HC - 3.00
MF - .65

Code-1
CR-81122
pages-18
Cat 19

N67-18011

Selection From Multivariate Normal Populations

by

Khursheed Alam and M. Haseeb Rizvi *

Indiana University and Ohio State University

0. Summary. This paper is concerned with the problems of selection and ranking of k non-central chi-squared and non-central F populations, defined in terms of their non-centrality parameters. We are interested in selecting the t largest of the k populations and a subset containing the t largest for which two procedures, named R_1 and R_2 are given. It is required that the probability of a correct selection using these procedures should be at least as large as any given number $P^* < 1$. We call this the "P^* condition." The main part of the problem is to determine the least favorable configurations of the parameter space for which the probability of a correct selection is minimum. The expression for the minimum value determines the smallest sample size needed to satisfy the P^* condition. The least favorable configurations and the corresponding expressions for the minimum of the probability of a correct selection are obtained for R_1 and R_2.

The selection procedures R_1 and R_2 suggest themselves naturally. Some operating characteristics of these procedures dealing with a stochastically ordered family of populations are shown.

* This author's work was supported in part by National Aeronautical and Space Administration Grant No. NGR 36-008-040 at The Ohio State University.

The ranking of k multivariate normal populations with mean column vectors μ_i and covariance matrices Σ_i $(i = 1,\ldots,k)$ in terms of the Mahalanobis [10] distance function $\theta_i = \mu_i' \Sigma_i^{-1} \mu_i$ reduces to ranking (with respect to the non-centrality parameters) the non-central chi-squared or non-central F populations. This parametric distance function has wide applications in multivariate analysis.

1. Introduction. Bechhofer [3] used R_1 (to be described below) to rank the means of several normal populations with known variances. Gupta [4] used R_2 (to be described below) to select a subset of the given normal populations containing the largest mean. These procedures have been also used for selection from binomial and some other populations. Hall [7] has shown some optimal properties of these rules.

Let $\pi_1,\ldots,\pi_k$ be $k \geq 2$ given populations which can be ordered by a real-valued parameter θ. Precisely, each population generates a random variable X having a (cumulative) distribution function $H(x,\theta)$ which, we assume, is non-increasing in θ for constant x. When this assumption holds we say that the given populations belong to a "stochastically ordered" family. Denote by θ_i the value of θ for π_i; $i = 1,\ldots,k$. We say that π_i is larger than π_j if $\theta_i > \theta_j$. It is assumed that no a priori information is available regarding the relative values of the θ_i's.

Two problems of selection are considered, which we propose to call Problem I and Problem II. In Problem I it is required to select the t largest of the k populations where $1 \le t < k$. In Problem II it is required to select a subset of the k populations which contains the t largest populations. In both the problems it is further required that the probability of a correct selection is not smaller than a pre-assigned quantity P^*, $1/\binom{k}{t} < P^* < 1$. We shall call this the "P^* condition". With regard to Problem II, it will be observed that the P^* condition can be satisfied by including all the populations in the selected subset. Therefore, any selection procedure that we may consider should be such that the size of the selected subset or its expected value, in case it is a random variable, is less than k.

If $\theta_i = \theta_j$, then π_i is not considered distinct from π_j, that is to say, the ranks of π_i and π_j can be interchanged. However, in the limiting case in which all θ_i's are equal and which is considered for evaluating the infimum of the probability of a correct selection, the definition of a correct selection is modified to mean the selection of a set of "tagged" populations.

Let x_i denote a real-valued observation or the value of such a statistic based on several observations taken from π_i. Order the k populations according to the values of x_i's. For Problem I, R_1 selects the t largest populations according to this ordering. In case of a tie between several populations for a given rank the selection between the competing members may be made by any random procedure not depending on the observations. If the distributions involved are continuous the probability of the occurrence of a tie is zero.

We denote the ordered values of a set of k numbers by using square brackets around the subscript. Thus, $x_{[i]}$ denotes the ith smallest number in the set $\{x_1,\ldots,x_k\}$. For Problem II, R_2 selects a subset of the k populations such that π_i is retained in the subset if and only if $d(x_i, x_{[k-t+1]}) \leq \epsilon$ where ϵ is a positive number and the function d represents a measure of distance. Two such functions will be considered, namely d_1 and d_2 given by $d_1(y,z) = z - y$ and $d_2(y,z) = z/y$. The value of ϵ is determined by the P^* condition. Clearly, the probability of a correct selection as well as the size of the selected subset tend to increase with ϵ.

Let $\underset{\sim}{\theta}$ denote the vector $(\theta_1,\ldots,\theta_k)$ and Ω the space of all admissible values of $\underset{\sim}{\theta}$. For example, Ω may be the k-dimensional Euclidean space or the sub-space $\{\underset{\sim}{\theta}\colon \theta_i \geq 0,\ i = 1,\ldots,k\}$. We denote by $P_1(\underset{\sim}{\theta})$ the probability of a correct selection for R_1 and by $P_2(\underset{\sim}{\theta})$ the probability of a correct selection for R_2 when $\underset{\sim}{\theta}$ is the unknown parameter.

Then

$$(1.1) \qquad P_1(\underset{\sim}{\theta}) = \sum_{j\in J} \int \prod_{u\in I, v\in J, v\neq j} H(x,\theta_u)\,[1 - H(x,\theta_v)]\, d\,H(x,\theta_j),$$

and $P_2(\underset{\sim}{\theta})$ is the coefficient of y^{t-1} in the expression

$$(1.2) \qquad \sum_{j\in J} \int \prod_{u\in I,\; v\in J,\; v\neq j} \{H(x,\theta_u) + y[1 - H(x,\theta_u)]\}$$

$$\{H(x,\theta_v) - H(x^o,\theta_v) + y[1 - H(x,\theta_v)]\}\, dH(x,\theta_j)$$

$$+ \sum_{i\in I} \int \prod_{u\in I,\; u\neq i,\; v\in J} \{H(x,\theta_u) + y[1 - H(x,\theta_u)]\}$$

$$\{H(x,\theta_v) - H(x^o,\theta_v) + y[1 - H(x,\theta_v)]\}\, d\,H(x,\theta_i),$$

where I denotes the set $\{[1],\ldots,[k-t]\}$, J the set $\{[k-t+1],\ldots,[k]\}$ and $x^o = d(\varepsilon,x)$.

For a stochastically ordered family of populations it is shown in section 2 that $P_1(\underset{\sim}{\theta})$ is non-increasing in each of the components $\theta_{[i]}$, $i = 1,\ldots,k-t$, and non-decreasing in each of the components $\theta_{[j]}$, $j = k-t+1,\ldots,k$. Therefore,

$$\inf_{\Omega} P_1(\underset{\sim}{\theta}) = P_1(\underset{\sim}{\theta}_o) = 1/\binom{k}{t},$$

where $\underset{\sim}{\theta}_o$ denotes any vector point in Ω whose components are all equal. Thus the P^* condition may be satisfied only on a subset of Ω which may be termed a "preference" zone. One such subset which we consider for the multivariate normal problem to be described below is $\Omega_3 = \Omega_1 \cap \Omega_2$, where

$$\Omega_1 = \{\underset{\sim}{\theta} \in \Omega:\ d_1(\theta_{[k-t]}, \theta_{[k-t+1]}) \geq \delta_1\},$$
$$\Omega_2 = \{\underset{\sim}{\theta} \in \Omega:\ d_2(\theta_{[k-t]}, \theta_{[k-t+1]}) \geq \delta_2\},$$

for some $\delta_1 > 0$, $\delta_2 > 1$. Such a preference zone has been considered by Sobel [11] for ranking Poisson populations.

For R_2 it is shown in section 2 that $P_2(\underset{\sim}{\theta})$ is non-increasing in $\theta_{[1]},\ldots,\theta_{[k-t]}$. It follows that for $t = 1$,

$$\inf_{\Omega} P_2(\underset{\sim}{\theta}) = \inf_{\Omega_o} P_2(\underset{\sim}{\theta}_o)$$

where Ω_o $(\subset \Omega)$ is the set of all points $\underset{\sim}{\theta}_o$.

Consider an application of these procedures to the multivariate normal populations. Let π_i represent a multivariate normal population with mean μ_i and covariance Σ_i, where μ_i is a column vector of p components and Σ_i is a positive definite $p \times p$ matrix, $i = 1,\ldots,k$. We rank the k populations according to the values of the parametric functions $\theta_i = \mu_i' \Sigma_i^{-1} \mu_i$, where μ_i' is the transpose of μ_i. Then π_i is called larger than π_j if $\mu_i' \Sigma_i^{-1} \mu_i > \mu_j' \Sigma_j^{-1} \mu_j$.

Suppose that a sample of size n_i is drawn from π_i. Denote the ith sample vector mean and covariance matrix by $\bar{x}_i$ and S_i respectively; these are maximum likelihood estimates respectively of μ_i and Σ_i. Let $u_i = \bar{x}_i' \Sigma_i^{-1} \bar{x}_i$ and $v_i = (\bar{x}_i' S_i^{-1} \bar{x}_i) . \dfrac{(n_i - p)}{n_i p}$, then $n_i U_i$ has the distribution of a non-central chi-squared random variable with p degrees of freedom and non-centrality parameter $n_i \mu_i' \Sigma_i^{-1} \mu_i$ and $n_i V_i$ has the non-central F distribution with p and $(n_i - p)$ degrees of freedom and non-centrality parameter $n_i \mu_i' \Sigma_i^{-1} \mu_i$ (see [1] pp. 113-114). Two cases may arise, according as the population covariance matrices are supposed to be known or unknown. In the first case we use u_i for x_i and carry out the procedures R_1 and R_2 as described above. In the second case we use v_i for x_i. Detailed analysis is given for the first case only.

We shall consider only the case when the sample size is same for each population, that is, $n_i = n$, say, for all i. From the point of view of the design of experiments the equality of the n_i's is suggested by the invariance of the problem and of the selection procedures under permutation of the labels of the populations. An expression is obtained giving the smallest value of n required to satisfy the P^* condition for the procedure R_1. Similar expression for the smallest value of ϵ required to satisfy the P^* condition is obtained for R_2 when $t = 1$.

Let S denote the size of the selected subset in Problem II. S is a random variable for R_2; denote its expected value by $E(S)$. Then $E(S)$ may be taken as a criterion for the suitability of the procedure R_2. An expression for $E(S)$ is given in (5.2). It is shown that

$$\operatorname{Sup}_{\Omega} E(S) = \operatorname{Sup}_{\Omega_0} E(S) = k$$

To carry out the procedure R_1 for a given P^* one needs to know the smallest value of n satisfying the P^* condition.

This is determined from equation (3.6) or (3.7). The values of n are under tabulation and will be published shortly.

For the procedure R_2 one needs to know the smallest value of ε satisfying the P^* condition. For $t = 1$ this is obtained from equation (4.2). Tables in Gupta [5] and Armitage and Krishnaiah [2] provide solutions of ε in some cases when the covariance matrices are known.

2. Operating Characteristics of R_1 and R_2. A few results on the minimization of $P_1(\underset{\sim}{\theta})$ and $P_2(\underset{\sim}{\theta})$ follow from the following lemma.

Lemma 2.1.** Let $X = (X_1,\ldots,X_k)$ be a vector-valued random variable of $k \geq 1$ independent components such that for each i the random variable X_i has the distribution function $H(x_i;\ \theta_i)$, which is non-increasing in θ_i for constant x_i, $i = 1, \ldots, k$. If $\psi(x)$ is a monotone function of x_i when the other components are fixed then $E\,\psi(X)$ is monotone in θ_i in the same direction.

Proof: For $k = 1$, see proof in [8]. For $k > 1$, suppose that $\psi(x)$ is non-decreasing in x_i. Let $\underset{\sim}{\theta}^* = (\theta_1,\ldots,\theta_{i-1},\theta_i^*,\ \theta_{i+1},\ldots,\theta_k)$, where $\theta_i^* \geq \theta_i$. Denoting by E_i the expectation with respect to X_i we get

$$\begin{aligned} E\{\psi(X);\underset{\sim}{\theta}\} &= E\ E_i\{\psi(X);\ x_1,\ldots,x_{i-1},x_{i+1},\ldots,x_k;\ \underset{\sim}{\theta}\} \\ &\leq E\ E_i\{\psi(X);x_1,\ldots,x_{i-1},x_{i+1},\ldots,x_k,\underset{\sim}{\theta}^*\} \\ &= E\ \{\psi(X);\ \underset{\sim}{\theta}^*\}. \end{aligned}$$

The case when $\psi(x)$ is non-increasing in x_i can be treated similarly. This completes the proof of the lemma.

** While this paper was in the process of publication, the authors learnt that Desu M. Mahamunulu had obtained a similar result in his paper "On a generalized goal in fixed-sample ranking and selection problems", Technical Report No. 72, Dept. of Statistics, Univ. of Minnesota, 1966.

Let us denote by $X_{(i)}$ the random variable of the ith smallest population. Note that the $X_{(i)}$'s are unknown quantities. Let

$$\psi(X) = 1 \quad \text{if } \max(X_{(1)},\dots,X_{(k-t)}) \le \min(X_{(k-t+1)}, \dots, X_{(k)})$$
$$= 0 \text{ , otherwise.}$$

Then $\psi(x)$ is non-increasing in $x_{(i)}$ for $i = 1,\dots,k-t$ and non-decreasing in $x_{(j)}$ for $j = k-t+1,\dots,k$ and $P_1(\underset{\sim}{\theta}) = E\,\psi(X)$. Therefore, by Lemma 2.1, $P_1(\underset{\sim}{\theta})$ is non-increasing in $\theta_{[i]}$ for $i = 1,\dots,k-t$ and non-decreasing in $\theta_{[j]}$ for $j = k-t+1,\dots, k$.

Similarly, define

$$\varphi(X) = 1 \quad \text{if } d(X_{(j)},X_{[k-t+1]}) \le \epsilon, j = k-t+1,\dots,k$$
$$= 0 \text{ otherwise.}$$

Then $\varphi(x)$ is non-increasing in $x_{(i)}$ for $i = 1,\dots,k-t$ and $P_2(\underset{\sim}{\theta}) = E\ \varphi(X)$. Therefore, by the above lemma $P_2(\underset{\sim}{\theta})$ is non-increasing in $\theta_{[i]}$ for $i = 1,\dots,k-t$. Thus we have

<u>Theorem 2.1</u>. For a stochastically ordered family of populations $P_1(\underset{\sim}{\theta})$ is non-increasing in $\theta_{[i]}$ for $i = 1,\dots,k-t$ and non-decreasing in $\theta_{[j]}$ for $j = k-t+1,\dots, k$; also $P_2(\underset{\sim}{\theta})$ is non-increasing in $\theta_{[i]}$ for $i = 1,\dots,k-t$.

<u>Corollary 2.1</u>. $\inf_{\Omega} P_1(\underset{\sim}{\theta}) = \inf_{\Omega_o} P_1(\underset{\sim}{\theta}_o) = 1/\binom{k}{t}$. For $t = 1$, $\inf_{\Omega} P_2(\underset{\sim}{\theta}) = \inf_{\Omega_o} P_2(\underset{\sim}{\theta}_o)$.

For a fixed i let p_i denote the probability that π_i is included in the subset selected by R_2. Then $p_i = E\,\eta(X)$, where

$$\eta(X) = 1 \quad \text{if } \pi_i \text{ is included in the subset selected by } R_2$$
$$= 0, \text{ otherwise.}$$

Clearly, $\eta(x)$ is non-decreasing in x_i and, therefore, by Lemma 2.1, p_i is non-decreasing in θ_i. Thus, a desirable characteristic of the procedure R_2 for any stochastically ordered family of populations is given by

Theorem 2.2. $p_i \geq p_j$ for $\theta_i \geq \theta_j$.

3. Problem I (Normal). Consider the problem (described in section 1) of selecting the t largest of k multivariate normal populations. First we suppose that the population covariance matrices are known. In this case we use u_i's to rank the populations. By Corollary 2.1 the P^* condition cannot be satisfied over Ω, the set of all k-dimensional vectors with non-negative components. Consider the infimum of $P_1(\underset{\sim}{\theta})$ over the subset Ω_1. Applying Theorem 2.1, we obtain

$$\inf_{\Omega_1} P_1(\underset{\sim}{\theta}) = \inf_{\theta \geq 0} t \int_0^{\infty} F_p^{k-t}(x,\theta) \{1-F_p(x,\theta + n\delta_1)\}^{t-1} f_p(x,\theta + n\delta_1)dx, \tag{3.1}$$

where $f_p(x,\theta)$ and $F_p(x,\theta)$ denote the density function and the distribution function, respectively, of the non-central chi-squared random variable with p degrees of freedom and non-centrality parameter θ. These functions can be written as (see [9], p. 312)

$$f_p(x,\theta) = e^{-\theta/2} \sum_{r=0}^{\infty} \frac{\theta^r}{2^r \; r!} f_{p+2r}(x), \; x > 0, \; \theta \geq 0$$

and

$$F_p(x,\theta) = e^{-\theta/2} \sum_{r=0}^{\infty} \frac{\theta^r}{2^r r \; !} F_{p+2r}(x), \quad x > 0, \; \theta \geq 0$$

where $f_\gamma(x) = \dfrac{x^{(\gamma/2)-1} e^{-x/2}}{2^{\gamma/2} \; \Gamma(\gamma/2)}$ represents the density function of a central chi-squared variable with γ degrees of freedom, and

$$F_\gamma(x) = \int_0^x f_\gamma(y)dy.$$

It is easily verified that

$$(3.2) \qquad 2\frac{\partial}{\partial\theta} f_p(x,\theta) = f_{p+2}(x,\theta) - f_p(x,\theta),$$

$$(3.3) \qquad 2\frac{\partial}{\partial\theta} F_p(x,\theta) = F_{p+2}(x,\theta) - F_p(x,\theta) = -2f_{p+2}(x,\theta).$$

Let A denote the integral on the right side of (3.1). Differentiating with respect to θ and making use of (3.2) and (3.3) we have

$$\frac{\partial A}{\partial\theta} = -(k-t) \int_0^\infty F_p^{k-t-1} (x,\theta) \; \{1-F_p(x,\theta+n \; \delta_1)\}^{t-1}$$

$$f_p(x,\theta+n \; \delta_1) f_{p+2}(x,\theta)dx$$

$$+ \; (t-1) \int_0^\infty F_p^{k-t} (x,\theta) \; \{1-F_p(x,\theta+n \; \delta_1)\}^{t-2}$$

$$f_{p+2}(x,\theta+n \; \delta_1) f_p(x,\theta+n \; \delta_1)dx \; +$$

$$\frac{1}{2} \int_0^\infty F_p^{k-t}(x,\theta)\ \{1-F_p(x,\theta+n\ \delta_1)\}^{t-1}$$

$$\{f_{p+2}(x,\theta+n\ \delta_1) - f_p(x,\theta+n\ \delta_1)\}\ dx.$$

Integrating by parts the third integral on the right in the above equation we have

$$\frac{\partial A}{\partial \theta} = (k-t) \int_0^\infty F_p^{k-t-1}(x,\theta)\ \{1-F_p(x,\theta+n\ \delta_1)\}^{t-1}$$

$$\{f_{p+2}(x,\theta+n\ \delta_1)\ \ f_p(x,\theta) - f_{p+2}(x,\theta) f_p(x,\theta+n\ \delta_1)\} dx.$$

By Lemma 3.1 given at the end of this section, $f_{p+2}(x,\theta)/f_p(x,\theta)$ is non-increasing in θ. Hence, $\frac{\partial A}{\partial \theta} \leq 0$.

Next consider the infimum of $P_1(\underset{\sim}{\theta})$ over Ω_2. Like (3.1) we obtain

$$(3.4) \qquad \inf_{\Omega_2} P_1(\underset{\sim}{\theta}) = \inf_{\theta \geq 0} t \int_0^\infty F_p^{\ k-t}(x,\theta) \{1-F_p(x,\delta_2\ \theta)\}^{t-1} f_p(x,\delta_2\theta) dx.$$

Denote the integral on the right side of (3.4) by B. Differentiating with respect to θ we have

$$\frac{\partial B}{\partial \theta} = (k-t) \int_0^\infty F_p^{\ k-t-1}(x,\theta)\ \{1 - F_p(x,\delta_2\theta)\}^{t-1}$$

$$[\delta_2\ f_p(x,\theta) f_{p+2}(x,\delta_2\theta) - f_{p+2}(x,\theta) f_p(x,\delta_2\theta)]\ dx$$

By the help of Lemma 3.1 it can be shown that the quantity inside the square brackets above is non-negative.

Since $\frac{\partial A}{\partial \theta} \leq 0$ and $\frac{\partial B}{\partial \theta} \geq 0$, we conclude that $P_1(\underset{\sim}{\theta})$ is minimized on Ω_3 at the vector point $\underset{\sim}{\lambda}$ whose components are given by

$$(3.5) \qquad \lambda_{[i]} = \delta_1/(\delta_2 - 1), \quad i = 1,\ldots,k-t$$
$$= \delta_1 \delta_2/(\delta_2 - 1), \; i = k-t+1,\ldots,k,$$

and the smallest n required to satisfy the P^* condition of the problem is obtained from the equation

$$(3.6) \qquad \inf_{\Omega_3} P_1(\underset{\sim}{\theta}) = t \int_0^\infty F_p^{k-t}(x, \frac{n\delta_1}{\delta_2 - 1}) \{1 - F_p(x, \frac{n\delta_1\delta_2}{\delta_2 - 1})\}^{t-1} f_p(x, \frac{n\delta_1\delta_2}{\delta_2 - 1}) dx = P^* .$$

Similarly in the second case where the population covariance matrices are unknown, using v_i's to rank the populations, the probability of a correct selection is again minimized at $\underset{\sim}{\lambda}$ in Ω_3. The smallest value of n required to satisfy the P^* condition in this case is obtained from the equation

$$(3.7) \qquad \inf_{\Omega_3} P_1(\underset{\sim}{\theta}) = t \int_0^\infty G_{p,n-p}^{k-t}(x, \frac{n\delta_1}{\delta_2 - 1}) \{1 - G_{p,n-p}(x, \frac{n\delta_1\delta_2}{\delta_2 - 1})\}^{t-1} g_{p,n-p}(x, \frac{n\delta_1\delta_2}{\delta_2 - 1}) dx = P^* ,$$

where $g_{p,q}(x,\theta)$ and $G_{p,q}(x,\theta)$ denote the density function and the distribution function respectively of the ratio of a non-central chi-squared variable with p degrees of freedom and non-centrality parameter θ and an independent central chi-squared variable with q degrees of freedom. These functions can be written as (see [1], p. 114)

$$g_{p,q}(x,\theta) = \frac{e^{-\theta/2}}{\Gamma(\frac{q}{2})} \sum_{r=o}^{\infty} \frac{x^{(p/2)+r-1}\Gamma(\frac{p}{2}+\frac{q}{2}+r)}{(1+x)^{\frac{p}{2}+\frac{q}{2}+r}\Gamma(\frac{p}{2}+r)} \cdot \frac{\theta^r}{2^r r!} \quad , x > 0 ,$$

$$G_{p,q}(x,\theta) = \frac{e^{-\theta/2}}{\Gamma(\frac{q}{2})} \sum_{r=o}^{\infty} \frac{\theta^r}{2^r r!} \int_0^x \frac{x^{(p/2)+r-1}\ \Gamma(\frac{p}{2}+\frac{q}{2}+r)}{(1+x)^{\frac{p}{2}+\frac{q}{2}+r}\ \Gamma(\frac{p}{2}+r)} dx, x > 0.$$

To derive the equation (3.7) we use the following relations, which are easily verified.

(3.8) $$2\frac{\partial}{\partial\theta} g_{p,q}(x,\theta) = g_{p+2,q}(x,\theta) - g_{p,q}(x,\theta),$$

(3.9) $$2\frac{\partial}{\partial\theta} G_{p,q}(x,\theta) = G_{p+2,q}(x,\theta) - G_{p,q}(x,\theta)$$

$$= -\frac{e^{-\theta/2}}{\Gamma(\frac{q}{2})} \sum_{r=o}^{\infty} \frac{\theta^r}{2^r r!} \int_0^x \frac{x^{\frac{p}{2}+r}\ \Gamma(\frac{p}{2}+\frac{q}{2}+r+1)\, dx}{(1+x)^{\frac{p}{2}+\frac{q}{2}+r+1}\ \Gamma(\frac{p}{2}+r+1)} = -\frac{2}{q-2}\, g_{p+2,q-2}(x,\theta)$$

for $q > 2$.

Summarizing the above discussion we have

Theorem 3.1. The probability of a correct selection using the procedure R_1 is minimized on Ω_3 at the point $\underset{\sim}{\lambda}$ given by (3.5). The smallest value of n required to satisfy the P^* condition is obtained from equation (3.6) or (3.7) according as the population covariance matrices are known or unknown.

The following Lemma has been cited above (for proof see [9] p. 313).

Lemma 3.1. Let $h(z) = (\sum_{i=o}^{\infty} b_i\ z^i)/\sum_{i=o}^{\infty} a_i z^i$, where the constants a_i, b_i are ≥ 0 and $\sum a_i z^i$ and $\sum b_i z^i$ converge for all $z > 0$. If the sequence $\{b_i/a_i\}$ is monotone then $h(z)$ is a monotone function of z in the same direction.

4. <u>Problem II</u> (Normal). Consider R_2 for the problem (described in section 1) of selecting a subset containing the t largest of k multivariate normal populations. If the difference d_1 is used for the distance function d describing R_2 then it is easily seen that the probability of a correct selection approaches its minimum value $1/\binom{k}{t}$ as the parameters become large. However, using the ratio d_2 for the distance function d we have from Corollary 2.1 for $t = 1$,

$$(4.1) \qquad \inf_{\Omega} P_2(\underset{\sim}{\theta}) = \inf_{\Omega} \int_0^{\infty} \Pi_{i=1}^{k-1} H(\varepsilon x, n\theta_{[i]}) dH(x, n\theta_{[k]})$$

$$= \inf_{\theta \geq 0} \int_0^{\infty} H^{k-1}(\varepsilon x, n\theta) dH(x, n\theta) \quad ,$$

where $H(\cdot,\cdot) = F_p(\cdot,\cdot)$ or $G_{p,n-p}(\cdot,\cdot)$ according as the population covariance matrices are known or unknown and where $\varepsilon > 1$. By the help of Lemma 3.1 the last integral on the right side of (4.1) can be shown to be non-decreasing in θ. The smallest value of ε required to satisfy the P^* condition is, therefore, determined by the equation

$$(4.2) \qquad \int_0^{\infty} H^{k-1}(\varepsilon x) \, dH(x) = P^*,$$

where $H(x) = H(x,0)$ represents the central chi-squared distribution function $F_p(\cdot)$ or the central F distribution function $G_{p,n-p}(\cdot)$. For the special case $k = 2$ and the population covariance matrices known, Gupta [6] obtains (4.2) with $k=2$ and $H(\cdot) = F_p(\cdot)$; he also treats problem II for any k when covariance matrices are known but restricts his discussion to large values of p only.

5. Size of the Selected Subset. The size of the subset selected by R_2 is a random variable. Denote the size by S and its expected value by $E(S)$. Then $E(S)$ may be taken as a measure of the efficiency of the procedure R_2. Let p_i denote the probability that π_i is included in the selected subset, then

$$E(S) = \sum_{i=1}^{k} p_i \tag{5.1}$$

Suppose that $\theta_i = \theta$ for $i = 1,\ldots,m$ and $\theta_i > \theta$ for $i = m+1,\ldots,k$. Then $E(S)$ is the coefficient of y^{t-1} in the polynomial expansion of

$$\frac{1}{1-y} \int_0^\infty m \prod_{j=m+1}^{k} \{H(\varepsilon x,\theta_j) + y[1 - H(\varepsilon x,\theta_j)]\} \{H(\varepsilon x,\theta) + y[1 - H(\varepsilon x,\theta)]\}^{m-1} d\,H(x,\theta) + \frac{1}{1-y} \sum_{i=m+1}^{k} \int_0^\infty \prod_{j=m+1,\, j\neq i}^{k} \{H(\varepsilon x,\theta_j) + y[1 - H(\varepsilon x,\theta_j)]\}\{H(\varepsilon x,\theta) + y[1 - H(\varepsilon x,\theta)]\}^m \, d\,H(x,\theta_i) \tag{5.2}$$

Let $H(\cdot,\cdot) = F_p(\cdot,\cdot)$. Differentiating with respect to θ we obtain $(\frac{\partial E(S)}{\partial \theta} - m)$ as the coefficient of y^{t-1} in

$$\int_0^\infty \Big[m \sum_{i=m+1}^{k} \prod_{j=m+1,\, j\neq i}^{k} \{F_p(\varepsilon x,\theta_j) + y[1 - F_p(\varepsilon x,\theta_j)]\} \{F_p(\varepsilon x,\theta) + y[1 - F_p(\varepsilon x,\theta)]\}^{m-1} \{\varepsilon f_{p+2}(x,\theta) f_p(\varepsilon x,\theta_i) - f_{p+2}(\varepsilon x,\theta) f_p(x,\theta_i)\}$$

$$+ m(m-1) \prod_{j=m+1}^{k} \{F_p(\varepsilon x, \theta_j) + y[1 - F_p(\varepsilon x, \theta_j)]\}$$

$$\{F_p(\varepsilon x, \theta) + y[1 - F_p(\varepsilon x, \theta)]\}^{m-2} \{\varepsilon f_{p+2}(x, \theta) f_p(\varepsilon x, \theta)$$

$$- f_{p+2}(\varepsilon x, \theta)\, f_p(x, \theta)\}\Big]\, dx .$$

By the help of Lemma 3.1 it can be shown that

$$\varepsilon\, f_{p+2}(x, \theta) f_p(\varepsilon x, \theta_i) - f_{p+2}(\varepsilon x, \theta)\, f_p(x, \theta_i) \geq 0 \quad \text{for} \quad \theta_i \geq \theta.$$

Hence

$$\frac{\partial\, E(S)}{\partial \theta} \geq 0 .$$

The same result holds for $H(\cdot,\cdot) = G_{p,n-p}(\cdot,\cdot)$. Therefore,

$$\sup_{\Omega} E(S) = \sup_{\Omega_0} E(S)$$

$$= \text{coefficient of } y^{t-1} \text{ in } \Big[\frac{k}{1-y} \lim_{\theta \to \infty}$$

$$\int_0^{\infty} \{H(\varepsilon x, \theta) + y[1 - H(\varepsilon x, \theta)]\}^{k-1} dH(x, \theta)\Big]$$

$$= \text{coefficient of } y^{t-1} \text{ in } \frac{k}{1-y}$$
$$= k .$$

Thus we have

<u>Theorem 5.1</u>. $\sup_{\Omega} E(S) = k$.

REFERENCES

[1] ANDERSON, T. W. (1958). <u>An Introduction to Multivariate Statistical Analysis</u>. John Wiley and Sons, New York.

[2] ARMITAGE, J. V., and KRISHNAIAH, P. R. (1964). Tables for the studentized largest chi-square distribution and their applications. ARL 64-188, Aerospace Research Laboratories, WP-AFB, Ohio.

[3] BECHHOFER, R. E. (1954). A single-sample multiple decision procedure for ranking means of normal populations with known variances. Ann. Math. Statist. 25 16-39.

[4] GUPTA, S. S. (1956). On a decision rule for a problem in ranking means. Mimeo. Series No. 150, Inst. of Statist., University of North Carolina.

[5] GUPTA, S. S. (1963). On a selection and ranking procedure for gamma populations. Ann. Inst. Statist. Math. Tokyo 14 199-216.

[6] GUPTA, S. S. (1965). On some selection and rankings procedures for multivariate normal populations using distance functions. Mimeo. Series No. 43, Department of Statistics, Purdue University.

[7] HALL, W. J. (1959). The most-economical character of some Bechhofer and Sobel decision rules. Ann. Math. Statist. 30 964-969.

[8] LEHMANN, E. L. (1955). Ordered families of distributions. Ann. Math. Statist. 26 399 - 419.

[9] LEHMANN, E. L. (1959). Testing Statistical Hypotheses. John Wiley and Sons, New York.

[10] MAHALANOBIS, P. C. (1930). On tests and measures of group divergence. J. Asiat. Soc. Beng. 26 541-588.

[11] SOBEL, M. (1963). Single sample ranking problems with Poisson populations. Tech. Report No. 19, Dept. of Statistics, University of Minnesota.