SOME MULTIPLE DECISION PROBLEMS IN ANALYSIS OF VARIANCE*

by

Shanti S. Gupta, Purdue University and

Deng-Yuan Huang, Academia Sinica, Taipei, Taiwan

Department of Statistics
Division of Mathematical Sciences
Mimeograph Series #458

July 1976

*This research was supported by the Office of Naval Research Contracts N00014-67-A-0226-0014 and N00014-75-C-0455 at Purdue University.

SOME MULTIPLE DECISION PROBLEMS IN ANALYSIS OF VARIANCE*

by

Shanti S. Gupta, Purdue University and

Deng-Yuan Huang, Academia Sinica, Taipei, Taiwan

1. Introduction

In most 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, in one-way ANOVA the hypothesis is rejected in actual application of the F-test, the resulting conclusion that the true means $\theta_1, \theta_2, \ldots, \theta_k$ are not all equal, would by itself usually be insufficient to satisfy the experimenter. In fact his problems would begin at this stage. The experimenter may desire to select the "best" population or a subset of the "good" populations; 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.

The extensive literature on selection and ranking procedures depends heavily on the use of independence between populations (block, treatments, etc.) in the analysis of variance. In practical applications, it is desirable to drop this assumption of independence and consider cases more general than the normal.

Our interest is to derive a method to construct locally best (in some sense) selection procedures to select a nonempty subset of the k populations

*This research was supported by the Office of Naval Research Contracts N00014-67-A-0226-0014 and N00014-75-C-0455 at Purdue University. Reproduction in whole or in part is permitted for any purpose of the United States Government.

containing the best population as ranked in terms of θ_i's (defined below) which control the size of the selected subset and maximizing the probability of selecting the best. We also consider the usual selection procedures in one way ANOVA based on the generalized least square estimates and apply the method to two way layout case. Some examples are discussed and some results on comparisons with other procedures are also considered.

2. Locally Best Selection Procedures

Let $\pi_1,\pi_2,\ldots,\pi_k$ represent k ($\geq$ 2) populations and let $X_{i1},\ldots,X_{in_i}$ be n_i independent random observations from π_i. The selection procedures will depend upon the observations through T_{ij} which are defined as follows.

Let $T_{ij} = T(X_{i1},\ldots,X_{in_i}; X_{j1},\ldots,X_{jn_j})$ be based on the n_i and n_j observations from π_i and π_j $(i,j = 1,2,\ldots,k)$, respectively. In a given problem the function T is so chosen as to indicate the differences between the populations in a reasonable way. For example, if the observations drawn from π_i are normally distributed with unknown mean θ_i, $(1 \leq i \leq k)$, and known variance σ^2, a choice of T_{ij} might be $\bar{X}_i-\bar{X}_j$, where $\bar{X}_i = \frac{1}{n_i}\sum_{\ell=1}^{n_i} X_{i\ell}$ and $\bar{X}_j = \frac{1}{n_j}\sum_{\ell=1}^{n_j} X_{j\ell}$. Now we assume that T_{ij} has a joint probability density function $g_{\tau_{ij}}(\cdot)$ depending on the parameter τ_{ij} and assume that τ_{ii}'s are known. Usually T_{ij}'s are chosen to obtain both sufficient and maximal invariant statistics for τ_{ij}'s. Let $\tau_i = \min_{j\neq i} \tau_{ij}$. Returning to the above normal means problem, we find that $\tau_{ij} = \theta_i-\theta_j$ and $\tau_i = \theta_i-\theta_{[k]}$ for $\theta_i < \theta_{[k]}$ and $\tau_i = \theta_i-\theta_{[k-1]}$ for $\theta_i = \theta_{[k]}$, where $\theta_{[1]} \leq \cdots \leq \theta_{[k]}$.

A population is said to be best if $\tau_i = \max_{1\leq j\leq k} \tau_j$. For the above normal means example, π_i is best if $\theta_i = \theta_{[k]}$ and in this case $\tau_{ii} = 0$ and $\tau_i = \theta_{[k]}-\theta_{[k-1]}$.

Let $\underline{\xi}_i = (\tau_{ij} | 1 \le j \le k,\ j \ne i)$, $\underline{\xi}_i^i = (\tau_{i\ell} | \tau_{i\ell} = \tau_{ii},\ 1 \le \ell \le k,\ \ell \ne i)$ and $\underline{z}_i = (z_{ij} | 1 \le j \le k,\ j \ne i)$ are all k-1 dimensional vectors. Assume that the joint density of T_{ij}, $j = 1,2,\ldots,k$, $j \ne i$, is $f_{\underline{\xi}_i}(\underline{z}_i)$, $1 \le i \le k$, (with respect to some σ-finite measure μ). Let δ_i be the probability of selecting π_i,

$$S_r = \{\delta: \delta = (\delta_1,\ldots,\delta_k),\ \sum_{i=1}^{k} \int \delta_i(\underline{z}_i) f_{\underline{\xi}_i^i}(\underline{z}_i) d\mu(\underline{z}_i) \le r\}$$

and

$$S_r' = \{\delta: \delta = (\delta_1,\ldots,\delta_k),\ \sum_{i=1}^{k} \int \delta_i(\underline{z}_i) f_{\underline{\xi}_i^i}(\underline{z}_i) d\mu(\underline{z}_i) = r\}.$$

__Theorem__: Let $\delta^0 = (\delta_1^0,\ldots,\delta_k^0) \in S_r'$ be defined by

$$\delta_i^0(\underline{z}_i) = \begin{cases} 1 & \text{if} \quad \min\limits_{\substack{1\le j\le k\\ j\ne i}} \frac{\partial}{\partial \tau_{ij}} f_{\underline{\xi}_i}(\underline{z}_i)\Big|_{\underline{\xi}_i^i} > c f_{\underline{\xi}_i^i}(\underline{z}_i), \\ \lambda_i & \qquad\qquad\qquad\qquad = \qquad\qquad , \\ 0 & \qquad\qquad\qquad\qquad < \qquad\qquad . \end{cases}$$

Then δ^0 maximizes $\sum_{i=1}^{k} \min_{j\ne i} \int \delta_i(\underline{z}_i) \frac{\partial}{\partial \tau_{ij}} f_{\underline{\xi}_i}(\underline{z}_i)\Big|_{\underline{\xi}_i^i} d\mu(\underline{z}_i)$ among all rules $\delta \in S_r$. δ^0 is called a locally best procedure in this sense.

__Proof__. For any $\delta \in S_r$,

$$\begin{aligned} &\sum_{i=1}^{k} \int \delta_i(\underline{z}_i) \min_{j\ne i} \frac{\partial}{\partial \tau_{ij}} f_{\underline{\xi}_i}(\underline{z}_i)\Big|_{\underline{\xi}_i^i} d\mu(\underline{z}_i) \\ &\quad - \sum_{i=1}^{k} \int \delta_i^0(\underline{z}_i) \min_{j\ne i} \frac{\partial}{\partial \tau_{ij}} f_{\underline{\xi}_i}(\underline{z}_i)\Big|_{\underline{\xi}_i^i} d\mu(\underline{z}_i) \\ &= \sum_{i=1}^{k} \int [\delta_i(\underline{z}_i) - \delta_i^0(\underline{z}_i)][\min_{j\ne i} \frac{\partial}{\partial \tau_{ij}} f_{\underline{\xi}_i}(\underline{z}_i)\Big|_{\underline{\xi}_i^i} - c f_{\underline{\xi}_i^i}(\underline{z}_i)] d\mu(\underline{z}_i) \\ &\quad + c \sum_{i=1}^{k} \int [\delta_i(\underline{z}_i) - \delta_i^0(\underline{z}_i)] f_{\underline{\xi}_i^i}(\underline{z}_i) d\mu(\underline{z}_i) \le 0. \end{aligned}$$

This proof is complete.

Example: Let $g_{\underline{\theta}}(\underline{x}) = \prod_{j=1}^{k} g_{\theta_j}(\bar{x}_j)$, where $g_{\theta_i}(\bar{x}_i) = \frac{\sqrt{n}}{\sqrt{2\pi}} e^{-\frac{n}{2}(\bar{x}_i-\theta_i)^2}$. Let $\tau_{ij} = \theta_i-\theta_j$, $1 \leq j \leq k$, $j \neq i$, $\tau_{ii} = 0$, and $Z_{ij} = \bar{X}_i-\bar{X}_j$, $j \neq i$. We know that a maximal invariant under a group G is $T_i = (\bar{X}_i-\bar{X}_1,\ldots,\bar{X}_i-\bar{X}_k)$ where G is the group of transformations

$$g\underline{z}_i = (z_{i1} + c,\ldots,z_{ik} + c),\ -\infty < c < \infty,$$

which, in the parameter space, induces the transformations $\bar{g}\,\tau_{ij} = \tau_{ij}+c$.

Since $\Sigma_{(k-1)x(k-1)} = \frac{1}{n}\begin{pmatrix} 2 & & 1 \\ & \ddots & \\ 1 & & 2 \end{pmatrix}$ is the covariance matrix of Z_{ij}'s. We know that Σ is positive definite, and

$$\Sigma^{-1} = \frac{n}{k}\begin{pmatrix} k-1 & & \\ & k-1 & -1 \\ -1 & \ddots & \\ & & k-1 \end{pmatrix}.$$

Hence

$$f_{\underline{\xi}_i}(\underline{z}_i) = (2\pi)^{-\frac{k-1}{2}} |\Sigma|^{-\frac{1}{2}} \exp\{(-\tfrac{n}{k})[((k-1)(z_{i1}-\tau_{i1})^2-(z_{i1}-\tau_{i1})(z_{i2}-\tau_{i2})$$
$$-\ldots-(z_{i1}-\tau_{i1})(z_{ik}-\tau_{ik})) +\ldots+ (-(z_{i1}-\tau_{i1})(z_{ik}-\tau_{ik})$$
$$-\ldots-(z_{i(k-1)}-\tau_{i(k-1)})(z_{ik}-\tau_{ik})+(k-1)(z_{ik}-\tau_{ik})^2)]\}.$$

Thus

$$\frac{\partial}{\partial\tau_{ij}} f_{\underline{\xi}_i}(\underline{z}_i)\Big|_{\underline{\xi}_i^i} = (2\pi)^{-\frac{k-1}{2}} |\Sigma|^{-\frac{1}{2}} \exp\{(-\tfrac{n}{k})[((k-1)z_{i1}^2$$
$$-z_{i1}z_{i2}-\ldots-z_{i1}z_{ik}) +\ldots+ (-z_{i1}z_{ik}-\ldots-z_{i(k-1)}z_{ik}$$
$$+(k-1)z_{ik}^2)]\}\cdot\{-2z_{i1}-\ldots-2z_{i(j-1)}+2(k-1)z_{ij}-2z_{i(j+1)}-$$
$$\ldots-2z_{ik}\}.$$

Hence

$$\delta_i^0(\underline{z}_i) = \begin{cases} 1 & \text{if } \min_{j \neq i}[-\sum_{\substack{\ell=1 \\ \ell \neq i,j}}^{k} Z_{i\ell} + (k-1)z_{ij}] \geq c, \\ 0 & \qquad\qquad\qquad\qquad\qquad\qquad < \quad , \end{cases}$$

or

$$\delta_i^0(\underline{z}_i) = \begin{cases} 1 & \text{if } \max_{\substack{1 \leq j \leq k \\ j \neq i}} \bar{X}_j \leq \frac{1}{k} \sum_{\ell=1}^{k} \bar{X}_\ell - c, \\ 0 & \qquad\qquad > \qquad\qquad . \end{cases}$$

3. Usual Approach to Selection Problems in One Way Layout

Let $\pi_1, \pi_2, \ldots, \pi_k$ be k populations. Let $X_{i1}, \ldots, X_{in_i}$ denote n_i independent observations from the ith population π_i. Let the joint density of $X_{11}, \ldots, X_{1n_1}; X_{21}, \ldots, X_{2n_2}; \ldots; X_{k1}, \ldots, X_{kn_k}$ be of the following form:

(3.1) $$c_k|\Lambda|^{-1}\, g((\underline{x}-\underline{\eta})'\Lambda^{-1}(\underline{x}-\underline{\eta}))$$

where $\underline{x}' = (x_{11}, \ldots, x_{1,n_1}; \ldots; x_{k1}, \ldots, x_{kn_k})$, $\underline{\eta}' = (\overbrace{\theta_1, \ldots, \theta_1}^{n_1}; \ldots; \overbrace{\theta_k, \ldots, \theta_k}^{n_k})$ and Λ is a known positive definite matrix and c_k is determined such that (3.1) is a density.

Let $\Omega = \{\underline{\theta}: \underline{\theta}' = (\theta_1, \ldots, \theta_k)\}$ and also, let

$$A_{N\times k} = \begin{pmatrix} \underline{1}_{n_1} & & & 0 \\ & \underline{1}_{n_2} & & \\ & & \ddots & \\ 0 & & & \underline{1}_{n_k} \end{pmatrix},$$

where $\underset{\sim}{1}'_{n_i} = (1,\ldots,1)$ with n_i components and $\sum_{i=1}^{k} n_i = N \geq k$.

We consider the analysis of variance problem in a one way layout; let

$$x_{ij} = \theta_i + e_{ij},\ j=1,\ldots,n_i;\ i=1,\ldots,k$$

which is in the form of the general linear model

$$\underset{\sim}{x} = A\underset{\sim}{\theta} + \underset{\sim}{e},$$

where $\underset{\sim}{e}' = (e_{11},\ldots,e_{1n_1};\ldots;e_{k1},\ldots,e_{kn_k})$ with $\mathrm{var}(\underset{\sim}{e}) = \Lambda$.

We know that the generalized least square estimator of $\underset{\sim}{\theta}$ is

$$\hat{\underset{\sim}{\theta}} = (A'\Lambda^{-1}A)^{-1}A'\Lambda^{-1}\underset{\sim}{x} = \begin{pmatrix} y_1 \\ \vdots \\ y_k \end{pmatrix} = \underset{\sim}{y}.$$

Since $\hat{\underset{\sim}{\theta}} = B\underset{\sim}{x}$, $B = (A'\Lambda^{-1}A)^{-1}A'\Lambda^{-1}$, the joint density of $\underset{\sim}{Y}' = (Y_1,\ldots,Y_k)$ is of the form

$$(3.2)\qquad b_k|\Lambda_1|^{-\frac{1}{2}}\, h((\underset{\sim}{y}-\underset{\sim}{\theta})'\Lambda_1^{-1}(\underset{\sim}{y}-\underset{\sim}{\theta}))$$

where $\Lambda_1 = B\Lambda B' = (\sigma_{ij})$.

The ordered θ_i's are denoted by $\theta_{[1]} \leq \cdots \leq \theta_{[k]}$. It is assumed that there is no prior knowledge of the correct pairing of the ordered and the unordered θ_i's. Let us denote by $\pi_{(i)}$ the population (unknown) associated with $\theta_{[i]}$, $i = 1,2,\ldots,k$. Our goal is to select a non-empty subset of the k populations so as to include the population associated with $\theta_{[k]}$. Defining any such selection as a correct selection, we wish to define a procedure R so that $P(CS|R)$, the probability of a correct selection, is at least a preassigned number $P^*(\frac{1}{k} < P^* < 1)$. We will refer to this requirement as the P^*-condition. We propose the following rule R based on Y_i, $1 \leq i \leq k$.

R: Retain π_i in the selected subset if and only if

$$Y_i \geq \max_{1 \leq j \leq k} (Y_j - c\sqrt{\sigma_{ii}+\sigma_{jj}-2\sigma_{ij}}),$$

where $c = c(k,P^*;n_i,\sigma_{ij},1 \leq i,\ j \leq k) > 0$ is chosen so as to satisfy the P^*-condition.

Let $Y_{(i)}$ and $\sigma_{(i)(i)}$ denote the observation and the variance associated with the population $\pi_{(i)}$ with mean $\theta_{[i]}$, $i = 1,2,\ldots,k$. Of course, both $Y_{(i)}$ and $\sigma_{(i)(i)}$ are unknown as in $\sigma_{(i)(j)}$, the covariance of $Y_{(i)}$ and $Y_{(j)}$. Thus

$$(3.3) \quad P(CS|R) = P\{Y_{(k)} \geq \max_{1 \leq j \leq k} (Y_{(j)} - c\sqrt{\sigma_{(k)(k)}+\sigma_{(j)(j)}-2\sigma_{(k)(j)}})\}$$

$$= P\{Z_{jk} \leq c+(\theta_{[k]}-\theta_{[j]})(\sigma_{(k)(k)}+\sigma_{(j)(j)}-2\sigma_{(k)(j)})^{-\frac{1}{2}},\ 1 \leq j \leq k-1\},$$

where for ℓ, $1 \leq \ell \leq k$, we define

$$(3.4) \quad Z_{r\ell} = (Y_{(r)}-Y_{(\ell)}-\theta_{[r]}+\theta_{[\ell]})(\sigma_{(r)(r)}+\sigma_{(\ell)(\ell)}-2\sigma_{(r)(\ell)})^{-\frac{1}{2}},$$

for $r = 1,2,\ldots,k$, $r \neq \ell$.

Let $\underline{Z}_\ell = \underline{Y} A_\ell$, where $\underline{Z}_\ell = (Z_{r\ell}: r = 1,2,\ldots,k,\ r \neq \ell)$ and $\underline{Y} = (Y_{(1)}-\theta_{[1]},\ldots,Y_{(k)}-\theta_{[k]})$.

The matrix A_ℓ with k rows and $(k-1)$ columns is defined as follows:

Let $\alpha_{r\ell} = (\sigma_{(r)(r)}+\sigma_{(\ell)(\ell)}-2\sigma_{(r)(\ell)})^{-\frac{1}{2}}$, $1 \leq r,\ \ell \leq k$, $r \neq \ell$;

$$(3.5) \quad A_\ell = \begin{pmatrix} \alpha_{1\ell} & 0 & \cdots & 0 & 0 & \vdots & 0 \\ 0 & \alpha_{2\ell} & \cdots & & & \vdots & \\ \vdots & 0 & & & & \vdots & \\ \vdots & \vdots & & & & \vdots & \\ 0 & 0 & & \alpha_{\ell-1,\ell} & 0 & \vdots & 0 \\ -\alpha_{1\ell} & -\alpha_{2\ell} & \cdots & -\alpha_{\ell-1,\ell} & -\alpha_{\ell+1,\ell} & \vdots & -\alpha_{k,\ell} \\ 0 & 0 & & 0 & \alpha_{\ell+1,\ell} & \vdots & 0 \\ & & & & 0 & \vdots & \\ 0 & 0 & & 0 & 0 & \vdots & 0 \\ & & & & & & \alpha_{k,\ell} \end{pmatrix}$$

and $\Sigma_1 = (\sigma_{(i)(j)})$.

Since $A_\ell'\Sigma_1 A_\ell = (\rho_{ij}^{(\ell)})$, i, j = 1,2,...,k; i, j ≠ ℓ, for $1 \leq \ell \leq k$ and A_ℓ is of rank k-1, hence the joint density of $\underline{Z}_\ell$ is

(3.6) $$\mathcal{L}(z_\ell) = a_k|A_\ell'\Sigma_1 A_\ell|^{-\frac{1}{2}} f(\underline{z}_\ell'(A_\ell'\Sigma_1 A_\ell)^{-1}\underline{z}_\ell).$$

For any given association between (σ_{ij}, i, j=1,2,...,k) and ($\sigma_{(i)(j)}$; i,j = 1,2,...,k) we can see from (3.3) that the infimum of P(CS|R) is attained when $\theta_{[1]} = \ldots = \theta_{[k]}$.

Hence,

(3.7) $$\inf_{\underline{\theta}\in\Omega} P(CS|R) = \min_{1\leq \ell \leq k} P\{Z_{r\ell} \leq c,\ r=1,2,\ldots,k, r\neq\ell;\ \{\rho_{ij}^{(\ell)}\}\}.$$

For $\ell < k$, let κ_{ij}^{ℓ} be such that

(3.8)

(i) $\kappa_{ij}^{\ell} \geq \kappa_{ij}^{k}$, i, j≠ℓ, i, j≠k;

(ii) $\kappa_{ij}^{\ell} \geq \kappa_{\ell j}^{k}$, j=1,2,...,k; j≠ℓ, k,

and for any ℓ, ($1 \leq \ell \leq k$), there exists an m, ($1 \leq m \leq k$), such that for any i, j, i, j=1,2,...,k, i, j≠ℓ, i≠j,

$$\kappa_{ij}^{(\ell)} = \rho_{r,s}^{(m)}$$

for some r, s, r≠s, r, s≠m, r, s=1,2,...,k.

<u>Lemma</u> [3]. If $\underline{X}' = (X_1,\ldots,X_p)$ has density $|\Sigma|^{-\frac{1}{2}}f(\underline{x}'\Sigma^{-1}\underline{x})$, then for any two positive definite (symmetric) pxp matrices $\Gamma_1 = (r_{ij})$ and $\Gamma_2 = (\sigma_{ij})$ such that $r_{ii} = \sigma_{ii}$, $1 \leq i \leq p$ and $r_{ij} \geq \sigma_{ij}$, $1 \leq i < j \leq p$,

$$P_{\Gamma_1}\{X_1 < \ell_1,\ldots,X_p < \ell_p\} \geq P_{\Gamma_2}\{X_1 \leq \ell_1,\ldots,X_p \leq \ell_p\}$$

for any real numbers $\ell_1,\ldots,\ell_p$.

By (3.8) and Lemma, we have

(3.9) $$\inf_{\underset{\sim}{\theta}\in\Omega} P_{\underset{\sim}{\theta}}(CS|R)$$

$$= \min_{1\le \ell\le k} P\{Z_{r\ell} \le c,\ r=1,2,\ldots,k;\ r\neq \ell;\ \{\kappa_{ij}^{\ell}\}\}$$

$$= P\{Z_{rk} \le c,\ r=1,2,\ldots,k-1;\ \{\kappa_{ij}^{k}\}\}$$

$$= \int_{-\infty}^{c} \cdots \int_{-\infty}^{c} a_k |\{\kappa_{ij}^{k}\}|^{-\frac{1}{2}} f(\underset{\sim}{z}_k'(\{\kappa_{ij}^{k}\})^{-1}\underset{\sim}{z}_k)dz_{1k}\cdots dz_{k-1,k}.$$

Discussion of Condition (3.8)

For computational convenience, we assume that $\Lambda=\text{diag}(\lambda_{11},\ldots,\lambda_{11};\ldots;\lambda_{kk},\ldots,\lambda_{kk})$, $\lambda_{ii} > 0$, $i=1,2,\ldots,k$. If also the components $X_{11},\ldots,X_{1n_1};\ldots;X_{k1},\ldots,X_{kn_k}$ are independent then the joint distribution g as in (3.1) is multivariate normal (see Kelker [9, p. 18]). Then

$$A'\Lambda^{-1}A = \text{diag}\left(\frac{n_1}{\lambda_{11}}, \frac{n_2}{\lambda_{22}},\ldots,\frac{n_k}{\lambda_{kk}}\right) \text{ and}$$

$$B = (A'\Lambda^{-1}A)^{-1}A'\Lambda^{-1} = \begin{pmatrix} n_1^{-1} & & \\ \vdots & & \\ n_1^{-1} & & \\ & n_2^{-1} & \\ & \vdots & \\ & n_2^{-1} & \\ & & n_k^{-1} \\ & & \vdots \\ & & n_k^{-1} \end{pmatrix},$$

$\Sigma = B\Lambda B' = \text{diag}(n_1^{-1}\lambda_{11}, n_2^{-1}\lambda_{22},\ldots,n_k^{-1}\lambda_{kk})$. Then $\sigma_{ii} = n_i^{-1}\lambda_{ii} = m_i^{-1}$ and $\sigma_{ij} = 0$ for all $i \neq j$. Let $m_{(i)}^{-1} = n_{(i)}^{-1}\lambda_{(i)(i)}$, $1 \le i \le k$ and $m_{[1]} \le\cdots\le m_{[k]}$. Then $\alpha_{r\ell} = (m_{(r)}^{-1} + m_{(\ell)}^{-1})^{-\frac{1}{2}}$, $1 \le r, \ell \le k$, $r \neq \ell$; and for any ℓ, $1 \le \ell \le k$,

$$\rho_{ij}^{(\ell)} = \alpha_{i\ell}\alpha_{j\ell}m_{(\ell)}^{-1} = \frac{1}{(1+\frac{m_{(\ell)}}{m_{(i)}})^{\frac{1}{2}}(1+\frac{m_{(\ell)}}{m_{(j)}})^{\frac{1}{2}}},$$

for $i \neq j$, $i, j \neq \ell$. Let

(3.10) $$\kappa_{ij}^{\ell} = \frac{1}{(1+\frac{m_{[\ell]}}{m_{[i]}})^{\frac{1}{2}}(1+\frac{m_{[\ell]}}{m_{[j]}})^{\frac{1}{2}}}, \quad i \neq j,\ i, j \neq \ell.$$

Then, it is easy to check that the condition (3.8) is satisfied.

Expected subset size for a special case

Let the joint density p as in (3.2) have the form

(3.11) $$p(\underline{x}) = h(\underline{y}'\Sigma^{-1}\underline{y})$$

where $\Sigma = (\sigma_{ij})$ is positive definite with $\sigma_{11} = \sigma_{22} = \ldots = \sigma_{kk} = \sigma^2$ and $\sigma_{ij} = \sigma^2\rho$ when $i \neq j$, σ and ρ are known. Let S be the size of the selected subset excluding the best population. Then the expected subset size is given by

$$E(S|R) = \sum_{i=1}^{k-1} P\{\text{Selecting } \pi_{(i)}|R\}$$

$$= \sum_{i=1}^{k-1} P\{Y_{(i)} \geq \max_{1\leq j\leq k} Y_{(j)} - c\sigma\sqrt{2(1-\rho)}\}$$

$$= \sum_{\ell=1}^{k-1} \int_{B_\ell+\underline{\theta}_\ell} a_k|(\rho_{ij}^{(\ell)})|^{-\frac{1}{2}} f(\underline{z}_\ell'(\rho_{ij}^{(\ell)})\underline{z}_\ell)d\underline{z}_\ell,$$

where $\underline{\theta}_\ell = (\theta_{[\ell]}-\theta_{[r]})[2\sigma^2(1-\rho)]^{-\frac{1}{2}}$, $r=1,2,\ldots,k$, $r\neq\ell$,

$B_\ell+\underline{\theta}_\ell = \{Z_{r\ell} \leq c+(\theta_{[\ell]}-\theta_{[r]})[2\sigma^2(1-\rho)]^{-\frac{1}{2}},$

$r = 1,2,\ldots,k$, $r\neq\ell\}$, $1 \leq \ell \leq k$,

and $\rho_{ij}^{(\ell)}$ defined as in (3.6) is

$$\rho_{ij}^{(\ell)} = \begin{cases} 1 & \text{if } i = j,\ i, j \neq \ell, \\ \frac{1}{2} & \text{if } i \neq j,\ i, j \neq \ell. \end{cases}$$

We assume that f is strictly decreasing. Then f is Schur-concave [8]. Since $\underline{y} \in B_\ell$ and $\underline{x} \prec \underline{y}$ implies $\underline{x} \in B_\ell$, hence

$$\int_{B_\ell + \underline{\theta}_\ell} a_k |(\rho_{ij}^{(\ell)})|^{-\frac{1}{2}} f(\underline{z}_\ell'(\rho_{ij}^{(\ell)})^{-1}\underline{z}_\ell)d\underline{z}_\ell$$

is a Schur-concave function of $\underline{\theta}_\ell$ [10]. From the fact that $(a_1,a_2,\ldots,a_n) \succ (\frac{\Sigma a_i}{n})(1,\ldots,1)$ for all vectors $\underline{a}$, where $\underline{a} = (a_1,\ldots,a_n)$, $\underline{b} = (b_1,\ldots,b_n)$ and $a_1 \geq a_2 \geq \ldots \geq a_n$, $b_1 \geq b_2 \geq \ldots \geq b_n$, $\underline{a} \succ \underline{b}$ means $\sum_{i=1}^{\ell} a_i \geq \sum_{i=1}^{\ell} b_i$, $\ell = 1,\ldots,n-1$, $\sum_{i=1}^{n} a_i = \sum_{i=1}^{n} b_i$. For any ℓ, $2 \leq \ell \leq k-1$,

$$(\theta_{[\ell]}-\theta_{[1]},\ldots,\theta_{[\ell]}-\theta_{[\ell-1]},\theta_{[\ell]}-\theta_{[\ell+1]},\ldots,\theta_{[\ell]}-\theta_{[k]})$$

$\succ (\theta,\ldots,\theta)$, for some θ.

But $\theta_{[\ell]}-\theta_{[j]} \leq 0$ for $j > \ell$ and $\theta_{[\ell]}-\theta_{[j]} \geq 0$ for $j \leq \ell$, hence it follows that the supremum of

$$\int_{B_\ell + \underline{\theta}_\ell} a_k |(\rho_{ij}^{(\ell)})|^{-\frac{1}{2}} f(\underline{z}_\ell'(\rho_{ij}^{(\ell)})^{-1}\underline{z}_\ell)d\underline{z}_\ell$$

over Ω occurs when $\theta_{[1]} = \ldots = \theta_{[k]}$. For $\ell = 1$,

$$(\theta_{[1]}-\theta_{[2]},\ldots,\theta_{[1]}-\theta_{[k]}) \leq (0,\ldots,0)$$

and $B_1+\theta_1 \subset B_1$. Hence

$$\sup_{\underline{\theta}\in\Omega} E_{\underline{\theta}}(S|R) = \sum_{\ell=1}^{k-1} \int_{B_\ell} a_k |(\kappa_{ij}^{\ell})|^{-\frac{1}{2}} f(\underline{z}_\ell'(\kappa_{ij}^{\ell})^{-1}\underline{z}_\ell)d\underline{z}_\ell$$

$= (k-1)P^*$ provided that

$$\inf_{\underline{\theta}\in\Omega} P_{\underline{\theta}}(CS|R) = P^*.$$

Remark: Let p be the multivariate normal density as in (3.11), then f has the required property.

3.1. Applications to Normal Populations

Let $\pi_1,\pi_2,\ldots,\pi_k$ be k independent normal populations with means $\mu_1,\mu_2,\ldots,\mu_k$ and variances $\sigma_1^2,\sigma_2^2,\ldots,\sigma_k^2$, respectively. Let $\sigma_1^2 =\ldots= \sigma_k^2 = \sigma^2$, where σ^2 may or may not be known.

Case (a): σ^2 known. We assume without any loss of generality that $\sigma^2 = 1$, and for this problem (3.2) assumes the following form:

$$p(\underline{x}) = (2\pi)^{-\frac{k}{2}}|D|^{-\frac{1}{2}}\, h((\underline{y}-\underline{\mu})'D^{-1}(\underline{y}-\underline{\mu})),$$

where $\underline{\mu}' = (\mu_1,\ldots,\mu_k)$, $h(x) = e^{-x}$, and $D = \mathrm{diag}(n_1^{-1},\ldots,n_k^{-1})$.

Gupta and Huang [6] proposed the following rule R_1 based on the sample means Y_i from π_i, $i = 1,2,\ldots,k$.

R_1: Retain π_i in the selected subset if and only if

$$Y_i \geq \max_{1\leq j\leq k} \left(Y_j - c_1\sqrt{\frac{1}{n_i}+\frac{1}{n_j}}\right),$$

where $c_1 = c_1(k,P^*,n_1,\ldots,n_k) > 0$ is chosen so as to satisfy the P*-condition.

For the condition (3.10), $\lambda_{ii} = 1$ implies $m_i^{-1} = n_i^{-1}$, $1 \leq i \leq k$. Therefore any ℓ, $1 \leq \ell \leq k$,

$$\kappa_{ij}^{\ell} = \left[\left(1+\frac{n_{[\ell]}}{n_{[k]}}\right)\left(1+\frac{n_{[\ell]}}{n_{[j]}}\right)\right]^{-\frac{1}{2}},\ i \neq j,\ i,j \neq \ell,$$

and

$$\kappa_{ii}^{\ell} = 1,\quad 1 \leq i \leq k,\ i \neq \ell.$$

Let $\beta_i = \left(1+\frac{n_{[k]}}{n_{[i]}}\right)^{-\frac{1}{2}}$, $i = 1,2,\ldots,k-1$. Thus $\kappa_{ij}^{k} = \beta_i\beta_j$, $i \neq j$, $i,j = 1,\ldots,k-1$ and $\kappa_{ii}^{k} = 1$, $1 \leq i \leq k-1$.

By (3.7), we have

(3.12) inf P(CS|R)

$$= \int_{-\infty}^{c_1} \cdots \int_{-\infty}^{c_1} (2\pi)^{-\frac{k}{2}} |(\kappa_{ij}^k)|^{-\frac{1}{2}} f(\underline{z}_k'(\kappa_{ij}^k)^{-1}\underline{z}_k) dz_{1k} \cdots dz_{k-1,k},$$

where $Z_{1k},\ldots,Z_{k-1,k}$ are standard normal random variables with correlation $\kappa_{rs}^k = \beta_r\beta_s$. It is known that $Z_{1k},\ldots,Z_{k-1,k}$ can be generated from k independent standard variates $Y_1,\ldots,Y_{k-1},Y$ by the transformation

$$Z_{jk} = (1-\beta_j^2)^{\frac{1}{2}} Y_j + \beta_j Y,$$

and then (3.9) is as follows:

$$\text{(3.13)} \qquad \inf P(CS|R_1) = \int \prod_{j=1}^{k-1} \Phi\left(\frac{c_1-\beta_j u}{(1-\beta_j^2)^{\frac{1}{2}}}\right) d\Phi(u).$$

<u>Case (b)</u>: σ^2 unknown. Let s_ν^2 denote the usual pooled estimate of σ^2 on ν degrees of freedom. Gupta and Huang [6] proposed the rule R_2 for selecting a subset containing the population associated with the largest μ_i's.

R_2: Retain π_i in the selected subset if and only if

$$Y_i \geq \max_{1\leq j\leq k} \left(Y_j - c_2 s_\nu \sqrt{\frac{1}{n_i} + \frac{1}{n_j}}\right),$$

where $c_2 = c_2(k,P^*,n_1,\ldots,n_k) > 0$ is to be determined so that the P^*-condition is satisfied.

Using the same argument as in case (a), we can obtain

$$\text{(3.14)} \qquad \inf P(CS|R_2) = \int_0^\infty \int_{-\infty}^\infty \prod_{j=1}^{k-1} \Phi\left[\frac{c_2 u - \beta_j x}{\sqrt{1-\beta_j^2}}\right] d\Phi(x) dQ_\nu(u)$$

where $Q_\nu(u)$ denotes the cdf of a $\chi_\nu/\sqrt{\nu}$ variate.

The Evaluation of the Constant c_1 Associated with R_1

Let $U_1,\ldots,U_{k-1}$ be k-1 standard normal random variables, and the correlation coefficient of U_i and U_j be ρ, i,j = 1,...,k-1, where

$$\rho = [(1 + \frac{n_{[k]}}{n_{[1]}})(1 + \frac{n_{[k]}}{n_{[2]}})]^{-\frac{1}{2}}.$$

Using the same notation as before, we have $\kappa_{ij}^k \geq \rho$, i,j = 1,...,k-1, hence

$$P\{Z_{ik} \leq c,\ i = 1,\ldots,k-1,\ \{\kappa_{ij}^k\}\}$$
$$\geq P\{U_i \leq c,\ i = 1,\ldots,k-1,\ \{\rho\}\}$$
$$= \int_{-\infty}^{\infty} \Phi^{k-1}\left(\frac{c-\rho^{\frac{1}{2}}u}{\sqrt{1-\rho}}\right)d\Phi(u).$$

Equating the above integral to P*, values of c are available for the equicorrelated U_i's from the tables in Gupta, Nagel and Panchapakesan [7]. These c-values will be greater than the exact c_1-values satisfying the equations by equating the left hand side of (3.13) to P*. Some of the exact c_1-values can be obtained from Table 1 of Gupta and Huang [6].

Some Results on Comparisons

Assume that $\sigma^2 = 1$. The procedure of Gupta and D. Y. Huang [6] is more efficient than Gupta and W. T. Huang [5] for the case of k = 2, $n_{[1]} = \alpha n_{[2]}$, $0 < \alpha < 1$.

For σ^2 unknown, Chen, Dudewicz and Lee [2], have proposed a class of procedures as follows:

R_a: Retain π_i in the selected subset if and only if

$$Y_i \geq \max_{1\leq j\leq k} Y_j - q_a s_\nu\sqrt{\frac{1}{n_i} + \frac{1}{a}},$$

where a is any fixed constant such that $0 < a < \infty$.

For any fixed P^*, $\frac{1}{k} < P^* < 1$, and $k = 2$, $n_1 \neq n_2$,

$$\inf_{\Omega} P(CS|R_a) = \int \Phi\left(\frac{\sqrt{\frac{1}{n_{[2]}} + \frac{1}{a}}}{\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}\, q_a x\right) dQ_\nu(x) = P^*,$$

and

$$\inf_{\Omega} P(CS|R_1 = \int \Phi(c_1 x) dQ_\nu(x) = P^*,$$

hence $c_1\sqrt{\frac{1}{n_1} + \frac{1}{n_2}} = q_a\sqrt{\frac{1}{n_{[2]}} + \frac{1}{a}}$. Since

$$\sup_{\Omega} E(S|R_a) = \sup_{\Omega} \sum_{i=1}^{2} P(Y_i \geq \max_{1\leq j\leq 2} Y_j - q_a s_\nu\sqrt{\frac{1}{n_i} + \frac{1}{a}})$$

$$> \sup_{\Omega} \sum_{i=1}^{2} P(Y_i \geq \max_{1\leq j\leq 2} Y_j - q_a s_\nu\sqrt{\frac{1}{n_{[2]}} + \frac{1}{a}})$$

$$= \sup_{\Omega} \sum_{i=1}^{2} P(Y_i \geq \max_{1\leq j\leq 2} Y_j - c_1 s_\nu\sqrt{\frac{1}{n_1} + \frac{1}{n_2}})$$

$$= \sup_{\Omega} E(S|R_1).$$

4. Selection for Small Variances of Normal Populations

Let $\pi_1,\pi_2,\ldots,\pi_k$ denote k independent normal populations with unknown variances $\sigma_1^2,\sigma_2^2,\ldots,\sigma_k^2$, respectively, ($\sigma_i > 0$, $i = 1,2,\ldots,k$), and with all means known or unknown. The ordered variances are denoted by $\sigma_{[1]}^2 \leq\cdots\leq \sigma_{[k]}^2$. It is assumed that there is no a priori information available about the correct pairing of the given populations and the ordered parameters $\sigma_{[i]}^2$. The population with variance equal to $\sigma_{[1]}^2$ is called the best population. The goal is to select a non empty subset of the k populations containing the best population. Any such selection will be called a correct selection (CS).

Let $s_1^2, s_2^2, \ldots, s_k^2$ denote the sample variances. Let $s_{(i)}^2$ denote the (unknown) sample variance that is associated with the ith smallest population variance, $\sigma_{[i]}^2$; let $\nu_{(i)}$ denote the number of degrees of freedom associated with $s_{(i)}^2$. Gupta and Sobel [8] have proposed a procedure for this goal. Gupta and Huang [6] obtained a lower bound on the infimum of the correct selection. We modify Gupta-Sobel procedure to obtain exact results to satisfy P*-condition asymptotically and apply the method of Section 2 to obtain an optimal procedure.

R_3: Retain π_i in the selected subset if and only if

$$s_i^2 \leq \min_{1 \leq j \leq k} \left[\left(\frac{1}{c_3}\right)^{\left(\frac{1}{\nu_i} + \frac{1}{\nu_j}\right)^{\frac{1}{2}}} s_j^2\right],$$

where c_3, $(0 < c_3 < 1)$, is the largest value satisfying the basic P*-condition.

We shall show how large sample theory can be used to find very good approximations to the required probabilities even for relatively small n. Our principal tools will be the use of the transformation $y = \log_e s^2$ (see [1]), and the approach of certain multivariate distributions to multivariate normal distributions.

Let $X_i = \log_e \frac{s_i^2}{\sigma_i^2}$, $i = 1,2,\ldots,k$. It is known (see [8]) that the expectation and variances are

$$EX_i = -\left(\frac{1}{\nu_i} + \frac{1}{3\nu_i^2}\right) + O(\nu_i^{-3}) \sim -\frac{1}{\nu_i}$$

$$\mathrm{Var}(X_i) = \frac{d^2}{dx^2}\left[\log_e \Gamma(x)\right]\Big|_{x=\frac{\nu_i}{2}}$$

$$= \frac{2}{\nu_i} + \frac{2}{\nu_i^2} + \frac{4}{3\nu_i^3} + O(\nu_i^{-5}) \sim \frac{2}{\nu_i},\quad i = 1,2,\ldots,k,$$

and $$E(X_i - X_j) \sim \frac{1}{\nu_j} - \frac{1}{\nu_i},$$

$$\mathrm{Var}(X_i - X_j) \sim \frac{2}{\nu_i} + \frac{2}{\nu_j}, \quad \text{for } j = 1,2,\ldots,k;\ j \neq i.$$

Thus

$$\left(\frac{\nu_i}{2}\right)^{\frac{1}{2}} \log \frac{s_i^2}{\sigma_i^2}$$

is asymptotically distributed as a standard normal variable Z_i as $\nu_i \to \infty$.

Since

$$\frac{s_i^2}{\sigma_i^2} \le \left(\frac{1}{c_3}\right)^{\left(\frac{1}{\nu_i} + \frac{1}{\nu_j}\right)^{\frac{1}{2}}} \frac{s_j^2}{\sigma_j^2}, \quad j = 1,2,\ldots,k,\ j \neq i$$

is equivalent to

$$Z_{ij} \le \frac{1}{\sqrt{2}} \log \frac{1}{c_3} + \frac{\frac{1}{\nu_i} - \frac{1}{\nu_j}}{\sqrt{\frac{2}{\nu_i} + \frac{2}{\nu_j}}}, \quad j = 1,2,\ldots,k,\ j \neq i,$$

where $Z_{ij} = (X_i - X_j + \frac{1}{\nu_i} - \frac{1}{\nu_j})(\frac{2}{\nu_i} + \frac{2}{\nu_j})^{-\frac{1}{2}}$, for $1 \le i, j \le k$, $i \neq j$, and

$$\kappa_{ij}^k = \frac{1}{\sqrt{\left(1 + \frac{\nu_{[k]}}{\nu_{[i]}}\right)\left(1 + \frac{\nu_{[k]}}{\nu_{[j]}}\right)}} = r_{ik}\, r_{jk}, \quad i,j = 1,\ldots,k-1,\ i \neq j.$$

Hence we can apply the results in Section 3 to prove

$$\inf P(CS|R_3) \approx \int_{-\infty}^{\infty} \prod_{j=1}^{k-1} \Phi\left(\frac{c_{kj} - r_{jk}u}{(1 - r_{jk}^2)^{\frac{1}{2}}}\right) d\Phi(u),$$

where

$$c_{kj} = \frac{1}{\sqrt{2}} \log \frac{1}{c_3} + \frac{\frac{1}{\nu_{[k]}} - \frac{1}{\nu_{[j]}}}{\sqrt{\frac{1}{\nu_{[k]}} + \frac{2}{\nu_{[j]}}}}, \quad j = 1,\ldots,k-1.$$

It should be pointed out that for equal sample size case, Gupta and Sobel [8] compared the exact value and the asymptotic values of the constant c_3 to see how close they are.

For any i, $1 \le i \le k$, let $\tau_{ij} = \frac{\sigma_j^2}{\sigma_i^2}$ and $T_{ij} = \frac{s_j^2}{s_i^2}$ for $1 \le j \le k$, $j \neq i$. We can find the joint density of T_{ij}, $j = 1,2,\ldots,k$, $j \neq i$. We can construct an optimal procedure based on T_{ij}'s using the method of Section 2.

5. Selection Procedures in Two-Way Layouts

Let $\pi_1, \pi_2, \ldots, \pi_k$ be k populations. For a two-factor complete block design with one observation per cell,we express the observable random variables $X_{i\alpha}$ ($i = 1,2,\ldots,k$; $\alpha = 1,\ldots,n$) as

$$(5.1) \qquad X_{i\alpha} = \mu + \beta_\alpha + \tau_i + \xi_{i\alpha}, \quad \sum_{i=1}^{k} \tau_i = 0,$$

where μ is the mean-effect, $\beta_1,\ldots,\beta_\alpha$ are the block effects (nuisance parameters for the fixed effects model or random variables for the mixed effects model), $\tau_1,\ldots,\tau_k$ are the treatment effects, and the $\xi_{i\alpha}$ are the error components. Let $X_{i1},\ldots,X_{in}$ denote the n independent observations from the ith population π_i. Let the joint density of $X_{11},\ldots,X_{1n}$; $X_{21},\ldots,X_{2k}$;...; $X_{k1},\ldots,X_{kn}$ be of the following form:

$$(5.2) \qquad c_k|\Lambda|^{-\frac{1}{2}} g((\underline{x}-\underline{\theta})'\Lambda^{-1}(\underline{x}-\underline{\theta}))$$

where $\underline{x}' = (x_{11},\ldots,x_{1n};\ldots;x_{k1},\ldots,x_{kn})$, and $\underline{\theta}' = (\theta_{11},\ldots,\theta_{1n};\ldots;\theta_{k1},\ldots,\theta_{kn})$, $\theta_{i\alpha} = \mu + \beta_\alpha + \tau_i$, $i = 1,2,\ldots,k$; $\alpha = 1,\ldots,n$, and Λ is a known positive definite matrix, c_k is determined such that (5.2) is a density.

Our purpose is to study some selection procedures to select a subset of a random size containing the "best" treatment. The quality of the treatment is judge by the largeness of the τ_i's.

Let $\tau_{[1]} \leq \ldots \leq \tau_{[k]}$ be the actual ranked τ's (which are unknown), and let

$$Z_i = X_i - \bar{X} \text{ where } X_i = \frac{1}{n}\sum_{j=1}^{n} X_{ij},\ i = 1,\ldots,k; \text{ and } \bar{X} = \frac{1}{k}\sum_{i=1}^{k} X_i.$$

We denote the ordered values of the Z_i's by $Z_{[1]} \leq \ldots \leq Z_{[k]}$ and let $Z_{(i)}$ be the random variable associated with $\tau_{[i]}$, $i = 1,\ldots,k$.

By a similar argument as in Section 3, we know that Z_i is the generalized least square estimator of τ_i.

Let $\underset{\sim}{Z} = E\underset{\sim}{Y}$, where $E' = (E_1,\ldots,E_k)_{knxk}$ with rank k, $\underset{\sim}{Y}' = (X_{11},\ldots,X_{1n};\ldots; X_{k1},\ldots,X_{kn})$, and $E_i' = (-\frac{1}{kn},\ldots,-\frac{1}{kn};\ldots;\frac{1}{n}(1-\frac{1}{k}),\ldots,\frac{1}{n}(1-\frac{1}{k});\ldots;-\frac{1}{kn},\ldots,-\frac{1}{kn})$, for $1 \leq i \leq k$.

Then the joint density of $Z_1,\ldots,Z_k$ is of the form:

$$b_k|\Sigma|^{-\frac{1}{2}} h((\underset{\sim}{z}-\underset{\sim}{\tau})'\Sigma^{-1}(\underset{\sim}{z}-\underset{\sim}{\tau})) \tag{5.3}$$

where $\underset{\sim}{z}' = (z_1,\ldots,z_k)$, $\underset{\sim}{\tau}' = (\tau_1,\ldots,\tau_k)$ and $\underset{kxk}{\Sigma} = E \wedge E' = (\sigma_{ij})$.

The methods to construct selection procedures are the same as in Section 3.

References

[1] Bartlett, M. S. and Kendall, D. G. (1946). The statistical analysis of variance - heterogeneity and the logarithmic transformation, supplement to the journal of the Royal Statistical Society 8, Part I, 128-38.

[2] Chen, H. J., Dudewicz, E. J. and Lee, Y. J. (1975). Subset selection procedures for normal means under unequal sample sizes. Mimeograph Series No. 75-5, Dept. of Math. Sciences, Memphis State Univ., Tenn.

[3] Das Gupta, S., Eaton, M. L., Olkin, I., Perlman, M., Savage, L. J. and Sobel, M. (1970). Inequalities on the probability content of convex regions for elliptically contoured distributions. Proc. of the Sixth Berkeley Symposium on Mathematical Statistics and Probability.

[4] Dudewicz, E. J. (1974). A note on selection procedures with unequal observation numbers. Zastosowania Mathematyki 14, No. 1, 31-35.

[5] Gupta, S. S. and Huang, W. T. (1974). A note on selecting a subset of normal populations with unequal sample sizes. Sankhyā, Ser. A, 36, 389-396.

[6] Gupta, S. S. and Huang, D. Y. (1976). Subset selection procedures for the means and variances of normal populations: unequal sample sizes case. To appear in Sankhyā.

[7] Gupta, S. S., Nagel, K. and Panchapakesan, S. (1973). On the order statistics from equally correlated normal variables. Biometrika 60, 403-413.

[8] Gupta, S. S. and Sobel, M. (1962). On selecting a subset containing the population with the smallest variance. Biometrika 49, 495-507.

[9] Kelker, D. (1968). Distribution theory of spherical distributions and some characterization theorems. Michigan State University, Technical Report Rm-210, DK-1.

[10] Marshall, A. W. and Olkin, I. (1974). Majorization in multivariate distributions. Ann. Math. Statist. 2, 1189-1200.

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER Mimeograph Series #458	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) Some Multiple Decision Problems in Analysis of Variance		5. TYPE OF REPORT & PERIOD COVERED Technical
		6. PERFORMING ORG. REPORT NUMBER Mimeo. Series #458
7. AUTHOR(s) Shanti S. Gupta and D. Y. Huang		8. CONTRACT OR GRANT NUMBER(s) N00014-67-A-0226 and N00014-75-C-0455
9. PERFORMING ORGANIZATION NAME AND ADDRESS Purdue University Deparmtent of Statistics W. Lafayette, IN 47907		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS NR 042-243
11. CONTROLLING OFFICE NAME AND ADDRESS Office of Naval Research Washington, DC		12. REPORT DATE July 1976
		13. NUMBER OF PAGES 20
14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office)		15. SECURITY CLASS. (of this report) Unclassified
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release, distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Locally best, selection procedures, correct selection, generalized LS estimates, Schur-concave functions, unequal sample sizes.

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

In most 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, in one-way ANOVA the hypothesis is rejected in actual application of the F-test, the resulting conclusion that the true means $\theta_1,\ldots,\theta_k$ are not all equal, would by itself usually be insufficient to satisfy the experimenter. In fact his problems would begin at this stage. The experimenter may desire to select the "best" population or a subset of the "good" populations; 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.

The extensive literature on selection and ranking procedures depends heavily on the use of independence between populations (block, treatments, etc.) in the analysis of variance. In practical applications, it is desirable to drop this assumption of independence and consider cases more general than the normal.

In the present paper, we derive a method to construct locally best (in some sense) selection procedures to select a non empty subset of the k populations containing the best population as ranked in terms of θ_i's which control the size of the selected subset and which maximizes the probability of selecting the best. We also consider the usual selection procedures in one-way ANOVA based on the generalized least squares estimates and apply the method to two-way layout case. Some examples are discussed and some results on comparisons with other procedures are also obtained.