0^1 \phi(u)\,du \right)^2 \tag{2.18}$$

Define

$$\zeta(u) = -\left[f'(F^{-1}(u)) / f(F^{-1}(u)) \right] \tag{2.19}$$

$$\gamma(\zeta,\phi) = \int_0^1 \zeta(u)\,\phi(u)\,du \tag{2.20}$$

$$\tau = A_\phi / \gamma(\zeta, \phi)\,. \tag{2.21}$$

For the least square estimates, we consider a test based on

$$l_n = \sum_{i=1}^{k} \left\{ \sum_{j=1}^{n_i} (t_{ij} - \bar{t}_{in})^2 \right\} (\hat{\hat{\beta}}_{in} - \tilde{\tilde{\beta}}_n)^2 / S_e^2 \tag{2.22}$$

where

$$S_e^2 = \sum_{i=1}^{k} \sum_{j=1}^{n_i} (Y_{ij} - \tilde{\tilde{\theta}}_{in} - \hat{\hat{\beta}}_{in} t_{ij})^2 / (n - 2k)$$

The two non-randomized preliminary tests for homogeneity of $\underset{\sim}{\beta}$ are defined by T_1 and T_2 as follows

$$T_1(\underset{\sim}{Y}) = 0 \quad \text{if } L_n < L_{n,\alpha}; \qquad = 1 \quad \text{otherwise} \tag{2.23}$$

$$T_2(\underset{\sim}{Y}) = 0 \quad \text{if } l_n < l_{n,\alpha}; \qquad = 1 \quad \text{otherwise} \tag{2.24}$$

where $L_{n,\alpha}$ and $l_{n,\alpha}$ are some constants. It is known (Adichie (1967), Sen (1969)) that as $n \to \infty$, $(L_n | H_0) \xrightarrow{\mathscr{L}} \chi_{k-1}^2$, $(l_n | H_0) \xrightarrow{\mathscr{L}} \chi_{k-1}^2$, the chi-square random variable with d.f. $k-1$. And also,

$L_{n,\alpha} \to \chi^2_{k-1,\alpha}$ $l_{n,\alpha} \to \chi^2_{k-1,\alpha}$, the $\alpha\%$ point of χ^2_{k-1}.

3. **Some subset selection procedures.** Based on a sample $\underset{\sim}{Y}$ drawn from the k populations (n_i sample from π_i), we select a subset (non-empty) of the k populations. It is required that the probability of correct selection (i.e. the population associated with the largest θ_i is included in the subset selected) is at least $P^*(1/k < P^* < 1)$ a preassigned value (P^*-condition). We propose the following non-randomized and randomized selection rules. For some given constants (to be chosen depending on P^*) $d_i(i = 1, 2, 3, 4, 5, 6)$ we define

R_1: Select π_i, if and only if

$$\tilde{\theta}_{in} \geq \max_{1\leq j\leq k} \tilde{\theta}_{jn} - \frac{d_1}{\sqrt{n}} \tag{3.1}$$

$\tilde{R}_1$: Select π_i, if and only if

$$\tilde{\tilde{\theta}}_{in} \geq \max_{1\leq j\leq k} \tilde{\tilde{\theta}}_{jn} - \frac{d_2}{\sqrt{n}} \tag{3.2}$$

R_2: If $T_1(\underset{\sim}{Y}) = 0$, select π_i, if and only if

$$\hat{\theta}_{in} \geq \max_{j} \hat{\theta}_{jn} - \frac{d_3}{\sqrt{n}} \tag{3.3}$$

and if $T_1(\underset{\sim}{Y}) = 1$, apply R_1 with constant d_4

$\tilde{R}_2$: $T_2(\underset{\sim}{Y}) = 0$, select π_i, if and only if

$$\hat{\hat{\theta}}_{in} \geq \max_{j} \hat{\hat{\theta}}_{jn} - \frac{d_5}{\sqrt{n}} \tag{3.4}$$

and if $T_2(\underset{\sim}{Y}) = 1$, apply $\tilde{R}_1$ with constant d_6.

Note that $L_{n,\alpha} \geq 0$ and $l_{n,\alpha} \geq 0$. Hence by choosing $L_{n,\alpha} = 0$ and $l_{n,\alpha} = 0$, R_2 and $\tilde{R}_2$ become, respectively, R_1 and $\tilde{R}_1$..

A correct selection (CS) is an event that at least one of the populations associated with the largest θ_i is included in the subset selected. Let $\theta_{[1]} \leq \theta_{[2]} \leq \cdots \leq \theta_{[k]}$ denote the ordered components of $\underset{\sim}{\theta}$. We use $\pi_{(i)}$ to denote some π_j which is associated with $\theta_{[i]}$. Also, we denote $\tilde{\theta}_{(i)n}$ ($\hat{\theta}_{(i)n}$ etc.) as the statistic of sample from $\pi_{(i)}$. Let $\underset{\sim}{n} = (n_1, n_2, \cdots, n_k)'$, $\underset{\sim}{t} = (t_1, t_2, \cdots, t_k)'$. Define the following parameter spaces.

(3.5) $$\Omega = \{(\underset{\sim}{\theta}', \underset{\sim}{\beta}')'\}$$

(3.6) $$\Omega_{01} = \{(\underset{\sim}{\theta}', \underset{\sim}{\beta}')' : \underset{\sim}{\theta} = \theta \underset{\sim}{1}_k,\ \underset{\sim}{\beta} \neq \beta \underset{\sim}{1}_k\}$$

(3.7) $$\Omega_{10} = \{(\underset{\sim}{\theta}', \underset{\sim}{\beta}')' : \underset{\sim}{\theta} \neq \theta \underset{\sim}{1}_k,\ \underset{\sim}{\beta} = \beta \underset{\sim}{1}_k\}$$

(3.8) $$\Omega_{00} = \{(\underset{\sim}{\theta}', \underset{\sim}{\beta}')' : \underset{\sim}{\theta} = \theta \underset{\sim}{1}_k,\ \underset{\sim}{\beta} = \beta \underset{\sim}{1}_k\}.$$

We have then the following

THEOREM 3.1. (i) $\lim_{n\to\infty} \min_{(\underset{\sim}{n},\{t_{ij}\})} \inf_{\Omega} P(\mathrm{CS}|R_1)$

(3.9) $$= \int_{-\infty}^{\infty} \sum_{i=2}^{k} \Phi\left(\frac{a_i u + d_1/v_i}{\sqrt{1-a_i^2}}\right) d\Phi(u)$$
$$= I(k-1, d_1;\ \underset{\sim}{a}, \underset{\sim}{v})$$

where $\underset{\sim}{a} = (a_1, a_2, \cdots, a_k)'$, $\underset{\sim}{v} = (v_1, v_2, \cdots, v_k)'$ *are defined by* (3.15) *and* (3.16).

(ii) $\lim_{n\to\infty} \min_{(n,\{t_{ij}\})} \inf_{\Omega} P(\mathrm{CS}|\tilde{R}_1) = I(k-1,\ d_2;\ \underset{\sim}{a},\ \underset{\sim}{v}_1)$

where $\underset{\sim}{v}_1 = (\sigma^2/\tau^2)\,\underset{\sim}{v}$, σ^2 *is the variance of* F, τ *defined by* (2.21).

Proof. (i) Since $\tilde{\theta}_{in}$ is translation invariant (Adichie (1967)) and it can be shown that $\tilde{\theta}_{in}$ is stochastically non-decreasing in θ_i, it follows (see Barr and Rizvi (1966)) that $P_\omega(\mathrm{CS}|R_1)$ attains its infimum on Ω_{01} (defined by (3.6)) for any $\omega \in \Omega$. Since $\sqrt{n}(\tilde{\theta}_{(i)n} - \theta_{[i]})$ is asymptotically normal $n(0, w_{(i)})$ (defined by (3.13)) and since

$$P(\mathrm{CS}|R_1) = P\{\sqrt{n}(\tilde{\theta}_{(k)n} - \theta_{[k]}) \geq (\max_j \sqrt{n}(\theta_{(j)n} - \theta_{[j]})) - d_1 - \sqrt{n}(\theta_{[k]} - \theta_{[j]})\},$$

we have then

(3.10) $$\lim_{n\to\infty} \inf_{\Omega} P(\mathrm{CS}|R_1) = P\{U_j \leq d_1,\ j = 1, 2, \cdots, k-1\}$$

where $\underset{\sim}{U} = (U_1, U_2, \cdots, U_{k-1})'$ is multivariate normal $n(\underset{\sim}{0}, \{\sigma_i \sigma_j \rho_{ij}\})$ where

(3.11) $$\rho_{ij} = \mathrm{Cor}(U_i, U_j) = \alpha_i \alpha_j$$

(3.12) $$\alpha_i = \left(1 + \frac{w_{(i)}}{w_{(k)}}\right)^{-1/2} \quad (w_{(i)} \text{ is associated with } \theta_{[i]}).$$

(3.13) $$w_i = \tau^2 \left(\frac{1}{\lambda_i} + \frac{t_i^2}{\lambda_i u_i}\right)$$

(3.14) $$\sigma_i^2 = w_{(i)} + w_{(k)}.$$

Now. by the Slepian's inequality (Slepian (1962) or see Gupta (1963)) that the RHS of (3.10) is non-decreasing with respect to each ρ_{ij}. Hence, $\liminf_{\Omega_{01}} P(\text{CS}|R_1)$ attains its minimum with respect to $\{\underset{\sim}{t}, \underset{\sim}{\lambda}\}$ when each ρ_{ij} is minimized. This can be attained when $w_{(k)} = \min_i w_i$. Let $w_{[1]} \leq w_{[2]} \leq \cdots \leq w_{[k]}$ denote the ordered values of $w_i's$. Let

(3.15) $$a_i = \left(1 + \frac{w_{[i]}}{w_{[1]}}\right)^{-1/2} \qquad i = 2, 3, \cdots, k$$

(3.16) $$v_i = (w_{[1]} + w_{[i]})^{1/2} \qquad i = 2, 3, \cdots, k.$$

Define

(3.17) $$X_i = v_i[(1-a_i^2)^{1/2} V_i - a_i V] \qquad i = 2, 3, \cdots, k$$

where $V, V_2, V_3, \cdots, V_k$ are iid standard normal. Then

$$\lim_{n\to\infty} \min_{(\underset{\sim}{n}, \{t_{ij}\})} \inf_{\Omega} P(\text{CS}|R_1) = P\{X_i \leq d_1, i = 2, 3, \cdots, k\}$$
$$\equiv I(k-1, d_1;\ \underset{\sim}{a}, \underset{\sim}{v}).$$

(ii) Follow analogous arguments in (i) and note that $\sqrt{n}(\tilde{\tilde{\theta}}_{ij} - \theta_i)$ is asymptotically normal with meah 0, variance $(\sigma^2/\tau^2) w_i$ (defined by (3.13)). The proof is thus complete.

THEOREM 3.2. *Taking $t_1 = t_2 = \cdots = t_k = t$, we have*

(i) $$\liminf_{n\to\infty} \inf_{\Omega} P(\text{CS}|R_2)$$
$$\geq \min\{I(k-1, d;\ \underset{\sim}{a}', \underset{\sim}{v}'), I(k-1, d;\ \underset{\sim}{a}, \underset{\sim}{v})\}$$

where $\underset{\sim}{a}'$ and $\underset{\sim}{v}'$ are defined by (3.18) *and* (3.19).

(ii) $$\liminf_{n\to\infty} \inf_{\Omega} P(\text{CS}|\tilde{R}_2)$$
$$\geq \min\{I(k-1, d;\ \underset{\sim}{a}', \underset{\sim}{v}_1'), I(k-1, d';\ \underset{\sim}{a}, \underset{\sim}{v}_1)\}$$

where $\underset{\sim}{v}_1' = (\sigma/\tau)^2 \underset{\sim}{v}'$, $\underset{\sim}{v}_1 = (\sigma/\tau)^2 \underset{\sim}{v}$, σ^2 is the variance of F.

Proof. (i) Note that $\hat{\theta}_n$ and $\tilde{\theta}_n$ are both translation invariant and stochastically non-decreasing with respect to $\underset{\sim}{\theta} = (\theta_1, \cdots, \theta_k)'$ componentwise. Hence, as $n \uparrow \infty$, $P(\text{CS}|R_2)$ attains its infimum on Ω_{01}. Now note that the test T_1 (defined by (2.23)) is consistent for H_0. Also $(L_n|H_0) \xrightarrow{\mathcal{L}} \chi^2_{k-1}$. Hence, for $\omega \in \Omega_{00}$,

$$\lim_{n\to\infty} P_\omega(\mathrm{CS}|R_2) = (1-\alpha)\lim_{n\to\infty} P_\omega\left(\hat\theta_{(k)n} \geq \max_j \hat\theta_{(j)n} - \frac{d_3}{\sqrt{n}}\right)$$
$$+\alpha\lim_{n\to\infty} P_\omega\left(\hat\theta_{(k)n} \geq \max_j \hat\theta_{(j)n} - \frac{d_4}{\sqrt{n}}\right)$$

where we take $L_{n,\alpha} = \chi^2_{k-1,\alpha}$ for the test T_1 in rule R_2.

For $\omega \in \Omega_{01}$, $\lim_{n\to\infty} P_\omega(\mathrm{CS}|R_2) = \lim_{n\to\infty} P_\omega(\mathrm{CS}|R_1)$.

Again $\sqrt{n}(\hat{\underset{\sim}{\theta}}_n - \underset{\sim}{\theta})$ is asymptotically normal $n(\underset{\sim}{0}, \Sigma_1)$, with $\sigma_{ij} = t_i t_j/\bar u$ $(i \neq j)$, $\bar u = \sum_1^k \lambda_i u_i$, $\sigma_{ii} = 1/\lambda_i + t_i/\bar u$ (Saleh and Sen (1978)). Define

$$U_i = \sqrt{n}\,[(\hat\theta_{(i)n} - \theta_{[i]}) - (\hat\theta_{(k)n} - \theta_{[k]})], \quad i = 1, 2, \cdots, k-1.$$

Then, for $t_1 = t_2 = \cdots = t_k = t$, $U = (U_1, U_2, \cdots, U_{k-1})'$ is $n(\underset{\sim}{0}, \Sigma_2)$ with $\sigma'_{ij} = (1/\lambda_{(k)})(i \neq j)$, and $\sigma'_{ii} = 1/\lambda_{(i)} + 1/\lambda_{(k)}$. Hence,

$$\omega \in \Omega_{01}, \quad \lim_{n\to\infty} P_\omega(\mathrm{CS}|R_2) = P\{U_i < d_3,\ i = 1, 2, \cdots, k-1\}$$

which attains its minimum when $\rho_{ij} \equiv \rho(U_i, U_j) = \alpha'_i \alpha'_j$ is minimized where $\alpha'_i = (1 + \lambda_{(k)}/\lambda_{(i)})^{-1/2}$ (by Slepian's inequality). let $\lambda_{[1]} \leq \lambda_{[2]} \leq \cdots \leq \lambda_{[k]}$ denote the ordered value of λ_i. Then ρ_{ij} is minimized when

$$(3.18) \qquad \alpha'_i = a'_{i+1} = \left(1 + \frac{\lambda_{[k]}}{\lambda_{[i]}}\right)^{-1/2}, \qquad i = 1, 2, \cdots, k-1.$$

Let

$$(3.19) \qquad v'_{i+1} = \tau^2\left(\frac{1}{\lambda_{[i]}} + \frac{1}{\lambda_{[k]}}\right), \qquad i = 1, 2, \cdots, k-1.$$

Follow analogous arguments given in (i) of Theorem 3.1, we concluds that for

$$\omega \in \Omega_{00}, \lim_{n\to\infty} P_\omega(\mathrm{CS}|R_2) = I(k-1, d_3;\ a', v').$$

On the other hand, we have, for

$$\omega \in \Omega_{01}, \quad \lim_{n\to\infty} P_\omega(\mathrm{CS}|R_2) = \lim_{n\to\infty} P_\omega(\mathrm{CS}|R_1)$$
$$\geq I(k-1, d_4;\ \underset{\sim}{a}, \underset{\sim}{v})$$

by (i) of Theorem 3.1. If

$$I(k-1, d;\ \underset{\sim}{a}, \underset{\sim}{v}) \geq I(k-1, d;\ \underset{\sim}{a}', \underset{\sim}{v}'),$$

then

$$(1-\alpha)I(k-1,\ d;\ \underset{\sim}{a},\ \underset{\sim}{v})+\alpha I(k-1,\ d;\ \underset{\sim}{a}',\ \underset{\sim}{v}')$$
$$\geq I(k-1,\ d;\ \underset{\sim}{a}',\ \underset{\sim}{v}').$$

Hence, we conclude

$$\liminf_{n\to\infty}\inf_{\Omega} P(\mathrm{CS}\,|\,R_2)$$
$$\geq \min\,(I(k-1,\ d;\ \underset{\sim}{a}',\ \underset{\sim}{v}'),\ I(k-1,\ d;\ \underset{\sim}{a},\ \underset{\sim}{v}))$$

This still holds if

$$I(k-1,\ d;\ a,\ v)<I(k-1,\ d;\ a',\ v').$$

For the proof of (ii), the arguments are analogous.

4. **Selection rules with indifference zone approach.** In this section we consider selection of one best (extension to t-best is analogous) with an indifference zone in parameter space. For given $\alpha>0$, we consider π_i is good if $\theta_i \geq \theta_{[k]}-\alpha$. Our parameter space now is defined to be $\Omega(\alpha)=\{(\underset{\sim}{\theta}',\ \underset{\sim}{\beta}'):\theta_{[k]}-\theta_{[k-1]}\geq\alpha\}$ the preference zone. Hence π_i is called the best if it is associated with $\theta_{[k]}$.. Consider a sequence of situations for increasing n and define π_i to be good if $\theta_i\geq\theta_{[k]}-\alpha^{(n)}$ where $\alpha^{(n)}$ depends on the sample size. This sequence is a device for approximating the actual situation since we are considering large sample case without completely specifying F.

We require that the probability of correct selection (CS) is not less than a preassigned value $P^*(1/k<P^*<1)$. In the preference zone, we need to determine sample size n so that the P^*-condition can be fulfilled. Our selection procedures are based on the maximum components of $\underset{\sim}{\tilde{\theta}}_n$ and $\underset{\sim}{\hat{\theta}}_n$ respectively. To be precise, we define selection rules R_{Ii} and $\tilde{R}_{Ii}$ as follows: R_{Ii} is the procedure R_i defined in section 3 by taking all d-values equal 0, $i=1,\,2$. And also, $\tilde{R}_{Ii}$ is the procedur $\tilde{R}_i$ by taking all d-values equal 0, $i=1,\,2$.

Hence one population is selected for each selection. Our criterion for selection is to satisy the P^*-condition i.e.

$$P_\omega(\text{selected population is good})\geq P^*\ \ \forall\ \omega\in\Omega(\alpha) \tag{4.1}$$

where $P^*(1/k<P^*<1)$ is preassigned and $\Omega(\alpha)$ is defined in

(4.2). We define following notation for convenience.

$$\Omega(\alpha) = \{(\underset{\sim}{\theta}', \underset{\sim}{\beta}') : \theta_{[k]} \geq \theta_{[k-1]} + \alpha\} \tag{4.2}$$

$$\Omega_{01}(\alpha) = \{(\underset{\sim}{\theta}', \underset{\sim}{\beta}') : \theta_{[1]} = \theta_{[k-1]} = \theta_{[k]} - \alpha\} \tag{4.3}$$

$$\Omega_{10}(\alpha) = \{(\underset{\sim}{\theta}', \underset{\sim}{\beta}') : \underset{\sim}{\beta} = \beta \underset{\sim}{1}_k\} \cap \Omega(\alpha) \tag{4.4}$$

$$\Omega_{00}(\alpha) = \Omega_{10}(\alpha) \cap \Omega_{01}(\alpha). \tag{4.5}$$

The sample size problems for large sample solution are given as follows.

THEOREM 4.1. (i) *As* $n \to \infty$, *if* $\alpha^{(n)} = \Delta_1/\sqrt{n} + O(1/\sqrt{n})$ *where* Δ_1 *satisfies* $I(k-1, \Delta_1; \underset{\sim}{a}, \underset{\sim}{v}) = P^*$ (*defined by* (3.9)), *then* $\lim_{n\to\infty} \inf_{\Omega(\alpha^{(n)})} P(CS|R_{I1}) = P^*$.

(ii) *As* $n \to \infty$, *if* $\alpha^{(n)} = \Delta_2/\sqrt{n} + O(1/\sqrt{n})$ *where* Δ_2 *satisfies* $I(k-1, \Delta_2; \underset{\sim}{a}, \underset{\sim}{v}_1) = P^*$, *then* $\lim_{n\to\infty} \inf_{\Omega(\alpha^{(n)})} P(CS|\tilde{R}_{I1}) = P^*$.

Proof. Note that $\tilde{\theta}_{in}$ is translation invariant and stochastically non-decreasing with respect to θ_i. Hence, by Barr and Rizvi (1966), $P(CS|R_{I1})$ attains its infinmum on $\Omega_{01}(\alpha^{(n)})$. Also, note that for

$$\omega \in \Omega_{01}(\alpha^{(n)}), \quad P_\omega(|R_{I1}) = P_r\{\sqrt{n}[(\tilde{\theta}_{(i)n} - \theta_{[i]}) - (\tilde{\theta}_{(k)n} - \theta_{[k]})] \leq \sqrt{n}\,\alpha^{(n)}, \quad i = 1, 2, \cdots, k-1\}.$$

Then, follow the analogous arguments in Theorem 3.1, and take $\sqrt{n}\,\alpha^{(n)} = \Delta_1$, we can conclude (i). Follow same arguments and by (ii) of Theorem 3.1, we conclude that (ii) holds.

Note that for very special case that $\underset{\sim}{\beta} = \underset{\sim}{0}$ and $\lambda_1 = \lambda_2 = \cdots = \lambda_k$, (ii) becomes Lemma 1 of Lehmann (1963).

THEOREM 4.2. *If* $t_1 = t_2 = \cdots = t_k = t$ *and* $\alpha^{(n)} = \Delta_3/\sqrt{n} + O(1/\sqrt{n})$, *then*

(i) $$\lim_{n\to\infty} \inf_{\Omega(\alpha^{(n)})} P(CS|R_{I2}) \geq \min\{I(k-1, \Delta_3; \underset{\sim}{a}', \underset{\sim}{v}'), I(k-1, \Delta_3; \underset{\sim}{a}, \underset{\sim}{v})\}$$

where $I(\cdot)$ *is defined by* (3.9), $\underset{\sim}{a}'$, $\underset{\sim}{v}'$ *defined by* (3.18), (3.19), $\underset{\sim}{a}$, $\underset{\sim}{v}$ *defined by* (3.13) *and* (3.14).

(ii) $$\lim_{n\to\infty} \inf_{\Omega(\alpha^{(n)})} P(CS|R_{I2}) \geq \{\min I(k-1, \Delta_3; \underset{\sim}{a}', \underset{\sim}{v}_1'), I(k-1, \Delta_3; \underset{\sim}{a}, \underset{\sim}{v}_1)\}$$

where $\underset{\sim}{v}_1' = (\sigma/\tau)^2 \underset{\sim}{v}'$, $\underset{\sim}{v}_1 = (\sigma/\tau)^2 \underset{\sim}{v}$, σ^2 *is the variance of F.*

It follows from Theorem 4.2 that $\Delta_3 = \max(\Delta_3(\underset{\sim}{a}', \underset{\sim}{v}'), \Delta_3(\underset{\sim}{a}, \underset{\sim}{v}))$ can be de determined so that the P^*-condition is satisfied.

Next, for given vectors $\underset{\sim}{\varepsilon}' = (\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_{k-1})$ $(\varepsilon_1 \geq \varepsilon_2 \geq \cdots \geq \varepsilon_{k-1})$ and $\underset{\sim}{\delta}' = (\delta_1, \delta_2, \cdots, \delta_k)$ $(\varepsilon_i \geq 0, \delta_i \geq 0)$, we consider a sequence of parameter vector

$$\Omega^{(n)}(\underset{\sim}{\varepsilon}, \underset{\sim}{\delta}) = \left\{(\underset{\sim}{\theta}', \underset{\sim}{\beta}') : \theta_{[k]} - \theta_{[j]} = \frac{\varepsilon_j}{\sqrt{n}},\ \beta_i = \beta + \frac{\delta_i}{\sqrt{n}}\right\} \tag{4.6}$$

We note that the asymptotic infima over $\Omega^{(n)}(\underset{\sim}{\varepsilon}, \underset{\sim}{\delta})$ of probability of correct selection applying R_{I1} and $\tilde{R}_{I1}$ are the same as the one over

$$\Omega^{(n)}(\underset{\sim}{\varepsilon}) = \left\{(\underset{\sim}{\theta}', \underset{\sim}{\beta}') : \theta_{[k]} - \theta_{[j]} = \frac{\varepsilon_j}{\sqrt{n}}\right\}$$

since $\tilde{\theta}_{in}$ and $\tilde{\tilde{\theta}}_{in}$ are both asymptotically independent of $\underset{\sim}{\beta}$.

In order to find the probability of correct selection asymptotically with respect to $\Omega^{(n)}(\underset{\sim}{\varepsilon}, \underset{\sim}{\delta})$ applying R_{I2}, we need to find the asymptotic joint distributions of $\{L_n, \sqrt{n}(\hat{\underset{\sim}{\theta}}_n - \underset{\sim}{\theta})\}$ and $\{L_n, \sqrt{n}(\tilde{\theta}_n - \underset{\sim}{\theta})\}$. By Sen (1969) and through some computations, we have that for $\omega_n \in \Omega^{(n)}(\underset{\sim}{\varepsilon}, \underset{\sim}{\delta})$, as $n \to \infty$,

$$\begin{aligned} L_n = \frac{1}{nA_\phi^2}\Bigg\{ & \sum_{i=1}^{k} \frac{1}{\lambda_i u_i}[L_{in}(\beta_i) + \sqrt{n}\,\delta_i \lambda_i u_i]^2 \\ & - \bar{u}^{-1}\left[\sum_{i=1}^{k}(L_{in}(\beta_i) + \sqrt{n}\,\delta_i \lambda_i u_i)\right]^2 + o_p(1)\Bigg\} \end{aligned} \tag{4.7}$$

where $\bar{u} = \sum_{i=1}^{k} \lambda_i u_i$. Again, it can be rewritten by

$$\begin{aligned} L_n = (n^{-1/2} A_\phi^{-1} & \underset{\sim}{L}(\underset{\sim}{\beta}) + \xi)'(D_{22} \\ & - \bar{u}^{-1} \underset{\sim}{1}_k \underset{\sim}{1}_k')(n^{-1/2} A_\phi^{-1} \underset{\sim}{L}(\underset{\sim}{\beta}) + \underset{\sim}{\xi}) \end{aligned} \tag{4.8}$$

where

$$\underset{\sim}{\xi} = (\xi_1, \xi_2, \cdots, \xi_k) \tag{4.9}$$

$$\xi_i = \gamma A_\phi^{-1} \delta_i \lambda_i u_i \tag{4.10}$$

$$\underset{\sim}{L}(\underset{\sim}{\beta}) = (L_{1n}(\beta_1), \cdots, L_{kn}(\beta_k))' \tag{4.11}$$

(4.12 $$D_{22} = \text{diag}\left(\frac{1}{\lambda_1 u_1}, \cdots, \frac{1}{\lambda_k u_k}\right).$$

Again, by Saleh and Sen (1978),

$$(4.13)\qquad (\sqrt{n}(\hat{\underset{\sim}{\theta}}_n - \underset{\sim}{\theta}),\ n^{-1/2}\underset{\sim}{L}(\beta)/A_\phi)$$

is asymptotically normal

$$n((\underset{\sim}{T}^*(\underset{\sim}{\delta} - \bar{\underset{\sim}{\delta}}),\ \tau^{-1}D_{22}^{-1}\underset{\sim}{\delta}),\ \Sigma_3)$$

where

$$(4.14)\qquad \underset{\sim}{T}^* = \mathrm{diag}\,(t_1, t_2, \cdots, t_k)$$

$$(4.15)\qquad \bar{\delta} = \sum_{i=1}^{k} \lambda_i u_i \delta_i/\bar{u}$$

$$(4.16)\qquad \Sigma_3 = \left(\begin{array}{c|c} \tau^2 D_{11} & D_{12}\tau \\ \hline D_{12}\tau & D_{22}^{-1} \end{array}\right)$$

$$(4.17)\qquad D_{11} = \mathrm{diag}\left(\frac{1}{\lambda_1}, \cdots, \frac{1}{\lambda_k}\right) + \bar{u}^{-1}\underset{\sim}{T}\,\underset{\sim}{T}'$$

$$(4.18)\qquad \underset{\sim}{T} = (t_1, t_2, \cdots, t_k)'$$

$$(4.19)\qquad D_{12} = -\bar{u}T^*\underset{\sim}{1}_k\underset{\sim}{1}_k' D_{22}^{-1}.$$

And also, we have

$$(4.20)\qquad (\sqrt{n}(\tilde{\underset{\sim}{\theta}}_n - \underset{\sim}{\theta}),\ n^{-1/2}\underset{\sim}{L}(\underset{\sim}{\beta})/A_\phi)$$

is asymptotically normal

$$n_{2k}((\underset{\sim}{0},\ \tau A_\Phi^{-1}D_{22}^{-1}\underset{\sim}{\delta}),\ \Sigma_4)$$

where

$$(4.21)\qquad \Sigma_4 = \left(\begin{array}{c|c} \tau^2 V^* & -T^*\tau \\ \hline -T^*\tau & D_{22}^{-1} \end{array}\right)$$

$$(4.22)\qquad V^* = \mathrm{diag}\left(\frac{1}{\lambda_1} + \frac{t_1^2}{\lambda_1 u_1}, \cdots, \frac{1}{\lambda_k} + \frac{t_k^2}{\lambda_k u_k}\right).$$

Hence, it can be shown that $\sqrt{n}(\hat{\underset{\sim}{\theta}}_n - \underset{\sim}{\theta})$ and $(D_{22} - \bar{u}^{-1}\underset{\sim}{1}_k\underset{\sim}{1}_k')\,n^{-1/2} \times \underset{\sim}{L}(\beta))$ are asymptotically independent since

$$(4.23)\qquad D_{12}(D_{22} - \bar{u}^{-1}\underset{\sim}{1}_k\underset{\sim}{1}_k') = \underset{\sim}{0}.$$

We know that (Sen (1969)) L_n is asymptotically chi-square on $\Omega^{(n)}(\underset{\sim}{\varepsilon}, \underset{\sim}{\delta})$ with non-centrality

$$(4.24)\qquad \Delta^* = \tau^{-2}\sum_{1}^{k} \lambda_i u_i(\delta_i - \bar{\delta})^2$$

We can conclude then that for $\omega_n \in \Omega^{(n)}(\underset{\sim}{\varepsilon}, \underset{\sim}{\delta})$

$$P_{\omega_n}(L_n < L_{n,\alpha},\ \hat{\theta}_{(k)n} = \max_j \hat{\theta}_{(j)n}) \tag{4.25}$$
$$\longrightarrow H_{k-1}(\chi^2_{k-1,\alpha};\ \Delta^*)\, P\{Z_i < \varepsilon_i - (\zeta_{(i)} + \zeta_{(k)}),\quad i = 1, 2, \cdots, k-1\}$$

where

$$\zeta_i = t_i(\delta_i - \bar{\delta}) \tag{4.26}$$

and $H_k(x;\ \Delta^*)$ is the cdf of non-central chi-square with non-centrality $\Delta^*((4.24))$ and $(Z_1, Z_2, \cdots, Z_{k-1})'$ is multivariate normal $n_{k-1}(\underset{\sim}{0},\ (\sigma_i\, \sigma_j\, \eta_i\, \eta_j))$

$$\sigma_i^2 = \left(\frac{1}{\lambda_{(i)}} + \frac{t^2_{(i)}}{\lambda_{(i)}\, u_{(i)}}\right)\tau^2 \tag{4.27}$$

$$\eta_i = \left(1 + \frac{\sigma_i}{\sigma_k}\right)^{-1/2} \tag{4.28}$$

Hence, we have for $\omega_n \in \Omega^{(n)}(\underset{\sim}{\varepsilon}, \underset{\sim}{\delta})$

$$P\{Z_i < \varepsilon_i - (\zeta_i + \zeta_{(k)}),\ i = 1, 2, \cdots, k-1\}$$
$$= \int_{-\infty}^{\infty} \sum_{i=1}^{k-1} \Phi\left(\frac{(\varepsilon_i - \zeta_{(i)} - \zeta_{(k)})/\sigma_i + \eta_i\, u}{\sqrt{1-\eta_i^2}}\right) d\Phi(u) \tag{4.29}$$

$$\equiv J(\underset{\sim}{\varepsilon},\ \underset{\sim}{\delta},\ \zeta,\ \underset{\sim}{\sigma},\ \eta) \tag{4.30}$$

where $\zeta_{(i)}$ is associated with $\pi_{(i)}$. Since $\zeta_{(i)}$ and $\zeta_{(k)}$ are unknown and also η_i is unknown. in order to minimize (4.30) with respect to all permutations of $(\zeta_1, \zeta_2, \cdots, \zeta_k)$ and $(\sigma_1^2, \sigma_2^2, \cdots, \sigma_k^2)$,it is sufficient to choose some j such that when $\sigma_k = 1/\lambda_j + t_j^2/\lambda_j u_j$ and $\zeta_{(k)} = t_j(\delta_j - \bar{\delta})$, (4.30) is minimized. Let

$$J_1(k-1, \underset{\sim}{\varepsilon},\ \underset{\sim}{\delta},\ \underset{\sim}{\lambda},\ \underset{\sim}{t}) \equiv \min_{(\underset{\sim}{\zeta}, \underset{\sim}{\sigma})} J(\underset{\sim}{\varepsilon},\ \underset{\sim}{\delta},\ \zeta,\ \underset{\sim}{\sigma},\ \eta). \tag{4.31}$$

On the other hand, taking $\underset{\sim}{W}_n = n^{-1/2} A_{\Phi}^{-1} \underset{\sim}{L}(\underset{\sim}{\beta})$, we have by (4.8) that

$$L_n = (\underset{\sim}{W}_n + \xi)'(D_{22} - \bar{u}^{-1} \underset{\sim}{1}_k \underset{\sim}{1}_k')(\underset{\sim}{W}_n + \xi). \tag{4.32}$$

Let

$$U_{in} = \sqrt{n}[(\tilde{\theta}_{(i)n} - \theta_i) - (\tilde{\theta}_{(k)n} - \theta_k)]. \tag{4.33}$$

Then, for $\omega_n \in \Omega^{(n)}(\underset{\sim}{\varepsilon}, \underset{\sim}{\delta})$

$$P_{\omega_n}\{L_n \geq L_{n,\alpha},\ \tilde{\theta}_{(k)n} = \max_j \theta_{(j)n}\}$$

(4.34)
$$\longrightarrow P\{U_i < \varepsilon_i,\ (\underset{\sim}{W} + \underset{\sim}{\xi})' B (\underset{\sim}{W} + \underset{\sim}{\xi}) > \chi^2_{k-1,\alpha}$$
$$i = 1, 2, \cdots, k-1\}$$

where, since by (4.20), (B defined by (4.41)) $W_n \to W$ and $U_{in} \to U_i$

(4.35) $(U, W)'$ is multivariate normal $n_{2k-1}(\underset{\sim}{0},\ \gamma\delta D_{22}^{-1}/A_\phi,\ \Sigma_5)$

where

(4.35)
$$\Sigma_5 = \left(\begin{array}{c|c} \tau^2 D_{33} & E \\ E & D_{22}^{-1} \end{array} \right)$$

where

(4.37)
$$D_{33} = \text{diag}\left(\frac{1}{\lambda_{(1)}} + \frac{1}{\lambda_{(k)}} + \frac{t_{(1)}^2}{\lambda_{(1)} u_{(1)}} + \frac{t_{(k)}^2}{\lambda_{(k)} u_{(k)}}, \cdots, \right.$$
$$\left. \cdot \frac{1}{\lambda_{(k-1)}} + \frac{1}{\lambda_{(k)}} + \frac{t_{(k-1)}^2}{\lambda_{(k-1)} u_{(k-1)}} + \frac{t_{(k)}^2}{\lambda_{(k)} u_{(k)}} \right)$$

(4.38)
$$E = \begin{pmatrix} -\tau t_{(1)} & 0 & \tau t_{(k)} \\ & \ddots & \tau t_{(k)} \\ & & \vdots \\ 0 & -\tau t_{(k-1)} & \tau t_{(k)} \end{pmatrix}_{(k-1)\times k}$$

It follows from (4.34) and (4.35) that for $\omega_n \in \Omega^{(n)}$ $(\underset{\sim}{\varepsilon}, \underset{\sim}{\delta})$

(4.39)
$$\lim_{n\to\infty} P_{\omega_n}\{L_n \geq L_{n,\alpha},\ \tilde{\theta}_{(k)n} = \max_j \tilde{\theta}_{(j)n}\}$$
$$= \int_\infty^{\varepsilon_1} \int_\infty^{\varepsilon_2} \cdots \int_\infty^{\varepsilon_{k-1}} d\underset{\sim}{u} \int_S n_{2k-1}(\underset{\sim}{u}, \underset{\sim}{w};\ (\underset{\sim}{0}, \tau^{-1} D_{22}^{-2} \underset{\sim}{\delta}), \Sigma_5)\, d\underset{\sim}{w}$$

where

(4.40)
$$S = \{\underset{\sim}{w} : (\underset{\sim}{w} + \underset{\sim}{\xi})' B (\underset{\sim}{w} + \underset{\sim}{\xi}) > \chi^2_{k-1,\alpha}\}$$

(4.41)
$$B = D_{22} - \bar{u}^{-1} \underset{\sim}{1}_k \underset{\sim}{1}_k'.$$

Let

(4.42) $J_2(k-1, \underset{\sim}{\varepsilon}, \underset{\sim}{\delta}, \Sigma_5)$ denote the minimum of the RHS of (4.39) over D_{33} (i.e. choosing some j so that when $\lambda_{(k)} = \lambda_j$ and $t_{(k)} = t_j$, $u_{(k)} = u_j$, the RHS of (4.39) is minimized with respect ta D_{33}).

Then, it follows from (4.29), (4.31), (4.34), (4.35), (4.39) and (4.42) that we conclude the following

THEOREM 4.3 *For* $\omega_n \in \Omega^{(n)}(\underset{\sim}{\varepsilon}, \underset{\sim}{\delta})$,

$$\lim_{n\to\infty} P_{\omega_n}(\text{CS}|R_{I2}) \geq \min\{J_1(k-1, \underset{\sim}{\varepsilon}, \underset{\sim}{\delta}, \underset{\sim}{\lambda}, \underset{\sim}{t})\, H_{k-1}(\chi^2_{k-1,\alpha}; \Delta^*), \quad J_2(k-1, \underset{\sim}{\varepsilon}, \underset{\sim}{\delta}, \Sigma_5)\}$$

where functions J_1 *and* J_2 *are defined by* (4.30), (4.31) *and* (4.42). *The analogous result for* $\tilde{R}_{I2}$ *can also be obtained.*

Some Comparisons Among the Proposed Rules

Suppose a value α_0 (>0) is given and we wish to find the smallest sample size n for which the P^*-condition holds on $\Omega(\alpha_0)$. Then it follows from Theorem 4.1 that when R_{I1} is applied, a large sample solution can be obtained by putting $\Delta_1/\sqrt{n} = \alpha_0$, or

$$n = \left(\frac{\Delta_1}{\alpha_0}\right)^2 \tag{4.43}$$

where Δ_1 satisfies $I(k-1, \Delta_1; \underset{\sim}{a}, \underset{\sim}{v}) = P^*$ defined by (3.9).

To compare two different procedures R_1 and R_2, we are interested in comparing the sample sizes n_1 and n_2, denoted respectively, by $n(R_1)$ and $n(R_2)$ when P^*-condition is satisfied for both rules. We define the efficiency of R_2 with respect to R_1 by

$$\text{Eff }(R_2 : R_1) = n(R_1)/n(R_2). \tag{4.44}$$

Suppose it is known that $H_0 : \beta_1 = \beta_2 = \cdots = \beta_k$ holds, i.e. the parameter space is $\Omega_{10}(\alpha)$, and we take $L_{n,\alpha} = \infty$ in R_{I2}. Then for $t_1 = t_2 = \cdots = t_k = t$, we have, by Theorem 4.1 and Theorem 4.2,

$$\begin{aligned}&\text{Eff}_{H_0}(R_{I2} : R_{I1})\\ &\quad = \frac{n(R_{I1})}{n(R_{I2}|H_0)} = \left(\frac{\Delta_1}{\alpha_0}\right)^2 \Big/ \left(\frac{\Delta_3}{\alpha_0}\right)^2 = \left(\frac{\Delta_1}{\Delta_3}\right)^2\end{aligned} \tag{4.45}$$

where Δ_1 and Δ_3 satisfies

$$\int_{-\infty}^{\infty} \prod_{i=2}^{k} \Phi\left(\frac{a_i u + \Delta_1/v_i}{\sqrt{1-a_2^i}}\right) d\Phi(u) = P^* \tag{4.46}$$

where

$$a_i = \left(1 + \frac{w_{[i]}}{w_{[1]}}\right)^{-1/2}, \quad v_i = w_{[1]} + w_{[i]}, \quad w_j = \tau^2\left(\frac{1}{\lambda_j} + \frac{t_j^2}{\lambda_j u_j}\right)$$

$$a'_{i+1} = \left(1 - \frac{\lambda_{[k]}}{\lambda_{[i]}}\right)^{-1/2}, \quad v'_{i+1} = \tau^2\left(\frac{1}{\lambda_{[i]}} + \frac{1}{\lambda_{[k]}}\right),$$

$$i = 2, 3, \cdots, k$$
$$j = 1, 2, \cdots, k.$$

We note that Δ_3 is independent of $t = (t_1, t_2, \cdots, t_k)'$. Hence,

$$\text{(4.47)} \qquad \mathrm{Eff}_{H_0}(R_{I2} : R_{I1}) > 1 \quad \text{if and only if} \quad \Delta_1 > \Delta_3.$$

To see a special case for $k = 2$, it follows from (4.46) to obtain

$$\text{(4.48)} \quad \mathrm{Eff}_{H_0}(R_{I2} : R_{I1}) = \left(\frac{\Delta_1}{\Delta_3}\right)^2 = \left(\frac{v_2}{v'_2}\right)^2 = \left(1 + \frac{t_1^2}{u_1}\right)^2 \geq 1$$
$$\text{if} \quad u_1 = u_2 \ (> 0).$$

The strict inequality holds if $t_1 \neq 0$.

Hence, when the information H_0 is utilized, the sample size required for P^*-condition can be reduced by using $\hat{\underset{\sim}{\theta}}_n$ instead of $\tilde{\underset{\sim}{\theta}}_n$ in the selection procedure.

It follows from Theorem 4.1, Theorem 4.2 and Theorem 4.3 that we have

THEOREM 4.4. (i) $\mathrm{Eff}(\tilde{R}_{I1} : R_{I1}) = (\tau/\sigma)^2$ *where* σ^2 *is the variance of* F *and* τ *is defined by* (2.21).

(ii) $\mathrm{Eff}(\tilde{R}_{I1} : R_{I1} | H_0) = (1-\alpha)(\Delta_1/\Delta_3)^2 + \alpha$ *where* $L_{n,\alpha} \to \chi^2_{k-1,\alpha}$, Δ_1 *satisfies* (4.46) *and* Δ_3 *satisfies* $(0 \leq \alpha \leq 1)$

$$(1-\alpha) I(k-1, \Delta_3, \underset{\sim}{a}', \underset{\sim}{v}') + \alpha I(k-1, \Delta_3; \underset{\sim}{a}, \underset{\sim}{v}) = P^*.$$

Proof. Note that $(L_n | H_0) \xrightarrow{\mathscr{L}} \chi^2_{k-1}$ and under H_0,

$$\inf_{\Omega} P(\mathrm{CS} | R_2) \longrightarrow (1-\alpha) I(k-1, \Delta_3, \underset{\sim}{a}', \underset{\sim}{v}')$$
$$+ \alpha I(k-1, \Delta_3; \underset{\sim}{a}, \underset{\sim}{v}).$$

5. **A Monte Carlo study of $\tilde{R}_1$ and $\tilde{R}_2$ for small sample size.** We study the behaviors of rule $\tilde{R}_1$ and $\tilde{R}_2$ for small sampl size by Monte Carlo method. We consider the case $k = 2$ and $k = 5$. For some special given value of $P^* = .90$, sample size n, $\underset{\sim}{\theta}$ and $\underset{\sim}{\beta}$, we draw n equal sample from $F(x - \theta_i - \beta_i t_{ij})$ where F is the standard normal cdf. We fix $\bar{t}_{in} = 1$, and consider equal sample

size case. Our t_{ij} is so chosen that $\bar{t}_{in}=1$ for all i, n, and each t_{ij} is equally spaced with center $\bar{t}_{in}$. We use s to denote the length of equal space and $\bar{t}$ to denote the center of t_{ij}.

For given value of α (which is used in $\tilde{R}_2$), we draw n sample from each normal population and apply $\tilde{R}_1$ and $\tilde{R}_2$ respectively to select the population associated with the largest θ_i. Here we tabulate Table 1 ($k=2$) and Table 2 ($k=5$). We repeat the

Table 1

$P^*=.90$ $(B1, B2)=(1.0, 2.0)$ $(A1, A2)$

N	α	$(\bar{t}, s)$	(0.0, 0.1)	(0.0, 0.5)	(0.1, 1.0)
10	.50	(1.0, 1.0)	1.000 2.000 .500 1.000 1.000 2.000 .500	1.000 2.000 .500 1.000 1.000 2.000 .500	1.000 2.000 .500 1.000 1.000 2.000 .500
		(1.0, .5)	1.000 2.000 .500 1.001 1.000 1.998 .501	1.000 1.966 .501 1.000 1.000 1.980 .505	1.000 1.979 .505 1.023 1.000 1.935 .517
		(1.0, .1)	.959 1.874 .512 1.028 .959 1.823 .526	.970 1.846 .525 1.062 .970 1.738 .558	.990 1.773 .558 1.129 .990 1.570 .631
	.80	(1.0, 1.0)	1.000 2.000 .500 1.000 1.000 2.000 .500	1.000 2.000 .500 1.000 1.000 2.000 .500	1.000 2.000 .500 1.000 1.000 2.000 .500
		(1.0, .5)	1.000 1.996 .501 1.000 1.000 1.996 .501	1.000 1.995 .501 1.001 1.000 1.994 .502	1.000 1.979 .505 1.000 1.000 1.979 .505
		(1.0, .1)	.953 1.863 .512 1.012 .953 1.840 .518	.975 1.838 .530 1.040 .975 1.767 .552	.980 1.781 .550 1.048 .980 1.700 .576
16	.50	(1.0, 1.0)	1.000 2.000 .500 1.000 1.000 2.000 .500	1.000 2.000 .500 1.000 1.000 2.000 .500	1.000 2.000 .500 1.000 1.000 2.000 .500
		(1.0, .5)	1.000 2.000 .500 1.000 1.000 2.000 .500	1.000 2.000 .500 1.000 1.000 2.000 .500	1.000 2.000 .500 1.000 1.000 2.000 .500
		(1.0, .1)	.967 1.897 .510 1.068 .967 1.776 .544	.995 1.813 .549 1.088 .995 1.666 .597	.999 1.674 .597 1.046 .999 1.600 .624
	.80	(1.0, 1.0)	1.000 2.000 .500 1.000 1.000 2.000 .500	1.000 2.000 .500 1.000 1.000 2.000 .500	1.000 2.000 .500 1.000 1.000 2.000 .500
		(1.0, .5)	1.000 2.000 .500 1.000 1.000 2.000 .500	1.000 2.000 .500 1.000 1.000 2.000 .500	1.000 2.000 .500 1.000 1.000 2.000 .500
		(1.0, .1)	.972 1.899 .512 1.013 .972 1.875 .518	.996 1.836 .512 1.015 .996 1.808 .551	1.000 1.695 .590 1.005 1.000 1.687 .593
25	.50	(1.0, 1.0)	1.000 2.000 .500 1.000 1.000 2.000 .500	1.000 2.000 .500 1.000 1.000 2.000 .500	1.000 2.000 .500 1.000 1.000 2.000 .500
		(1.0, .5)	1.000 2.000 .500 1.000 1.000 2.000 .500	1.000 2.000 .500 1.000 1.000 2.000 .500	1.000 2.000 .500 1.000 1.000 2.000 .500
		(1.0, .1)	.990 1.944 .509 1.020 .990 1.906 .519	.998 1.845 .541 1.008 .998 1.831 .545	1.000 1.532 .653 1.000 1.000 1.532 .653
	.80	(1.0, 1.0)	1.000 2.000 .500 1.000 1.000 2.000 .500	1.000 2.000 .500 1.000 1.000 2.000 .500	1.000 2.000 .500 1.000 1.000 2.000 .500
		(1.0, .5)	1.000 2.000 .500 1.000 1.000 2.000 .500	1.000 2.000 .500 1.000 1.000 2.000 .500	1.000 2.000 .500 1.000 1.000 2.000 .500
		(1.0, .1)	.988 1.948 .507 1.001 .988 1.947 .507	.999 1.831 .546 1.000 .999 1.831 .546	1.000 1.554 .644 1.000 1.000 1.554 .644

process of sampling and selecting 1000 times and consider the frequency of correct selection in 1000 times as our probability of correct selection. We also compute the arithmetic mean of the numbe of populations in subset selected in 1000 times as our expected mean of subset selected. We denote it by a. Let the frequency of correct selection be denoted by p. Then, we define our efficiency of rule $\tilde{R}_i$ by the quantity $e_i = p/a$, $i = 1, 2$. Also, we consider the effieiency of $\tilde{R}_2$ with respect to $\tilde{R}_1$ by the quantity $e = e_2/e_1$. Let p_i and a_i denote respectively the quantity of p and a by using $\tilde{R}_i$, $i = 1, 2$.

Table 2

$P^* = .90$ $(B1, B2, B3, B4, B5) = (1.0, 1.5, 2.0, 2.5, 3.0)$ $(A1, A2, A3, A4, A5)$

N	α	$(\bar{t}, s)$	(0.0, 0.0, 0.0, 0.0, 0.5)	(0.0, 0.2, 0.4, 0.6, 0.8)	(0.0, 0.5, 0.5, 0.5, 1.0)
5	.60	(1.0, .2)	.893 4.132 .216 1.096 .893 3.771 .237	.902 4.162 .217 1.129 .902 3.685 .245	.924 4.116 .224 1.121 .924 3.672 .252
		(1.0, .8)	1.000 4.995 .200 1.109 1.000 4.503 .222	1.000 4.989 .200 1.130 1.000 4.416 .226	1.000 4.990 .200 1.138 1.000 4.384 .228
	.90	(1.0, .2)	.891 4.124 .216 1.001 .891 4.120 .216	.886 4.115 .215 1.002 .886 4.108 .216	.913 4.120 .222 1.001 .913 4.114 .222
		(1.0, .8)	1.000 4.996 .200 1.000 1.000 4.996 .200	1.000 4.987 .201 1.000 1.000 4.987 .201	1.000 4.990 .200 1.000 1.000 4.990 .200
13	.60	(1.0, .2)	.999 4.772 .209 1.007 .999 4.738 .211	.994 4.691 .212 1.010 .994 4.645 .214	.999 4.617 .216 1.016 .999 4.546 .202
		(1.0, .8)	1.000 5.000 .200 1.000 1.000 5.000 .200	1.000 5.000 .200 1.000 1.000 5.000 .200	1.000 5.000 .200 1.000 1.000 5.000 .200
	.90	(1.0, .2)	.998 4.756 .210 1.000 .998 4.756 .210	.996 4.684 .213 1.000 .996 4.684 .213	.999 4.581 .218 1.000 .999 4.581 .218
		(1.0, .8)	1.000 5.000 .200 1.000 1.000 5.000 .200	1.000 5.000 .200 1.000 1.000 5.000 .200	1.000 5.000 .200 1.000 1.000 5.000 .200
21	.60	(1.0, .2)	1.000 4.960 .202 1.000 1.000 4.960 .202	1.000 4.926 .203 1.000 1.000 4.926 .203	1.000 4.835 .207 1.000 1.000 4.835 .207
		(1.0, .8)	1.000 5.000 .200 1.000 1.000 5.000 .200	1.000 5.000 .200 1.000 1.000 5.000 .200	1.000 5.000 .200 1.000 1.000 5.000 .200
	.90	(1.0, .2)	1.000 4.957 .202 1.000 1.000 4.957 .202	1.000 4.907 .204 1.000 1.000 4.907 .204	1.000 4.820 .207 1.000 1.000 4.820 .207
		(1.0, .8)	1.000 5.000 .200 1.000 1.000 5.000 .200	1.000 5.000 .200 1.000 1.000 5.000 .200	1.000 5.000 .200 1.000 1.000 5.000 .200

In Table 1, associated with $P^* = .90$, sample size, test size α of T_2 in $\tilde{R}_2$, $(\bar{t}, s)$, $(B1, B2)$ ($= (\beta_1, \beta_2) = (1.\ 0, 0.\ 20)$), $(A1, A2)$ $(=(\theta_1, \theta_2))$, the entry of the first row are respectively p_1, a_1, e_1 and e and the second row are respctively p_2, a_2, e_2. It can be seen in most cases, $e \geq 1$. Table 2, the notations are analogousuly defined for $k = 5$.

REFERENCES

1. J.N. Adichie, *Estimates of regression parameters based on rank tests,* Ann. Math. Statist., **38** (1967), 894-904.
2. Capt. D.R. Barr and M.H. Rizvi, *An introduction to ranking and selection procedures,* J. Amer. Statist. Assoc., **61** (1966), 640-646.
3. R.E. Bechhofer, *The effect of preliminary test of significance on the size and power of certain tests of univaratie linear hypothesis,* (1951) Ph. D. Thesis.
4. H. Bozivich, T.A. Bancroft and H. Hartley, *Power of analysis of variance test procedures for certain incompletely specified models,* Ann. Math. Statist., **27** (1956), 1017-1043.
5. S.S. Gupta, *Bibliography on the multivariate normal integrals and related topics,* Ann. Math. Statist., **34** (1963), 829-838.
6. S. S. Gupta and S. Panchapakesan, *Multiple Decision Procedures: Theory and Methodology of Selecting and Ranking Populations,* Wiley, New York (1979).
7. E.L. Lehmann, *A class of selection procedures based on ranks,* Math. Ann., **150** (1963), 268-275.
8. A. K. Md. Ehsanes Saleh and P. K. Sen, *Nonparametric estimation of location parameter after a preliminary test on regression,* Ann. Statist., **6** (1978), 154-168.
9. P. K. Sen and A. K. Md. Ehsanes Saleh, *Nonparametric estimation of location parameter after a preliminary test on regression in the multivariate case,* J. Multivariate Anal., **9** (1979), 322-331.

Wen-Tao Huang
Institute of Statistics
Academia Sinica
Taipei, Taiwan, R.O.C.

A.K. Md. Ehsanes Saleh
Department of Mathematics and Statistics
Carleton University
Ottawa, CANADA