November 1963

RANKING AND SELECTION PROBLEMS OF NORMAL POPULATIONS USING THE ABSOLUTE VALUES OF THEIR MEANS: FIXED SAMPLE SIZE CASE*

M. Haseeb Rizvi

Technical Report No. 31

University of Minnesota

Minneapolis, Minnesota

*This research was supported in part by the National Science Foundation under Grant Number G-19126 and by the U. S. Air Force under Contract No. AF33(616)-6503 while the author was at the Aerospace Research Laboratories, Wright-Patterson Air Force Base during summer 1963.

ACKNOWLEDGEMENTS

The author is deeply indebted to Professor Milton Sobel who introduced him to the field of ranking and selection problems and whose invaluable suggestions have gone a long way in shaping this work. Thanks are also due to Professor I. Richard Savage for some helpful suggestions and to Dr. Khursheed Alam for some stimulating discussions that led to the present form of Theorem 1. Great appreciation is recorded here for the generous help given by Mr. Roy C. Milton in the programming of the tables. Mr. Bale S. Gurunanjappa also helped with the tables and his efforts are gratefully acknowledged.

ABSTRACT

RANKING AND SELECTION PROBLEMS OF NORMAL POPULATIONS USING THE ABSOLUTE VALUES OF THEIR MEANS: FIXED SAMPLE SIZE CASE

Consider $k \geq 2$ normal populations $\pi_i (i=1,2,\ldots,k)$ with unknown means $\mu_i (i=1,2,\ldots,k)$ and a common unit variance; let $\theta_{[1]} \leq \theta_{[2]} \leq \ldots \leq \theta_{[k]}$ be the ordered values of $\theta_i = |\mu_i|$. This investigation is concerned with the ranking and selection problems of these normal populations according to the unknown ordering of $\theta_i (i=1,2,\ldots,k)$. The proposed procedures are based on the statistics $W_i = |\bar{X}_i| (i=1,2,\ldots,k)$, where $\bar{X}_i$ is the sample mean of a common number n of independent observations from π_i. Chapter I of this work studies some properties of W_i; in particular, its density is shown to possess a strict monotone likelihood ratio with θ_i as the parameter.

The fixed sample size "indifference zone" formulation for the problem of selecting $t(<k)$ populations with t largest θ-values is studied in Chapter II along the lines of Bechhofer (<u>Ann. Math. Statist.</u> 25 (1954) 16-39). With correct selection (CS) defined in an obvious manner, a procedure R_t is required so as to satisfy the condition $P\{CS|R_t, \theta_{[k-t+1]} - \theta_{[k-t]} \geq \delta^*\} \geq P^*$, where P^* and $\delta^* > 0$ are pre-assigned. The proposed procedure R_t ranks w_i and selects the populations with the t largest w_i as populations with t largest θ-values. Then n is determined so that R_t satisfies the probability condition. Certain bounds on $P\{CS|R_t\}$ are obtained. Tables for special cases, t=1 and t=k-1, give values of the infimum of $P\{CS|R_t\}$ for k=2(1)10 and $\lambda = n^{\frac{1}{2}}\delta^* = 0(0.1)7.0$. In addition the values of λ are given for k=2(1)10 and several P^*'s. The allocation of sample sizes is discussed for different variance set-ups. The decision rule R_t is shown to be most economical by demonstrating its minimax and admissible nature respective to a simple loss function.

The "subset" formulation for the problem of selecting a small non-empty subset containing the population with the largest parameter $\theta_{[k]}$ is considered in Chapter III along the lines of Gupta (Mimeo. Series No. 150, 1956, Inst. of Statist., Univ. of North Carolina). Any selection of a subset which contains at least one population with parameter $\theta_{[k]}$ is regarded as a correct selection (CS). Then, for a pre-assigned probability P^*, a procedure R is required so as to satisfy the condition $P\{CS|R\}\geq P^*$ regardless of true unknown θ-values. The proposed procedure R is: Retain π_i in the selected subset if and only if $w_i\geq w_{max}-d$, where $d\geq 0$ is determined subject to R satisfying the probability condition. The solution $\gamma=n^{\frac{1}{2}}d$ of the probability condition is tabulated in Bechhofer (Ann. Math. Statist. 25 (1954) 16-39). The expected size of the selected subset, regarded as a criterion of efficiency, is derived and its supremum obtained. Two secondary problems of determining n required to control the expected size of the retained subset are also treated. The effect of different variance set-ups on the problem is discussed. The related problem of selecting the population with the smallest parameter $\theta_{[1]}$ is also considered and certain directions for generalizations indicated.

Finally, for the subset formulation of ranking and selection problems of populations with monotone likelihood ratio, some theorems are proved in Chapter IV. It is shown there that the $P\{CS\}$-function is monotonic in the parameters for a certain class of procedures and that this class of procedures possesses a certain monotonicity property.

TABLE OF CONTENTS

Chapter		Page
	ACKNOWLEDGEMENTS	ii
	ABSTRACT	iii
	TABLE OF CONTENTS	v
I	INTRODUCTION	1
	1. Introductory Remarks	1
	2. Discussion of the Statistic Used	3
II	INDIFFERENCE ZONE FORMULATION	5
	1. Formal Statement of the Problem	5
	2. Proposed Procedure R_t	6
	3. Probability of Correct Selection and its Infimum	6
	4. Certain Bounds on $P\{CS\|R_t\}$	12
	5. Special Cases and Tables	13
	6. Allocation of Sample Sizes	15
	7. Some Properties of the Procedure R_t	17
III	SUBSET FORMULATION	22
	1. Formal Statement of the Problem	22
	2. Proposed Procedure R	22
	3. Probability of Correct Selection and its Infimum	23
	4. Expected Size of the Selected Subset and its Supremum	26
	5. Monotonicity Property of the Procedure R	34
	6. Different Variance Set-ups	36
	7. The Worst Population Problem	37
	8. Directions of Certain Generalizations	39

Chapter		Page
IV	SOME REMARKS ON SUBSET FORMULATION FOR POPULATIONS WITH MONOTONE LIKELIHOOD RATIO	41
	TABLES	44
	REFERENCES	50

CHAPTER I: INTRODUCTION

1. Introductory Remarks.

In recent years new techniques of ranking populations and selecting subsets of populations, based on the ordered values of unknown parameters, have been developed and their "operating characteristics" have been studied; see, for example, Bechhofer [1], Bechhofer, Dunnett and Sobel [2], Gupta [6] and Hall [10]. These procedures are formulated as multiple-decision procedures within the general framework of Wald's decision theory. It has been pointed out that they can be used as alternatives to the classical tests of homogeneity in the Analysis of Variance; a more general discussion of the philosophy and uses of these procedures can be found in Bechhofer [1].

We are interested in ranking $k \geq 2$ independent normal populations with unknown means and a common known variance, say unity, according to the unknown ordering of the absolute values of the means. Since many different functions of the original parameters could be considered, all of which lead to different problems, a little motivation for the particular choice of the absolute value of the original parameters is in order. Suppose we are interested in ranking k p-variate normal populations with vector means $\underset{\sim}{\mu}_i$ $(i=1,2,\ldots,k)$ and a common known covariance matrix Σ. Then an interesting way of ranking these k multivariate normal populations is according to the values of the parametric function $\underset{\sim}{\mu}'\Sigma^{-1}\underset{\sim}{\mu}$ when $\underset{\sim}{\mu}=\underset{\sim}{\mu}_i(i=1,2,\ldots,k)$. This parametric function has been regarded (see, for example, Mahalanobis [17]) as a measure of the distance between two multivariate normal populations, one with vector mean $\underset{\sim}{\mu}$ and covariance matrix Σ and another with vector mean $\underset{\sim}{0}=(0,0,\ldots,0)$ and the same covariance matrix Σ. When p=1 and the common variance σ^2 of the k univariate normal populations is unity, this measure of distance clearly reduces to μ^2 or equivalently $|\mu|$. To indicate an application, suppose there are k instruments each of which gives independent measurements either greater or smaller than some known true value. We might then be interested in selecting that instrument for which the absolute value of the expectation of the difference between the true and observed value is smallest.

In Chapter II we consider the fixed sample size "indifference zone" formulation along the lines of Bechhofer [1] and Bechhofer and Sobel [3] for the problem of selecting the $t(<k)$ "best" of k populations, where the best population is defined as the one whose mean has the largest absolute value. In this formulation an indifference zone in the parameter space is pre-assigned and the common number of observations needed from each population to satisfy the requirement of the procedure is then determined. It should be noted that the problem of selecting the t' worst of k populations using the same criterion is mathematically equivalent to selecting the k-t' best populations.

In Chapter III we consider the "subset" formulation for the problem of selecting a subset of k populations which contains the best population and this is done along the lines of Gupta [6] and Gupta and Sobel [8]. In this formulation the number of observations is given beforehand and the constant needed for satisfying the requirement of the procedure is then determined. The problem of selecting a subset containing the worst population is also treated.

Chapter IV gives theorems dealing with the subset formulation of the ranking and selection problems of populations having "monotone likelihood ratio" (for definition of monotone likelihood ratio see, for example, Lehmann [15, p. 68]).

2. Discussion of the Statistic Used.

The procedures proposed in Chapters II and III are based on the absolute value of the sample mean. This statistic, besides having strong intuitive recommendation for our problems, has many desirable properties that lend themselves to proving certain optimal characteristics of the proposed procedures.

Let $X_1, X_2, \ldots, X_n$ be a sample of n independent observations from a normal population with mean $\mu(-\infty<\mu<\infty)$ and variance unity. Let $W=|\overline{X}|$ where $\overline{X}=\sum_{i=1}^{n} X_i/n$. Then the cumulative distribution function (c.d.f.) of W is

$$(2.1) \qquad H(w,\mu)=\begin{cases} F(n^{\frac{1}{2}}(w-\mu))-F(n^{\frac{1}{2}}(-w-\mu)), & w>0 \\ 0, & \text{otherwise} \end{cases}$$

and the probability density function (p.d.f.) of W is

$$(2.2) \qquad h(w,\mu)=\begin{cases} n^{\frac{1}{2}}\{f(n^{\frac{1}{2}}(w-\mu))+f(n^{\frac{1}{2}}(w+\mu))\}, & w>0 \\ 0, & \text{otherwise} \end{cases}$$

where $F(x)=\int_{-\infty}^{x} f(u)du$ and $f(u)=(2\pi)^{-\frac{1}{2}}e^{\frac{-u^2}{2}}, -\infty<u<\infty$ are the standard normal c.d.f. and p.d.f. respectively. We shall use this notation throughout this discussion.

Thus W has a non-central chi-distribution with one degree of freedom and non-centrality parameter equal to $n^{\frac{1}{2}}|\mu|$. This distribution is also referred to as the "folded normal" distribution with folding at the origin (see, for example, Elandt [5], Johnson [11] and Leone, Nelson and Nottingham [16]).

It is easy to see that both $H(w,\mu)$ and $h(w,\mu)$ are even functions of μ and hence letting $\theta=|\mu|\geq 0$ we can write

$$(2.3) \qquad H(w,\mu)=H(w,\theta)=F(n^{\frac{1}{2}}(w-\theta))-F(n^{\frac{1}{2}}(-w-\theta)), \quad w>0$$

(2.4) $$h(w,\ \mu) = h(w,\ \theta) = n^{\frac{1}{2}}\{f(n^{\frac{1}{2}}(w-\theta))+f(n^{\frac{1}{2}}(w+\theta))\}, \quad w > 0$$

and both are zero for $w \leqq 0$.

A very useful property of W is that its p.d.f. $h(w,\ \theta)$ has a "strict monotone likelihood ratio" in w (for definition of strict monotone likelihood ratio see Karlin [12]). For $\theta > 0$ this follows because $\frac{\partial^2 \log h(w,\ \theta)}{\partial\theta\partial w}$ exists and is positive for all $\theta > 0$ and $w > 0$ as shown below.

(2.5) $$\log h(w,\ \theta) = \log 2(n/2\pi)^{\frac{1}{2}} - \frac{n}{2}(w^2+\theta^2) + \log \cosh(nw\theta).$$

(2.6) $$\frac{\partial}{\partial w} \log h(w,\ \theta) = -nw + n\theta \tanh(nw\theta).$$

(2.7) $$\frac{\partial^2 \log h(w,\ \theta)}{\partial\theta\partial w} = n \tanh(nw\theta) + n^2 w\theta \operatorname{sech}^2(nw\theta) > 0.$$

Moreover, since $h(w,\ \theta)$ is right-continuous in θ at $\theta=0$, it follows that we have a strict monotone likelihood ratio for $\theta \geqq 0$ and $w > 0$. We shall use this property in proving certain results in Sections 3 and 7 of Chapter II and Section 5 of Chapter III.

CHAPTER II: INDIFFERENCE ZONE FORMULATION

1. Formal Statement of the Problem.

Let π_i denote a normal population with unknown mean μ_i $(i=1,2,\ldots,k)$ and common known variance which we take without loss of generality to be unity. Let the ordered values of the parameters $\theta_i = |\mu_i|$ $(i=1,2,\ldots,k)$ be denoted by

$$(1.1) \qquad 0 \leq \theta_{[1]} \leq \theta_{[2]} \leq \ldots \leq \theta_{[k]}.$$

It is assumed that there is no a priori information available about the correct pairing of the k populations π_i and the ordered parameters $\theta_{[i]}$ $(i=1,2,\ldots,k)$.

Our goal is to select t ($< k$) "best" populations in an unordered manner; a "better" population is defined to be one with a larger θ-value. When some or all equalities in (1.1) hold the choice of any t populations with parameters equal to the t largest θ-values is regarded as a correct selection (CS).

Let $\underset{\sim}{\theta} = (\theta_{[1]},\theta_{[2]},\ldots,\theta_{[k]})$ denote a point in the parameter space Ω, which is partitioned into a "preference zone" $\Omega^+(\delta^*)$ defined by

$$(1.2) \qquad \Omega^+(\delta^*) = \{\underset{\sim}{\theta}: \theta_{[k-t+1]}-\theta_{[k-t]} \geq \delta^* > 0\}$$

and its complement, the "indifference zone" $\Omega^-(\delta^*)$. The quantity $\delta^* > 0$ is specified in advance by the experimenter. Thus it is assumed that the experimenter is indifferent between the populations with parameter value $\theta_{[k-t+1]}$ and any other population with parameter value $\theta_{[j]}$ $(j=1,2,\ldots,k-t)$ if $\theta_{[k-t+1]}-\theta_{[j]} < \delta^*$; also he is indifferent between any population with parameter value $\theta_{[i]}$ $(i=k-t+1,\ldots,k)$ and the population with parameter value $\theta_{[k-t]}$ if $\theta_{[i]}-\theta_{[k-t]} < \delta^*$. In addition to specifying δ^*, the experimenter also specifies a constant P^*, $1/\binom{k}{t} < P^* < 1$.

After specifying δ^* and P^* the experimenter requires a procedure R_t for which the probability of a correct selection satisfies the condition

$$P\{CS|R_t, \underset{\sim}{\theta}\} \geq P^* \qquad \text{for all } \underset{\sim}{\theta} \in \Omega^+(\delta^*). \tag{1.3}$$

In the next section we propose a procedure and in subsequent sections we study its properties.

2. Proposed Procedure R_t.

Let $\bar{x}_i$ be the sample means based on a common pre-determined number n of independent observations from each π_i and let $w_i = |\bar{x}_i|$ $(i=1,2,\ldots,k)$. The ranked w_i are denoted by

$$0 \leq w_{[1]} \leq w_{[2]} \leq \ldots \leq w_{[k]}. \tag{2.1}$$

A tie in two or more w_i is an event of probability zero. (The tied w_i, if any, should be ranked by using a randomized procedure which assigns equal probability to each ordering.) We then assert that the populations corresponding to the t largest w_i are the t "best" populations and the remaining k-t populations are the "worst" populations.

Now n is determined as the smallest integer greater than or equal to the solution $n_o(k, t, P^*, \delta^*)$ of (1.3) with equality holding. This value of n is the common number of observations to be taken from each population.

3. Probability of Correct Selection and its Infimum.

Let $W_{(i)}$ denote the statistic (absolute value of the sample mean) from the population with parameter $\theta_{[i]} \geq 0$ $(i=1,2,\ldots,k)$. Then we have

$$P\{CS|R_t\} = P\{\max(W_{(1)},\ldots,W_{(k-t)}) < \min(W_{(k-t+1)},\ldots,W_{(k)})\}$$

$$= \sum_{j=k-t+1}^{k} \int_0^\infty \prod_{1 \leq \beta \leq k-t} H(w, \theta_{[\beta]}) \prod_{\substack{k-t+1 \leq \alpha \leq k \\ \alpha \neq j}} \{1-H(w, \theta_{[\alpha]})\} h(w, \theta_{[j]}) dw$$

$$(3.1) \quad = \sum_{j=k-t+1}^{k} \int_0^\infty \prod_{1 \le \beta \le k-t} \{F(n^{\frac{1}{2}}(w-\theta_{[\beta]})) - F(n^{\frac{1}{2}}(-w-\theta_{[\beta]}))\}$$

$$\prod_{\substack{k-t+1 \le \alpha \le k \\ \alpha \neq j}} \{F(n^{\frac{1}{2}}(-w+\theta_{[\alpha]})) + F(n^{\frac{1}{2}}(-w-\theta_{[\alpha]}))\}$$

$$n^{\frac{1}{2}}\{f(n^{\frac{1}{2}}(w-\theta_{[j]})) + f(n^{\frac{1}{2}}(w+\theta_{[j]}))\}dw.$$

This expression viewed as a function of $\underset{\sim}{\theta}$ can be regarded as the "operating characteristic" of the procedure. We shall be interested in finding infimum of $P\{CS|R_t\}$ over all $\underset{\sim}{\theta} \in \Omega^+(\delta^*)$ and the $\underset{\sim}{\theta}$ for which this infimum is attained. We shall call this $\underset{\sim}{\theta}$ the "least favorable configuration" of the parameters. We minimize $P\{CS|R_t\}$ in two steps. Toward this end we state Lemma 1 without proof and prove Lemma 2 and Theorem 1 below. A proof for Lemma 1 can be found in Lehmann [15, p. 74].

Lemma 1:

Let $p(x, \theta)$ be a family of densities on the real line with monotone likelihood ratio in x for the scalar parameter θ. If ψ is a nondecreasing (nonincreasing) function of x, then $E_\theta\psi(X)$ is a nondecreasing (nonincreasing) function of θ.

Lemma 2:

Let $p(x_i, \theta_i)$ be a family of densities on the real line with monotone likelihood ratio in x_i $(i=1,2,\ldots,k)$. Let $X_1,X_2,\ldots,X_k$ be independently distributed with densities $p(x_1, \theta_1), p(x_2, \theta_2),\ldots,p(x_k, \theta_k)$, respectively, and for any fixed i let ψ be a nondecreasing (nonincreasing) function of x_i holding all x_j $(j\neq i)$ fixed. Then $E\psi(X_1,X_2,\ldots,X_k)$ is a nondecreasing (nonincreasing) function of θ_i.

Proof:

For any fixed i, we can write

$$(3.2) \qquad E\psi(X_1,X_2,\ldots,X_k) = EE_{\theta_i}\{\psi(X_1,X_2,\ldots,X_k)|x_1,\ldots,x_{i-1}, x_{i+1},\ldots,x_k\}$$

where the outside expectation is over all x_j ($j \neq i$).

Now since ψ is a nondecreasing (nonincreasing) function of x_i holding all x_j ($j \neq i$) fixed, it follows from Lemma 1 above that for the fixed i the function $E_{\theta_i}\{\psi(X_1,X_2,\ldots,X_k)|x_1,\ldots,x_{i-1}, x_{i+1},\ldots,x_k\}$ is nondecreasing (nonincreasing) in θ_i. This holds for each value of $(x_1,\ldots,x_{i-1}, x_{i+1},\ldots,x_k)$. Then the right member of (3.2), and hence also the left member, is a nondecreasing (nonincreasing) function of θ_i.

Theorem 1:

Let $p(x_i, \theta_i)$ be a family of densities on the real line with monotone likelihood ratio in x_i ($i=1,2,\ldots,k$), $k \geq 2$. Let X_i be independently distributed with density $p(x_i, \theta_i)$ ($i=1,2,\ldots,k$), respectively. Suppose the population with density $p(x_i, \theta_i)$ is denoted by π_i ($i=1,2,\ldots,k$) and the choice of any $t < k$ populations with parameters equal to t largest θ-values is regarded as a correct selection (CS). Also, suppose procedure R_t is defined by ordering x_i and asserting that the populations corresponding to the t largest x_i ($i=1,2,\ldots,k$) are the populations with the t largest parameters. Then $P\{CS|R_t\}$ is a nondecreasing function of $\theta_{[\alpha]}$ ($\alpha=k-t+1, k-t+2,\ldots,k$) and a nonincreasing function of $\theta_{[\beta]}$ ($\beta=1,2,\ldots,k-t$), where $\theta_{[1]} \leq \theta_{[2]} \leq\ldots\leq \theta_{[k]}$ are the ordered θ-values.

Proof:

Let $X_{(i)}$ denote the statistic from the population with parameter $\theta_{[i]}$ ($i=1,2,\ldots,k$). Define a random variable ψ as follows:

$$(3.3) \qquad \psi = \begin{cases} 1 & \text{if } \max(X_{(1)},\ldots,X_{(k-t)}) < \min(X_{(k-t+1)},\ldots,X_{(k)}) \\ 0, & \text{otherwise.} \end{cases}$$

Then, obviously, $E(\psi) = P\{CS|R_t\}$. It is easy to see from (3.3) that for each α (α=k-t+1, k-t+2,...,k) ψ is a nondecreasing function of $X_{(\alpha)}$ holding all $X_{(i)}$ ($i \neq \alpha$) fixed and for each β (β=1,2,...,k-t) it is a nonincreasing function of $X_{(\beta)}$ holding all $X_{(j)}$ ($j \neq \beta$) fixed. Hence, from Lemma 2, it follows that $E(\psi) = P\{CS|R_t\}$ is a nondecreasing function of $\theta_{[\alpha]}$ (α=k-t+1, k-t+2,...,k) and a nonincreasing function of $\theta_{[\beta]}$ (β=1,2,...,k-t).

Now the statistic $W = |\bar{X}|$ in our problem has a monotone likelihood ratio as shown in Section 2, Chapter I and Theorem 1 applies. Consequently, for any fixed non-negative value of $\theta_{[k-t]} = \theta^*$ (say), (3.1) is minimized subject to the δ^*-condition in (1.2) by setting

$$(3.4) \qquad \theta_{[1]} = \theta_{[2]} = \ldots = \theta_{[k-t]} = \theta^*, \quad \theta_{[k-t+1]} = \theta_{[k-t+2]} = \ldots = \theta_{[k]} = \theta^*+\delta^*.$$

Using (3.1) and (3.4), renaming $n^{\frac{1}{2}}w$ as u, $n^{\frac{1}{2}}\theta^*$ as θ and $n^{\frac{1}{2}}\delta^*$ as λ and letting $T(\lambda)$ denote the infimum of $P\{CS|R_t\}$ over all $\underset{\sim}{\theta} \in \Omega^+(\delta^*)$, we obtain

$$(3.5) \qquad T(\lambda) = \inf_{\theta \geq 0} t \int_0^\infty \{F(u-\theta)-F(-u-\theta)\}^{k-t}\{F(-u+\theta+\lambda)+F(-u-\theta-\lambda)\}^{t-1}$$
$$\{f(u-\theta-\lambda)+f(u+\theta+\lambda)\}du.$$

Integrating (3.5) by parts, we have

$$(3.6) \qquad T(\lambda) = \inf_{\theta \geq 0} (k-t) \int_0^\infty \{F(u-\theta)-F(-u-\theta)\}^{k-t-1}\{F(-u+\theta+\lambda)+F(-u-\theta-\lambda)\}^{t}$$
$$\{f(u-\theta)+f(u+\theta)\}du.$$

For the second step of minimization of $P\{CS|R_t\}$ we need the following theorem.

Theorem 2:

$$I(\theta, \lambda) = \int_0^\infty \{F(u-\theta)-F(-u-\theta)\}^{k-t-1}\{F(-u+\theta+\lambda)+F(-u-\theta-\lambda)\}^t f(u-\theta)du$$
$$+ \int_0^\infty \{F(u-\theta)-F(-u-\theta)\}^{k-t-1}\{F(-u+\theta+\lambda)+F(-u-\theta-\lambda)\}^t f(u+\theta)du \tag{3.7}$$

is a strictly increasing function of θ for $\theta \geqq 0$ and fixed λ.

Proof:

Putting $u-\theta = y$ in the first and $u+\theta = y$ in the second integral in (3.7), we obtain

$$I(\theta, \lambda) = \int_{-\theta}^\infty \{F(y)-F(-y-2\theta)\}^{k-t-1}\{F(-y+\lambda)+F(-y-2\theta-\lambda)\}^t f(y)dy$$
$$+ \int_\theta^\infty \{F(y-2\theta)-F(-y)\}^{k-t-1}\{F(-y+2\theta+\lambda)+F(-y-\lambda)\}^t f(y)dy. \tag{3.8}$$

In order to show that $I(\theta, \lambda)$ is an increasing function of θ for fixed λ, we shall differentiate (3.8) with respect to θ, which is permissible, and show that the partial derivative is positive for all $\theta \geqq 0$. Differentiation gives

$$\frac{\partial I(\theta, \lambda)}{\partial \theta}$$
$$= 2(k-t-1) \int_{-\theta}^\infty \{F(y)-F(-y-2\theta)\}^{k-t-2}\{F(-y+\lambda)+F(-y-2\theta-\lambda)\}^t f(y+2\theta)f(y)dy$$
$$- 2t \int_{-\theta}^\infty \{F(y)-F(-y-2\theta)\}^{k-t-1}\{F(-y+\lambda)+F(-y-2\theta-\lambda)\}^{t-1} f(y+2\theta+\lambda)f(y)dy \tag{3.9}$$
$$- 2(k-t-1) \int_\theta^\infty \{F(y-2\theta)-F(-y)\}^{k-t-2}\{F(-y+2\theta+\lambda)+F(-y-\lambda)\}^t f(y-2\theta)f(y)dy$$
$$+ 2t \int_\theta^\infty \{F(y-2\theta)-F(-y)\}^{k-t-1}\{F(-y+2\theta+\lambda)+F(-y-\lambda)\}^{t-1} f(y-2\theta-\lambda)f(y)dy.$$

Substituting back $y = u-\theta$ in the first and second integral and $y = u+\theta$ in the third and fourth integral, the first and the third integral cancel each other and we obtain

$$(3.10) \qquad \frac{\partial I(\theta, \lambda)}{\partial \theta} = 2t \int_0^\infty \{F(u-\theta)-F(-u-\theta)\}^{k-t-1}\{F(-u+\theta+\lambda)+F(-u-\theta-\lambda)\}^{t-1}$$

$$\{f(u-\theta-\lambda)f(u+\theta) - f(u+\theta+\lambda)f(u-\theta)\}du.$$

But it is easily seen from the strict monotone likelihood ratio property of the normal p.d.f. that $f(u-\theta-\lambda)f(u+\theta) > f(u+\theta+\lambda)f(u-\theta)$ for $\lambda > 0$, $u > 0$. Hence $\frac{\partial I(\theta, \lambda)}{\partial \theta} > 0$ for all $\theta \geq 0$ and $I(\theta, \lambda)$ is therefore a strictly increasing function of θ for $\theta \geq 0$ and fixed λ.

Now using (3.4) and applying Theorem 2, the least favorable configuration of the parameters for our problem is

$$(3.11) \qquad \theta_{[1]} = \theta_{[2]} = \ldots = \theta_{[k-t]} = 0, \quad \theta_{[k-t+1]} = \theta_{[k-t+2]} = \ldots = \theta_{[k]} = \delta^* > 0$$

and the infimum of $P\{CS|R_t\}$ is obtained by setting θ equal to zero in (3.6). Thus,

$$(3.12) \qquad T(\lambda) = 2(k-t) \int_0^\infty \{2F(u)-1\}^{k-t-1}\{F(-u+\lambda)+F(-u-\lambda)\}^t f(u)du$$

where $\lambda = n^{\frac{1}{2}}\delta^*$.

Hence, a solution to our problem is obtained by finding the solution $\lambda = \lambda(k, t, P^*)$ of

$$(3.13) \qquad 2(k-t) \int_0^\infty \{2F(u)-1\}^{k-t-1}\{F(-u+\lambda)+F(-u-\lambda)\}^t f(u)du = P^*$$

and taking n to be the smallest integer greater than or equal to $(\lambda/\delta^*)^2$. The existence of this solution follows from the fact that the left member of (3.13) approaches unity as λ goes to infinity. The uniqueness of the solution follows from the fact that the left member of (3.13) is a strictly increasing function of λ, which is easily seen by differentiation under the integral sign. Thus there exists a unique n satisfying (1.3).

4. Certain Bounds on $P\{CS|R_t\}$.

There is an analogue of "power" in our formulation; it is the $P\{CS|R_t\}$ viewed as a function of the true parameter values. We shall be interested in studying the $P\{CS\}$-function $(k-t)\, I(\theta, \lambda)$ for the configuration (3.4) of the parameters, where both $\theta = n^{\frac{1}{2}}\theta^* \geq 0$ and $\lambda = n^{\frac{1}{2}}\delta^* \geq 0$ can vary.

First let λ be fixed. Then $I(\theta, \lambda)$ is a strictly increasing function of θ by Theorem 2. The infimum of $(k-t)\, I(\theta, \lambda)$, attained at $\theta=0$, is given by the left member of (3.13). The supremum of $(k-t)\, I(\theta, \lambda)$ corresponds to $\theta=\infty$ and using (3.8) we obtain

$$\begin{aligned}
\sup_{\theta \geq 0}\ (k-t)\, I(\theta, \lambda) &= (k-t) \int_{-\infty}^{\infty} F^{k-t-1}(y) F^{t}(-y+\lambda) f(y)\,dy \\
&= t \int_{-\infty}^{\infty} F^{k-t}(y)[1-F(y-\lambda)]^{t-1} f(y-\lambda)\,dy \qquad (4.1)\\
&= t \int_{-\infty}^{\infty} F^{k-t}(x+\lambda)[1-F(x)]^{t-1} f(x)\,dx.
\end{aligned}$$

This is the integral given in Bechhofer [1, (20)] and this is an upper bound on $T(\lambda)$ for any pair (k, t).

Next let θ be fixed. Then we shall prove that $I(\theta, \lambda)$ is a strictly increasing function of λ by showing that the partial derivative of $I(\theta, \lambda)$ with respect to λ is positive for all $\lambda \geq 0$. Thus

$$(4.2) \qquad \frac{\partial I(\theta, \lambda)}{\partial \lambda} = t \int_{0}^{\infty} \{F(u-\theta)-F(-u-\theta)\}^{k-t-1}\{F(-u+\theta+\lambda)+F(-u-\theta-\lambda)\}^{t-1}$$

$$\{f(-u+\theta+\lambda)-f(u+\theta+\lambda)\}\{f(u-\theta)+f(u+\theta)\}du.$$

But $f(-u+\theta+\lambda) > f(u+\theta+\lambda)$ for $u > 0$, $\theta \geqq 0$, $\lambda \geqq 0$ (excepting $\theta=\lambda=0$ where equality holds) and therefore $\frac{\partial I(\theta, \lambda)}{\partial\lambda} > 0$ for all $\lambda \geqq 0$ and fixed θ. Hence the assertion.

By considering $\lambda=0$ we find as a consequence of the above that a lower bound on the $P\{CS|R_t\}$ over all parameter points is $1/\binom{k}{t}$, which is the result to be expected. Also by letting $\lambda \rightarrow \infty$ we obtain unity, which is the supremum of the $P\{CS|R_t\}$ over all parameter points.

5. Special Cases and Tables.

The following two cases are of special interest.

Case A: $t=1$, i.e., we want to select the population with the largest θ-value.

Case B: $t=k-1$; in our formulation this is equivalent to selecting the population with the smallest θ-value.

We now set $t=1,k-1$ successively in (3.13). Thus for Case A we determine λ from the equation

$$(5.1) \qquad 2(k-1) \int_0^\infty \{2F(u)-1\}^{k-2}\{F(-u+\lambda)+F(-u-\lambda)\}f(u)du = P^*$$

and for Case B from the equation

$$(5.2) \qquad 2 \int_0^\infty \{F(-u+\lambda)+F(-u-\lambda)\}^{k-1}f(u)du = P^*.$$

The equation (5.1) can be solved for λ by the use of Gauss-Legendre quadrature after truncating the upper limit of the integral at a suitable finite value (for our Tables we have used 10 as an upper truncation point) and making a transformation which brings the limits of integration to the standard form. A similar method is used for solving (5.2).

Tables for the above two special cases have been prepared and are given after Chapter IV. We have used 32 point Gauss-Legendre quadrature formula in constructing these tables. Table I gives the probability of a correct selection for Case A based on the left member of (5.1) for k=2(1)10 and $\lambda = n^{\frac{1}{2}}\delta^* = 0(0.1)7.0$. Table II is also based on (5.1) and gives the value of $\lambda = n^{\frac{1}{2}}\delta^*$ associated with specified probabilities $P^* = .5000$, .7500, .9000, .9500, .9750, .9900, .9950, .9990, .9995 and .9999 for k=2(1)10. Tables III and IV are to be used in place of Tables I and II respectively for Case B. Entries in all our Tables are accurate to all the six decimal places that are given. In view of (4.1) it should be observed that the entries of Table II are greater than the corresponding entries of Table I of Bechhofer [1]. To find the sample size n for Case A (Case B) we first find λ corresponding to the given values of P^* and k from Table II (Table IV) and then determine n as the smallest positive integer greater than or equal to the solution of $\lambda = n^{\frac{1}{2}}\delta^*$ where δ^* is known.

For k=2, equations (5.1) and (5.2) are the same, since the two problems are then equivalent, and we can further simplify the common left member

$$(5.3) \qquad T(\lambda) = 2\int_0^\infty \{F(-u+\lambda)+F(-u-\lambda)\}f(u)du.$$

Differentiating under the integral sign with respect to λ, which is permissible, we obtain

$$(5.4) \qquad \frac{dT(\lambda)}{d\lambda} = 2\int_0^\infty \{f(u-\lambda)-f(u+\lambda)\}f(u)du$$

$$= 2f\left(\frac{\lambda}{\sqrt{2}}\right)\int_0^\infty \{f(\sqrt{2}(u-\tfrac{\lambda}{2}))-f(\sqrt{2}(u+\tfrac{\lambda}{2}))\}du$$

$$= \sqrt{2}\,f\left(\frac{\lambda}{\sqrt{2}}\right)\left\{F\left(\frac{\lambda}{\sqrt{2}}\right)-F\left(-\frac{\lambda}{\sqrt{2}}\right)\right\}.$$

Hence

$$(5.5) \qquad T(\lambda) = F^2\left(\frac{\lambda}{\sqrt{2}}\right)+F^2\left(-\frac{\lambda}{\sqrt{2}}\right) = 1-2F\left(\frac{\lambda}{\sqrt{2}}\right)+2F^2\left(\frac{\lambda}{\sqrt{2}}\right),$$

the constant of integration being zero.

Letting $y = F\left(\frac{\lambda}{\sqrt{2}}\right)$ we can write the equation (5,1) or (5.2) with $k=2$ as the quadratic equation

$$(5.6) \qquad 2y^2-2y+1-P^* = 0$$

where $\frac{1}{2} < P^* < 1$.

Since $y > \frac{1}{2}$, the admissible root of this equation is

$$(5.7) \qquad y = F\left(\frac{\lambda}{\sqrt{2}}\right) = \tfrac{1}{2}\{1+(2P^*-1)^{\frac{1}{2}}\}.$$

It should be pointed out that (5.7) served as a check of the accuracy of the quadrature employed for solving (5.1) or (5.2) for λ. Let us consider an illustration. Consider $k=2$ populations with a common variance $\sigma^2=100$ and let $P^*=0.95$. The relation $\lambda = n^{\frac{1}{2}}\delta^*$ is now replaced by $\lambda = n^{\frac{1}{2}}\delta^*/\sigma$ and we obtain from Table II, $\frac{n^{\frac{1}{2}}\delta^*}{\sigma} = 2.756050$ so that if $\delta^*=1$, we need $n=760$ observations from each of the two populations. Now as a check, using equation (5,7) with $P^*=0.95$ we obtain $y = F\left(\frac{\lambda}{\sqrt{2}}\right) = 0.97434$. This gives $\lambda=2.7560$ so that if $\delta^*=1$, we need $n=760$ observations from each of the two populations.

6, Allocation of Sample Sizes.

Case I: Variances equal with the common value known.

This is the case considered above. All the k normal populations are assumed to have the same known variance σ^2, which, just for the sake of notational convenience, we have taken as unity. The consideration of invariance under permutations of labels of the populations for single-sample procedures suggests the use of a common number n of observations from each population. We choose n as the smallest positive integer greater than or equal to the positive number n_o satisfying equality in (1.3).

Case II: Variances known and unequal.

If the k normal populations have variances $\sigma_i^2 = a_i\sigma^2$ $(i=1,2,\ldots,k)$ where σ^2 and a_i are known constants (all a_i equal corresponds to Case I above), then it may be desirable to choose the sample sizes n_i $(i=1,2,\ldots,k)$ in such a way as to make the variances of the sample means equal. No optimal properties for this choice are known at present, but it has an important practical advantage, namely, that the tables for the Case I mentioned above become applicable. We act as if the k populations have the common variance $\sigma^2(= \frac{\sigma_i^2}{a_i})$, which is given, and then find the required common number n_o of groups of observations, where n_o is positive but not necessarily an integer. The common variance of the sample is then

$$(6.1) \qquad \frac{\sigma_i^2}{n_i} = \frac{a_i}{n_i}\sigma^2 = \frac{\sigma^2}{n_o} \quad (i=1,2,\ldots,k).$$

Thus we can choose n_i as the smallest positive integer greater than or equal to the solution of $n_i = n_o a_i$ $(i=1,2,\ldots,k)$ and the probability requirement (1.3) will be satisfied. It may happen that for some values of i (but not all) we can take the largest integer less than the solution of $n_i = n_o a_i$ and still satisfy (1.3); this is related to the fact that the solution is not in general unique.

Case III: Variances equal with the common value unknown.

If $\sigma_1^2 = \sigma_2^2 = \ldots = \sigma_k^2 = \sigma^2$ (say) but the common value σ^2 is unknown, the proposed single-stage procedure R_t with preference zone now defined as

$$(6.2) \qquad \Omega^{++} = \{\underset{\sim}{\theta}: \frac{\theta_{[k-t+1]}-\theta_{[k-t]}}{\sigma} \geq \delta^* > 0\}$$

again satisfies (1.3). This, however, puts the burden on the experimenter to specify the quantity δ^* in standardized units.

It should be pointed out that Case III and the case where variances are unknown and unequal are more realistic from a practical point of view than Cases I and II above because one rarely knows variances or even ratios of variances without knowing means.

7. Some Properties of the Procedure R_t.

7.1. General remarks.

Consider a multiple decision problem. Let A_i $(i=1,2,\ldots,N)$ denote the N actions. Suppose $\underset{\sim}{X} = (X_1,X_2,\ldots,X_k)$ has c.d.f. $P(\underset{\sim}{x}, \underset{\sim}{\theta})$ where $\underset{\sim}{X} \in M$ and $\underset{\sim}{\theta} = (\theta_1,\theta_2,\ldots,\theta_k) \in \Omega$; we shall assume that $P(\underset{\sim}{x}, \underset{\sim}{\theta})$ has a joint density given by $p(\underset{\sim}{x}, \underset{\sim}{\theta})$. Let $L_i(\underset{\sim}{\theta})$ denote the loss in taking action A_i in the presence of $\underset{\sim}{\theta}$. A decision rule $\underset{\sim}{\varphi}(\underset{\sim}{x}) = (\varphi_1(\underset{\sim}{x}),\varphi_2(\underset{\sim}{x}),\ldots,\varphi_N(\underset{\sim}{x}))$ is a function satisfying the conditions that $0 \leq \varphi_i(\underset{\sim}{x}) \leq 1$ $(i=1,2,\ldots,N)$ and $\sum_{i=1}^{N} \varphi_i(\underset{\sim}{x}) = 1$ for each $\underset{\sim}{x}$. Let Φ be the set of all decision functions; the risk function r is a function defined on $\Phi \times \Omega$ by

$$(7.1) \qquad r(\underset{\sim}{\varphi}, \underset{\sim}{\theta}) = \sum_{i=1}^{N} L_i(\underset{\sim}{\theta}) \int_M \varphi_i(\underset{\sim}{x})dP(\underset{\sim}{x}, \underset{\sim}{\theta}).$$

Consider a prior distribution $\eta(\underset{\sim}{\theta})$ on Ω. The Bayes risk with respect to $\eta(\underset{\sim}{\theta})$ is given by

$$(7.2) \qquad \rho(\underset{\sim}{\varphi}, \eta) = \int_\Omega r(\underset{\sim}{\varphi}, \underset{\sim}{\theta})d\eta(\underset{\sim}{\theta}) = \int_\Omega \int_M \sum_{i=1}^{N} L_i(\underset{\sim}{\theta})\varphi_i(\underset{\sim}{x})dP(\underset{\sim}{x}, \underset{\sim}{\theta})d\eta(\underset{\sim}{\theta}).$$

A decision rule $\underset{\sim}{\varphi} \in \Phi$ that minimizes $\rho(\underset{\sim}{\varphi}, \eta)$ is called a Bayes decision rule. It is well known (see, for example, Wald [20, p. 124]) that $\underset{\sim}{\varphi}$ is a Bayes decision rule if and only if $\sum_i \varphi_i(\underset{\sim}{x}) = 1$ where the summation is over those i-values such that

$$(7.3) \qquad \int_\Omega L_i(\underset{\sim}{\theta})p(\underset{\sim}{x}, \underset{\sim}{\theta})d\eta(\underset{\sim}{\theta}) = \min_{j=1,\ldots,N} \int_\Omega L_j(\underset{\sim}{\theta})p(\underset{\sim}{x},\underset{\sim}{\theta})d\eta(\underset{\sim}{\theta})$$

and $\varphi_i(\underset{\sim}{x}) = 0$ if i is such that an inequality holds in (7.3).

7.2 Admissibility and minimax nature of R_t.

We shall first find a Bayes decision rule for our problem. In our problem the $N = \binom{k}{t}$ actions correspond to the set of all possible selections of t best out of k populations. We define the simple loss function

$$(7.4) \qquad L_i(\underset{\sim}{\theta}) = 0, \quad \text{for correct selection}$$
$$= 1, \quad \text{otherwise}$$

for $\underset{\sim}{\theta} \in \Omega$ and $i=1,2,\ldots,N$. Consider the prior distribution $\eta^o(\underset{\sim}{\theta})$ on Ω which puts equal mass at each of the N points obtained by permuting the components of $\underset{\sim}{\theta}^o = (0,\ldots,0, \delta^*,\ldots,\delta^*)$ where t of the components of $\underset{\sim}{\theta}^o$ are equal to δ^* and the remaining (k-t) are zero. We label these N points as $\underset{\sim}{\theta}^o_\alpha$ ($\alpha=1,2,\ldots,N$) so that the action A_α is correct in the presence of $\underset{\sim}{\theta}^o_\alpha$ ($\alpha=1,2,\ldots,N$).

Our problem remains invariant under a group G of transformations (of the space of sample means) where an arbitrary element $g \in G$ is defined by

$$(7.5) \qquad g(\bar{X}_1,\bar{X}_2,\ldots,\bar{X}_k) = (a_1\bar{X}_1,a_2\bar{X}_2,\ldots,a_k\bar{X}_k), \quad a_\beta = \pm 1 \ (\beta=1,2,\ldots,k)$$

where $\bar{X}_\beta$ is the sample mean from the normal population π_β with mean μ_β and variance unity ($\beta=1,2,\ldots,k$). This transformation in turn induces the group $\bar{G}$ of transformations on the parameter space with elements $\bar{g} \in \bar{G}$ given by

$$(7.6)\qquad \bar{g}(\mu_1,\mu_2,\ldots,\mu_k)=(a_1\mu_1,a_2\mu_2,\ldots,a_k\mu_k),\quad a_\beta=\pm 1\quad(\beta=1,2,\ldots,k).$$

Clearly, the family of the underlying distributions and the structure of the loss function (7.4) remain invariant under G. Then the principle of invariance restricts consideration to the class of invariant decision rules. Using the definition of maximal invariant (see, for example, Lehmann [15, p. 215]), it is easy to check that $\underset{\sim}{W}=(W_1,W_2,\ldots,W_k)$, where $W_\beta = |\bar{X}_\beta|$ $(\beta=1,2,\ldots,k)$, is a maximal invariant with respect to G. So we base our decision rule on $\underset{\sim}{W}$ only. It should be noted that the distribution of $\underset{\sim}{W}$ depends only on $\underset{\sim}{\theta}=(\theta_1,\theta_2,\ldots,\theta_k)$, $\theta_\beta = |\mu_\beta|$ $(\beta=1,2,\ldots,k)$, which is the maximal invariant of the induced group $\bar{G}$, as it should according to Lehmann [15, Theorem 3, p. 220].

To obtain an invariant Bayes rule with respect to the prior distribution $\eta^{o}(\underset{\sim}{\theta})$ defined above using (7.3) and following Karlin and Truax [13], we consider the expressions

$$(7.7)\qquad d_{ij}=\sum_{\alpha=1}^{N}\{[L_j(\underset{\sim}{\theta}^{o}_{\alpha})-L_i(\underset{\sim}{\theta}^{o}_{\alpha})]\prod_{\beta=1}^{k}h(w_\beta,\ \theta^{o}_{\alpha,\beta})\},\quad i,j=1,2,\ldots,N,\ i\neq j$$

where $h(w,\ \theta)$ is the p.d.f. of W defined by (2.4), Chapter I and $\theta^{o}_{\alpha,\beta}$ $(\beta=1,2,\ldots,k)$ are the components of $\underset{\sim}{\theta}^{o}_{\alpha}$ $(\alpha=1,2,\ldots,N)$. For the loss function (7.4) the expression d_{ij} $(i\neq j)$ simplifies to

$$(7.8)\qquad d_{ij}=\prod_{\beta=1}^{k}h(w_\beta,\ \theta^{o}_{i,\beta})-\prod_{\beta=1}^{k}h(w_\beta,\ \theta^{o}_{j,\beta}).$$

Let $\theta^{o}_{i_1},\theta^{o}_{i_2},\ldots,\theta^{o}_{i_t}$ be those components of $\underset{\sim}{\theta}^{o}_{i}$ which are equal to δ^* and let c_{ij} (possibly unity) be the factor common to the two product terms on the right side of (7.8). Then (7.8) can be written as

$$(7.9)\qquad d_{ij}=c_{ij}\left\{\prod_{s=1}^{\tau}h(w_{\ell_s},\ \delta^*)h(w_{m_s},\ 0)-\prod_{s=1}^{\tau}h(w_{\ell_s},\ 0)h(w_{m_s},\ \delta^*)\right\}$$

where $\{\ell_1,\ell_2,\ldots,\ell_\tau\}$ is a subset of $\{i_1,i_2,\ldots,i_t\}$ and $\{m_1,m_2,\ldots,m_\tau\}$ is a subset of the complement of $\{i_1,i_2,\ldots,i_t\}$. It should be noted that the number of elements in the sets $\{\ell_1,\ell_2,\ldots,\ell_\tau\}$ and $\{m_1,m_2,\ldots,m_\tau\}$ is the same because each of the two sets $\{\theta_{i,\beta}\}$ and $\{\theta_{j,\beta}\}$ $(\beta=1,2,\ldots,k)$ has t elements equal to δ^* and $(k-t)$ elements equal to zero. Now the expression d_{ij} given by (7.9) is non-negative for all $i\neq j$ if

$$\max_{\nu\neq i_1,i_2,\ldots,i_t}(w_\nu) < \min(w_{i_1},w_{i_2},\ldots,w_{i_t}). \tag{7.10}$$

This is a consequence of the monotone likelihood ratio property of the p.d.f. of W_β $(\beta=1,2,\ldots,k)$. A similar argument, in a different context, appears in Savage [18]. Since (7.10) describes the proposed procedure R_t, it follows that R_t is an invariant Bayes decision rule with respect to the prior distribution $\eta^o(\underset{\sim}{\theta})$. Moreover, since we have a strict monotone likelihood ratio property, the procedure R_t is the unique invariant Bayes decision rule with respect to this prior distribution.

The risk of the decision rule R_t with respect to the loss function (7.4) is given by

$$r(R_t, \underset{\sim}{\theta}) = 1-P\{CS|R_t\}. \tag{7.11}$$

Now we have demonstrated in Section 3 that the $P\{CS|R_t\}$ is minimized for the configuration (3.11) and, consequently, $r(R_t, \underset{\sim}{\theta})$ is maximized at each of the points $\underset{\sim}{\theta}^o_\alpha$ $(\alpha=1,2,\ldots,N)$. We now state without proof a lemma due to Lehmann [14, p. 4-19]. In what follows let $\underset{\sim}{\delta}_\eta$ denote a Bayes decision rule with respect to the prior distribution $\eta(\underset{\sim}{\theta})$.

<u>Lemma 3</u>:

Let $\underset{\sim}{X}$ have distribution $P(\underset{\sim}{x}, \underset{\sim}{\theta})$, $\underset{\sim}{\theta} \in \Omega$. Suppose there is a distribution $\eta(\underset{\sim}{\theta})$ over Ω and a set $\omega \subset \Omega$ such that $\eta(\omega) = 1$ and

$$r(\underset{\sim}{\delta}_\eta, \underset{\sim}{\theta}) = \sup_{\underset{\sim}{\theta}' \in \Omega} r(\underset{\sim}{\delta}_\eta, \underset{\sim}{\theta}') \qquad \text{for all } \underset{\sim}{\theta} \in \omega. \tag{7.12}$$

Then $\underset{\sim}{\delta}_\eta$ is minimax. Further if $\underset{\sim}{\delta}_\eta$ is unique Bayes then $\underset{\sim}{\delta}_\eta$ is unique minimax and hence also admissible.

With $\eta(\underset{\sim}{\theta}) = \eta^o(\underset{\sim}{\theta})$ and $\omega = \{\underset{\sim}{\theta}^o_\alpha \ (\alpha=1,2,\ldots,N)\}$ and $\underset{\sim}{X}$ replaced by $\underset{\sim}{W}$, Lemma 3 is immediately applicable to our problem within the framework of invariant procedures. Consequently we conclude that the proposed decision rule R_t is unique minimax and hence also admissible in the class of all invariant procedures. Since the group G of transformations (7.5) is finite, it follows from Blackwell and Girshick [4, p. 227] that R_t is minimax in the class of all procedures and from Theorem 8.6.6 of Blackwell and Girshick [4, p. 228] that R_t is admissible in the class of all procedures.

7.3 Most economical character of R_t.

We shall show that R_t is a most economical decision rule (see Hall [9] and [10]), that is, no other rules can satisfy (1.3) with a smaller fixed sample size.

Let n be the smallest common positive integer such that the rule $R_t(n)$ satisfies (1.3). If $R_t(n)$ is not most economical, then there exists a rule $\delta_t(n')$ which satisfies (1.3) with a common positive integer $n' < n$. Now consider the rule $R_t(n')$. Since R_t is minimax it follows that $R_t(n')$ satisfies (1.3) but this contradicts the above assumption that n is the smallest positive integer such that $R_t(n)$ satisfies (1.3). Hence $R_t(n)$ is a most economical rule.

CHAPTER III: SUBSET FORMULATION

1. Formal Statement of the Problem.

Let π_i denote a normal population with unknown mean $\mu_i(i=1,2,\ldots,k)$ and variance unity. Let the order values of $\theta_i = |\mu_i|(i=1,2,\ldots,k)$ be denoted by

(1.1) $$0 \leq \theta_{[1]} \leq \theta_{[2]} \leq \ldots \leq \theta_{[k]}.$$

It is assumed that there is no a priori information available about the correct pairing of the k populations and the ordered parameters $\theta_{[i]}$.

Any population with parameter equal to $\theta_{[k]}$ is called a "best" population. The goal is to select a non-empty subset of the k populations containing a best population; we would like this subset to be small and yet large enough to satisfy a certain probability requirement given below. Any selection of a subset which contains at least one population with a parameter value equal to $\theta_{[k]}$ will be called a correct selection (CS).

The problem is to find a rule R such that for a pre-assigned probability $P^*(\frac{1}{k}<P^* \leq 1)$

(1.2) $$P\{CS|R\} \geq P^*$$

regardless of the true unknown value of $\underset{\sim}{\theta} = (\theta_{[1]},\theta_{[2]},\ldots,\theta_{[k]}) \in \Omega$.

2. Proposed Procedure R.

Let $\bar{x}_i(i=1,2,\ldots,k)$ be the sample means based on a common number n of independent observations from each π_i and let $w_i = |\bar{x}_i|$; the c.d.f. and the p.d.f. of a typical W are given respectively by (2.3) and (2.4) of Chapter I. The ranked w_i are denoted by

(2.1) $$w_{[1]} \leq w_{[2]} \leq \ldots . \leq w_{[k]}.$$

Then the procedure R is defined as follows. Retain π_i in the selected subset if and only if

$$(2.2) \qquad w_i \geq w_{[k]} - d, \qquad (i=1,2,\ldots,k),$$

where $d = d(n, k, P^*)$ is a non-negative constant determined in advance of the experimentation. The constant d is chosen to be the smallest non-negative value satisfying (1.2) for all $\underset{\sim}{\theta} \in \Omega$.

3. Probability of Correct Selection and its Infimum.

Let $W_{(i)}$ denote the statistic (absolute value of the sample mean) associated with the population with parameter value $\theta_{[i]}$ $(i=1,2,\ldots,k)$. Then

$$P\{CS|R\} = P\{W_{(k)} \geq W_{[k]}-d\} = P\{W_{(i)} < W_{(k)}+d,\ i=1,2,\ldots,k-1\}$$

$$(3.1) \qquad = \int_0^\infty \prod_{i=1}^{k-1} \{F(n^{\frac{1}{2}}(w+d-\theta_{[i]}))-F(n^{\frac{1}{2}}(-w-d-\theta_{[i]}))\}$$
$$n^{\frac{1}{2}}\{f(n^{\frac{1}{2}}(w-\theta_{[k]}))+f(n^{\frac{1}{2}}(w+\theta_{[k]}))\}dw.$$

Now $F(n^{\frac{1}{2}}(w+d-\theta_{[i]}))-F(n^{\frac{1}{2}}(-w-d-\theta_{[i]}))$ is a strictly decreasing function of $\theta_{[i]}$ and hence as a first step for obtaining the infimum of $P\{CS|R\}$ over all $\underset{\sim}{\theta} \in \Omega$ we can take, for any fixed non-negative value of $\theta_{[k]} = \theta^*$(say),

$$(3.2) \qquad \theta_{[1]} = \theta_{[2]} = \ldots = \theta_{[k-1]} = \theta_{[k]} = \theta^*.$$

With (3.2) and renaming $n^{\frac{1}{2}}w$ as u, $n^{\frac{1}{2}}\theta^*$ as θ and $n^{\frac{1}{2}}d$ as γ we can write

$$(3.3) \qquad \inf_{\underset{\sim}{\theta} \in \Omega} P\{CS|R\} = \inf_{\theta \geq 0} \int_0^\infty \{F(u+\gamma-\theta)-F(-u-\gamma-\theta)\}^{k-1}\{f(u-\theta)+f(u+\theta)\}du.$$

Now we prove the following theorem.

Theorem 3:

$$(3.4)\qquad J(\theta,\gamma)=\int_0^\infty \{F(u+\gamma-\theta)-F(-u-\gamma-\theta)\}^{k-1}\{f(u-\theta)+f(u+\theta)\}du$$

is a nonincreasing function of θ for $\theta \geqq 0$ and fixed $\gamma \geqq 0$ and is strictly decreasing for $\gamma > 0$.

Proof:

$$(3.5)\qquad J(\theta,\gamma)=\int_0^\infty \{F(u+\gamma-\theta)-F(-u-\gamma-\theta)\}^{k-1}f(u-\theta)du$$

$$+\int_0^\infty \{F(u+\gamma-\theta)-F(-u-\gamma-\theta)\}^{k-1}f(u+\theta)du.$$

Putting $u-\theta=x$ in the first and $u+\theta=x$ in the second integral of (3.5) we obtain

$$(3.6)\qquad J(\theta,\gamma)=\int_{-\theta}^\infty \{F(x+\gamma)-F(-x-\gamma-2\theta)\}^{k-1}f(x)dx$$

$$+\int_\theta^\infty \{F(x+\gamma-2\theta)-F(-x-\gamma)\}^{k-1}f(x)dx.$$

Now the partial differentiation of (3.6) with respect to θ, which is permissible, yields

$$(3.7)\qquad \frac{\partial J}{\partial\theta}(\theta,\gamma)=2(k-1)\int_{-\theta}^\infty \{F(x+\gamma)-F(-x-\gamma-2\theta)\}^{k-2}f(x+\gamma+2\theta)f(x)dx$$

$$-2(k-1)\int_\theta^\infty \{F(x+\gamma-2\theta)-F(-x-\gamma)\}^{k-2}f(x+\gamma-2\theta)f(x)dx.$$

Substituting back $x=u-\theta$ in the first and $x=u+\theta$ in the second integral of (3.7) we obtain

$$(3.8)\qquad \frac{\partial J}{\partial\theta}(\theta,\gamma)=2(k-1)\int_0^\infty \{F(u+\gamma-\theta)-F(-u-\gamma-\theta)\}^{k-2}\{f(u+\gamma+\theta)f(u-\theta)$$

$$-f(u+\gamma-\theta)f(u+\theta)\}du.$$

Now $f(u+\gamma+\theta)f(u-\theta) \leq f(u+\gamma-\theta)f(u+\theta)$ and so also $\frac{\partial J}{\partial \theta}(\theta, \gamma) \leq 0$ for $\theta \geq 0$ and fixed $\gamma \geq 0$. In particular, for $\gamma > 0$ and $\theta > 0$, the inequality is strict. This proves Theorem 3.

Thus, the least favorable configuration of the parameters for our problem is obtained by letting

$$\theta_{[1]} = \theta_{[2]} = \ldots = \theta_{[k]} = \theta^* \to \infty \tag{3.9}$$

and from (3.6) the infimum of $P\{CS|R\}$ is obtained as

$$\inf_{\underset{\sim}{\theta} \in \Omega} P\{CS|R\} = \lim_{\theta \to \infty} \Big[\int_{-\theta}^{\infty} \{F(x+\gamma)-F(-x-\gamma-2\theta)\}^{k-1} f(x)dx$$

$$+ \int_{\theta}^{\infty} \{F(x+\gamma-2\theta)-F(-x-\gamma)\}^{k-1} f(x)dx\Big]$$

$$= \int_{-\infty}^{\infty} F^{k-1}(x+\gamma)f(x)dx. \tag{3.10}$$

Therefore, a solution to our problem is obtained by solving for γ the equation

$$\int_{-\infty}^{\infty} F^{k-1}(x+\gamma)f(x)dx = P^*, \tag{3.11}$$

and using the relation $\gamma = n^{\frac{1}{2}}d$ to solve for d. The existence of this solution follows from the fact that the left member of (3.11) approaches unity as γ goes to infinity. The uniqueness of the solution follows from the fact that the left member of (3.11) is a strictly increasing function of γ. Thus there exists a unique non-negative d satisfying (1.2). The left member of (3.11) is the same as integral (20) of Bechhofer [1] with t=1; the quantity γ obtained as the solution of (3.11) is tabulated by Bechhofer [1, Table I] with t=1. The left member of (3.11) is tabulated extensively in Teichroew [19].

Let us further examine the function $J(\theta, \gamma)$ defined by (3.4) and viewed as a function of θ and γ where both $\theta = n^{\frac{1}{2}}\theta^* \geq 0$ and $\gamma = n^{\frac{1}{2}}d \geq 0$ can vary.

The function $J(\theta, \gamma)$ is strictly decreasing in θ for fixed $\gamma > 0$ (Theorem 3). Its infimum is given by (3.10) and the supremum by

$$(3.12) \qquad \sup_{\theta \geq 0} J(\theta, \gamma) = 2 \int_0^{\infty} \{2F(u+\gamma)-1\}^{k-1} f(u)du.$$

For fixed θ, it is easily seen by differentiation that $J(\theta, \gamma)$ is strictly increasing in γ. Its supremum is unity and the infimum is

$$(3.13) \qquad \inf_{\theta \geq 0} J(\theta, \gamma) = \int_0^{\infty} \{F(u-\theta)-F(-u-\theta)\}^{k-1}\{f(u-\theta)+f(u+\theta)\}du = \frac{1}{k}$$

which is to be expected for $\gamma = n^{\frac{1}{2}}d = 0$.

4. Expected Size of the Selected Subset and its Supremum.

For the procedure R, the size S of the selected subset is a chance variable which can take on only integer values from 1 to k, inclusive. For any fixed values of n, k and P^*, the expected value of S is a function of the true configuration $\underset{\sim}{\theta}$ and this function will be regarded as a criterion of the efficiency of any procedure which satisfies (1.2).

4.1. Exact expression for the expected size.

Let $\chi(\underset{\sim}{w}; A_i)$ be the indicator function of the set $A_i = \{\underset{\sim}{w}: w_i \geq w_{[k]}-d\}$, that is,

$$(4.1) \qquad \chi(\underset{\sim}{w}; A_i) = \begin{cases} 1, & \underset{\sim}{w} \in A_i \\ 0, & \text{otherwise} \end{cases}$$

where $\underset{\sim}{w} = (w_1, w_2, \ldots, w_k)$ and $i=1,2,\ldots,k$. Then, for any values of n, k, P^* and $\underset{\sim}{\theta}$, and using $W_{(i)}$ defined in Section 3,

$$E\{S\} = E\{\sum_{i=1}^{k} \chi(\underset{\sim}{W};\, A_i)\} = \sum_{i=1}^{k} E\{\chi(\underset{\sim}{W};\, A_i)\} = \sum_{i=1}^{k} P\{W_i \geqq W_{[k]}-d\}$$

$$(4.2) \quad = \sum_{i=1}^{k} P\{W_{(i)} \geqq W_{[k]}-d\} = \sum_{i=1}^{k} P\{W_{(j)} < W_{(i)}+d,\ j=1,2,\ldots,k;\ j\neq i\}$$

$$= \sum_{i=1}^{k} \int_0^\infty \prod_{\substack{j=1\\ j\neq i}}^{k} \{F(n^{\frac{1}{2}}(w+d-\theta_{[j]}))-F(n^{\frac{1}{2}}(-w-d-\theta_{[j]}))\}$$

$$n^{\frac{1}{2}}\{f(n^{\frac{1}{2}}(w-\theta_{[i]}))+f(n^{\frac{1}{2}}(w+\theta_{[i]}))\}dw.$$

If we rename $n^{\frac{1}{2}}w$ as u, $n^{\frac{1}{2}}\theta_{[j]}$ as θ_j' and $n^{\frac{1}{2}}d$ as γ, then (4.2) becomes

$$(4.3) \qquad E\{S\} = \sum_{i=1}^{k} \int_0^\infty \prod_{\substack{j=1\\ j\neq i}}^{k} \{F(u+\gamma-\theta_j')-F(-u-\gamma-\theta_j')\}\{f(u-\theta_i')+f(u+\theta_i')\}du.$$

4.2. <u>Supremum of $E\{S\}$.</u>

In order to find the supremum of $E\{S\}$ we prove the following theorem as a first step.

<u>Theorem 4:</u>

For any fixed $\gamma \geqq 0$,

$$(4.4) \qquad Q = E\{S \,|\, \theta_k' = \theta_{k-1}' = \ldots = \theta_{k-m+1}' = \theta \text{ (say)},\quad 1 \leqq m < k\}$$

is a nonincreasing function of θ for $\theta \geqq 0$. For $\gamma > 0$, it is strictly decreasing in θ for $\theta \geqq 0$.

<u>Proof:</u>

Setting $\theta_k' = \theta_{k-1}' = \ldots = \theta_{k-m+1}' = \theta$ (say), $1 \leqq m < k$, (4.3) gives

$$(4.5) \quad Q = \sum_{i=1}^{k-m} \int_0^\infty \{F(u+\gamma-\theta)-F(-u-\gamma-\theta)\}^m \prod_{\substack{j=1\\ j\neq i}}^{k-m} \{F(u+\gamma-\theta_j')-F(-u-\gamma-\theta_j')\}$$

$$\{f(u-\theta_i')+f(u+\theta_i')\}du$$

$$+ m \int_0^\infty \{F(u+\gamma-\theta)-F(-u-\gamma-\theta)\}^{m-1} \prod_{j=1}^{k-m} \{F(u+\gamma-\theta_j')-F(-u-\gamma-\theta_j')\}$$

$$\{f(u-\theta)+f(u+\theta)\}du.$$

The last term is broken up into two integrals and we substitute $u-\theta=y$ in the first integral and $u+\theta=y$ in the second integral obtaining

$$Q = \sum_{i=1}^{k-m} \int_0^\infty \{F(y+\gamma-\theta)-F(-y-\gamma-\theta)\}^m \prod_{\substack{j=1\\j\neq i}}^{k-m} \{F(y+\gamma-\theta_j')-F(-y-\gamma-\theta_j')\}$$

$$\{f(y-\theta_i')+f(y+\theta_i')\}dy$$

(4.6) $$+ m \int_{-\theta}^\infty \{F(y+\gamma)-F(-y-\gamma-2\theta)\}^{m-1} \prod_{j=1}^{k-m} \{F(y+\gamma+\theta-\theta_j')-F(-y-\gamma-\theta-\theta_j')\}f(y)dy$$

$$+ m \int_{\theta}^\infty \{F(y+\gamma-2\theta)-F(-y-\gamma)\}^{m-1} \prod_{j=1}^{k-m} \{F(y+\gamma-\theta-\theta_j')-F(-y-\gamma+\theta-\theta_j')\}f(y)dy.$$

We now differentiate Q with respect to θ and thus we have

$$\frac{dQ}{d\theta} = m \sum_{i=1}^{k-m} \int_0^\infty \{F(y+\gamma-\theta)-F(-y-\gamma-\theta)\}^{m-1} \prod_{\substack{j=1\\j\neq i}}^{k-m} \{F(y+\gamma-\theta_j')-F(-y-\gamma-\theta_j')\}$$

$$\{f(y-\theta_i')+f(y+\theta_i')\}\{f(y+\gamma+\theta)-f(y+\gamma-\theta)\}dy$$

$$+ 2m(m-1) \int_{-\theta}^\infty \{F(y+\gamma)-F(-y-\gamma-2\theta)\}^{m-2} \prod_{j=1}^{k-m} \{F(y+\gamma+\theta-\theta_j')-F(-y-\gamma-\theta-\theta_j')\}$$

$$f(y+\gamma+2\theta)f(y)dy$$

(4.7) $$+ m \int_{-\theta}^\infty \{F(y+\gamma)-F(-y-\gamma-2\theta)\}^{m-1} \sum_{i=1}^{k-m} \left[\prod_{\substack{j=1\\j\neq i}}^{k-m} \{F(y+\gamma+\theta-\theta_j')-F(-y-\gamma-\theta-\theta_j')\} \right.$$

$$\left. \{f(y+\gamma+\theta-\theta_i')+f(y+\gamma+\theta+\theta_i')\} \right] f(y)dy$$

$$- 2m(m-1) \int_{\theta}^{\infty} \{F(y+\gamma-2\theta)-F(-y-\gamma)\}^{m-2} \prod_{j=1}^{k-m} \{F(y+\gamma-\theta-\theta_j')-F(-y-\gamma+\theta-\theta_j')\}$$
$$f(y+\gamma-2\theta)f(y)dy$$

$$- m \int_{\theta}^{\infty} \{F(y+\gamma-2\theta)-F(-y-\gamma)\}^{m-1} \sum_{i=1}^{k-m} \Bigg[\prod_{\substack{j=1 \\ j\neq i}}^{k-m} \{F(y+\gamma-\theta-\theta_j')-F(-y-\gamma+\theta-\theta_j')\}$$
$$\{f(y+\gamma-\theta-\theta_i')+f(-y-\gamma+\theta-\theta_i')\}\Bigg] f(y)dy.$$

Substituting back $y=u-\theta$ in the second and third integrals and $y=u+\theta$ in the fourth and fifth integrals we obtain, after combining terms,

$$\frac{dQ}{d\theta} = 2m(m-1) \int_{0}^{\infty} \{F(u+\gamma-\theta)-F(-u-\gamma-\theta)\}^{m-1} \prod_{j=1}^{k-m} \{F(u+\gamma-\theta_j')-F(-u-\gamma-\theta_j')\}$$
$$\{f(u+\gamma+\theta)f(u-\theta) - f(u+\gamma-\theta)f(u+\theta)\}du$$

$$(4.8) \qquad + m \sum_{i=1}^{k-m} \int_{0}^{\infty} \{F(u+\gamma-\theta)-F(-u-\gamma-\theta)\}^{m-1} \prod_{\substack{j=1 \\ j\neq i}}^{k-m} \{F(u+\gamma-\theta_j')-F(-u-\gamma-\theta_j')\}$$
$$[\{f(u-\theta_i')+f(u+\theta_i')\}\{f(u+\gamma+\theta)-f(u+\gamma-\theta)\}$$
$$+\{f(u+\gamma-\theta_i')+f(u+\gamma+\theta_i')\}\{f(u-\theta)-f(u+\theta)\}]du.$$

We shall, presently, demonstrate that $\frac{dQ}{d\theta} < 0$ for $\theta > 0$, $\gamma > 0$. Now $f(u+\gamma+\theta)f(u-\theta) \leq f(u+\gamma-\theta)f(u+\theta)$ for every $u > 0$ and $\theta \geqq 0$, $\gamma \geqq 0$ (strict inequality holding for $\theta > 0$, $\gamma > 0$). Thus the first term in (4.8) is negative for $m > 1$, $\theta > 0$, $\gamma > 0$ and is zero when at least one of the equalities $m=1$, $\theta=0$, $\gamma=0$ holds. Next we want to show that the expression

$$(4.9) \qquad L = \{f(u-\theta_i')+f(u+\theta_i')\}\{f(u+\gamma+\theta)-f(u+\gamma-\theta)\}$$
$$+\{f(u+\gamma-\theta_i')+f(u+\gamma+\theta_i')\}\{f(u-\theta)-f(u+\theta)\}$$

is non-positive, that is, we want to show that

$$(4.10)\quad \{f(u-\theta_i')+f(u+\theta_i')\}f(u+\gamma+\theta)+\{f(u+\gamma-\theta_i')+f(u+\gamma+\theta_i')\}f(u-\theta)$$

$$\leq \{f(u-\theta_i')+f(u+\theta_i')\}f(u+\gamma-\theta)+\{f(u+\gamma-\theta_i')+f(u+\gamma+\theta_i')\}f(u+\theta)$$

But this follows immediately as a consequence of the following relations. For $0 \leq \theta_i' \leq \theta$ $(i=1,2,\ldots,k-m)$, $\gamma \geq 0$,

$$(4.11)\qquad f(u+\theta_i')f(u+\gamma+\theta) \leq f(u+\gamma+\theta_i')f(u+\theta),$$

$$(4.12)\qquad f(u+\gamma-\theta_i')f(u-\theta) \leq f(u-\theta_i')f(u+\gamma-\theta),$$

$$(4.13)\qquad f(u+\gamma+\theta_i')f(u-\theta) \leq f(u+\theta_i')f(u+\gamma-\theta),$$

$$(4.14)\qquad f(u-\theta_i')f(u+\gamma+\theta) \leq f(u+\gamma-\theta_i')f(u+\theta).$$

Hence $L \leq 0$ for every $u > 0$ and consequently the second term in (4.8) is negative for $\theta > 0$, $\gamma > 0$ and zero when either $\theta=0$ or $\gamma=0$ or both. Therefore, $\frac{dQ}{d\theta} \leq 0$ for $\theta \geq 0$, $\gamma \geq 0$ (with strict inequality for $\theta > 0$, $\gamma > 0$). It should be noted that $\gamma = n^{\frac{1}{2}}d > 0$ for $\frac{1}{k} < P^* \leq 1$. Thus Theorem 4 is proved.

By this theorem Q has a supremum when $\theta = \theta_{k-m}'$ and, since this holds for any integer $m < k$, the supremum of $E\{S\}$ is given by

$$(4.15)\qquad \sup_{\underset{\sim}{\theta} \in \Omega} E\{S\} = \sup_{\theta \geq 0} k \int_0^\infty \{F(u+\gamma-\theta)-F(-u-\gamma-\theta)\}^{k-1}\{f(u-\theta)+f(u+\theta)\}du$$

$$= \sup_{\theta \geq 0} k\, J(\theta, \gamma)$$

where $J(\theta, \gamma)$ is defined in (3.4). By Theorem 3, $J(\theta, \gamma)$ is strictly decreasing in θ for fixed $\gamma > 0$. It follows that

$$(4.16)\qquad \sup_{\underset{\sim}{\theta} \in \Omega} E\{S\} = 2k \int_0^\infty \{2F(u+\gamma)-1\}^{k-1} f(u)du.$$

Thus, subject to the basic requirement (1.2), the procedure R satisfies the condition that the expected size of the subset retained is bounded above by the right side of (4.16) for all $\underset{\sim}{\theta} \in \Omega$. This bound, however, exceeds kP^*.

4.3. Two secondary problems.

In analogy with power function considerations, one secondary problem is to find the smallest common sample size n necessary to control $E\{S\}$ at some pre-assigned level for a particular alternative in the parameter space; alternatively, it may be desired to control the supremum of $E\{S\}$ over all parameter points in the subspace defined by

$$(4.17) \qquad \Omega(\xi) = \{\underset{\sim}{\theta}: \theta_{[k]}-\theta_{[i]} \geq \xi > 0;\ i=1,2,\ldots,k-1\}$$

where $\Omega(\xi) \subset \Omega$ and $\xi > 0$ is pre-assigned.

The procedure R depends on the common number n of observations from each population; we denote the procedure by R(n). Let $\varepsilon > 0$ be pre-assigned. Then the first secondary problem is to find the smallest n such that for some particular $\underset{\sim}{\theta}_o \in \Omega$,

$$(4.18) \qquad E\{S|k, \underset{\sim}{\theta}_o, P^*, R(n)\} \leq 1+\varepsilon.$$

Thus n is determined as the smallest positive integer greater than or equal to the solution of the equation obtained by putting $\underset{\sim}{\theta} = \underset{\sim}{\theta}_o$ in the right side of (4.3) and equating it to $1+\varepsilon$.

The second problem is to find the smallest n such that

$$(4.19) \qquad \sup_{\underset{\sim}{\theta} \in \Omega(\xi)} E\{S|k, \underset{\sim}{\theta}, P^*, R(n)\} \leq 1+\varepsilon.$$

It can further be shown by a method similar to that used in Theorem 4 that in the subspace $\Omega(\xi)$ of the parameter space the function $E\{S\}$ takes on its supremum when

$$(4.20) \qquad \theta_{[k-1]} = \theta_{[k-2]} = \dots = \theta_{[1]} = \theta^* \text{ (say)}, \quad \theta_{[k]} = \theta^* + \xi.$$

Hence, from (4.3) we can write

$$(4.21) \qquad \sup_{\underset{\sim}{\theta} \in \Omega(\xi)} E\{S\} = \sup_{\theta \geq 0} \Big[\int_0^\infty \{F(u+\gamma-\theta)-F(-u-\gamma-\theta)\}^{k-1}\{f(u-\theta-\lambda)+f(u+\theta+\lambda)\}du$$

$$+ (k-1) \int_0^\infty \{F(u+\gamma-\theta)-F(-u-\gamma-\theta)\}^{k-2}$$

$$\{F(u+\gamma-\theta-\lambda)-F(-u-\gamma-\theta-\lambda)\}\{f(u-\theta)+f(u+\theta)\}du\Big]$$

where $\theta = n^{\frac{1}{2}}\theta^*$, $\gamma = n^{\frac{1}{2}}d$ and $\lambda = n^{\frac{1}{2}}\xi$. We now prove the following theorem.

<u>Theorem 5:</u>

$$B(\theta, \gamma, \lambda) = \int_0^\infty [\{F(u+\gamma-\theta)-F(-u-\gamma-\theta)\}^{k-1}f(u-\theta-\lambda)+(k-1)\{F(u+\gamma-\theta)-F(-u-\gamma-\theta)\}^{k-2}$$

$$\{F(u+\gamma-\theta-\lambda)-F(-u-\gamma-\theta-\lambda)\}f(u-\theta)]du$$

$$(4.22) \qquad + \int_0^\infty [\{F(u+\gamma-\theta)-F(-u-\gamma-\theta)\}^{k-1}f(u+\theta+\lambda)+(k-1)\{F(u+\gamma-\theta)-F(-u-\gamma-\theta)\}^{k-2}$$

$$\{F(u+\gamma-\theta-\lambda)-F(-u-\gamma-\theta-\lambda)\}f(u+\theta)]du$$

is a nonincreasing function of θ for $\theta \geq 0$ and fixed $\gamma \geq 0$ and fixed $\lambda > 0$ and it is strictly decreasing for $\gamma > 0$.

<u>Proof:</u>

Putting $u-\theta=x$ in the first and $u+\theta=x$ in the second integral of (4.22) and taking the partial derivative with respect to θ, we obtain

$$\frac{\partial B}{\partial \theta}(\theta, \gamma, \lambda) = 2(k-1) \int_{-\theta}^\infty \{F(x+\gamma)-F(-x-2\theta-\gamma)\}^{k-2}f(x+2\theta+\gamma)f(x-\lambda)dx$$

$$(4.23) \qquad + 2(k-1)(k-2) \int_{-\theta}^\infty \{F(x+\gamma)-F(-x-2\theta-\gamma)\}^{k-3}\{F(x+\gamma-\lambda)-F(-x-2\theta-\gamma-\lambda)\}$$

$$f(x+2\theta+\gamma)f(x)dx$$

$$+ 2(k-1) \int_{-\theta}^{\infty} \{F(x+\gamma)-F(-x-2\theta-\gamma)\}^{k-2} f(x+2\theta+\gamma+\lambda)f(x)dx$$

$$- 2(k-1) \int_{\theta}^{\infty} \{F(x-2\theta+\gamma)-F(-x-\gamma)\}^{k-2} f(x-2\theta+\gamma)f(x+\lambda)dx$$

$$- 2(k-1)(k-2) \int_{\theta}^{\infty} \{F(x-2\theta+\gamma)-F(-x-\gamma)\}^{k-3}\{F(x-2\theta+\gamma-\lambda)-F(-x-\gamma-\lambda)\} f(x-2\theta+\gamma)f(x)dx$$

$$- 2(k-1) \int_{\theta}^{\infty} \{F(x-2\theta+\gamma)-F(-x-\gamma)\}^{k-2} f(x-2\theta+\gamma-\lambda)f(x)dx.$$

Substituting back $x=u-\theta$ in the first three integrals and $x=u+\theta$ in the last three integrals of (4.23) and then combining the first and sixth integrals, the second and fifth integrals and the third and fourth integrals, it is easily seen that each of the three terms thus formed are non-positive for $\gamma \geqq 0$. Their sum is non-positive for $\gamma \geqq 0$ and negative for $\gamma > 0$. This proves Theorem 5.

Applying this theorem and using (4.20) we conclude that in the subspace $\Omega(\xi)$ the function $E\{S\}$ takes on its supremum when

(4.24) $$\theta_{[k-1]} = \theta_{[k-2]} = \ldots = \theta_{[1]} = 0, \quad \theta_{[k]} = \xi.$$

Also, (4.21) now becomes

(4.25) $$\sup_{\underset{\sim}{\theta} \in \Omega(\xi)} E\{S\} = \int_{0}^{\infty} \{2F(u+\gamma)-1\}^{k-1}\{f(u-\lambda)+f(u+\lambda)\}du$$

$$+ 2(k-1) \int_{0}^{\infty} \{2F(u+\gamma)-1\}^{k-2}\{F(u+\gamma-\lambda)-F(-u-\gamma-\lambda)\}f(u)du$$

where $\gamma = n^{\frac{1}{2}}d$ and $\lambda = n^{\frac{1}{2}}\xi$. Note that the first term on the right side of (4.25) is the $P\{CS|R\}$ for the configuration (4.24). In order to find the smallest n satisfying (4.19) we should equate the right member of (4.25) to $1+\mathcal{E}$ and determine n as the smallest positive integer greater than or equal to the solution of the equation thus formed. Such a value of n must exist because for fixed $\xi > 0$, the right member of (4.25) goes to unity as $n \to \infty$. This is easy to check because as $n \to \infty$, λ goes to infinity ($\xi > 0$ being fixed) and γ obtained as the solution of equation (3.11) is fixed and hence the first term in (4.25) approaches unity while the second term goes to zero.

The expected size of the retained subset is regarded as analogous with the complement of the "power" of the test of a hypothesis and both (4.18) and (4.19) are conditions which insure good power. It is assumed here that both $\underset{\sim}{\theta}_o$ and ε (or both ξ and ε) can be specified by the experimenter.

5. Monotonicity Property of the Procedure R.

We prove the following theorem in a rather general framework.

Theorem 6:

Let π_i $(i=1,2,\ldots,k)$ denote $k \geq 2$ independent populations with c.d.f.'s $F(x_i, \theta_i)$ on the real line and let $\theta_{[1]} \leq \theta_{[2]} \leq \ldots \leq \theta_{[k]}$ be the ordered θ-values. Suppose $F(x, \theta_{[j]}) \leq F(x, \theta_{[i]})$ for all x and $\theta_{[j]} \geq \theta_{[i]}$. Suppose it is desired to select a non-empty subset of the k populations which contains at least one population with parameter $\theta_{[k]}$ and suppose the following procedure R is employed: Retain π_i in the selected subset if and only if $x_i \geq \max\{x_1, x_2, \ldots, x_k\} - d$, where d is a non-negative constant determined in advance of the experimentation. Let q_i denote the probability of including the population with parameter $\theta_{[i]}$ in the subset thus selected. Then $q_j \geq q_i$ for $\theta_{[j]} \geq \theta_{[i]}$.

Proof:

Let $q_{j,i}$ be the probability that the subset includes the population with parameter $\theta_{[j]}$ but does not include the one with parameter $\theta_{[i]}$. Suppose that the random variable associated with the population with parameter $\theta_{[i]}$ is denoted by $X_{(i)}$ and $U = \max\{X_{(\ell)};\ \ell \neq i,j\}$. Let $G(u)$ denote the c.d.f. of U. Then,

$$q_j - q_i = q_{j,i} - q_{i,j} = P\{X_{(i)} < U-d \leq X_{(j)} < U\} + P\{X_{(i)} < X_{(j)}-d,\ X_{(j)} \geq U\}$$

$$-P\{X_{(j)} < U-d \leq X_{(i)} < U\} - P\{X_{(j)} < X_{(i)}-d,\ X_{(i)} \geq U\}$$

$$(5.1)\qquad = \int_{-\infty}^{\infty} F(u-d,\ \theta_{[i]})\{F(u,\ \theta_{[j]})-F(u-d,\ \theta_{[j]})\}dG(u)$$

$$+ \int_{-\infty}^{\infty} F(u-d,\ \theta_{[i]})G(u)dF(u,\ \theta_{[j]})$$

$$- \int_{-\infty}^{\infty} F(u-d,\ \theta_{[j]})\{F(u,\ \theta_{[i]})-F(u-d,\ \theta_{[i]})\}dG(u)$$

$$- \int_{-\infty}^{\infty} F(u-d,\ \theta_{[j]})G(u)dF(u,\ \theta_{[i]}).$$

Let $F_i(x)$ stand for $F(x,\ \theta_{[i]})$. Note that $G(x)F_i(x)$ is a c.d.f.. Let Y_i be a random variable with c.d.f. $G(y)F_i(y)$. Now (5.1) can be written as

$$(5.2)\qquad q_j - q_i = \int_{-\infty}^{\infty} F_i(u-d)d(GF_j)(u) - \int_{-\infty}^{\infty} F_j(u-d)d(GF_i)(u).$$

But

$$(5.3)\qquad \int_{-\infty}^{\infty} F_i(u-d)d(GF_j)(u) = P\{X_{(i)} < Y_j-d\} \geq P\{X_{(j)} < Y_j-d\}$$

$$\geq P\{X_{(j)} < Y_i-d\} = \int_{-\infty}^{\infty} F_j(u-d)d(GF_i)(u)$$

which shows that $q_j - q_i \geqq 0$, thus proving the theorem.

Now this theorem applies to our problem. The statistic W has a monotone likelihood ratio property which, as can easily be checked, implies that $H(w, \theta_{[j]}) \leqq H(w, \theta_{[i]})$ for all w and for $\theta_{[j]} \geqq \theta_{[i]}$; here $H(w, \theta)$ is the c.d.f. of W defined by (2.3) in Chapter I. Thus the proposed procedure R for our problem has the above monotonicity property, namely that, $q_j \geqq q_i$ for $\theta_{[j]} \geqq \theta_{[i]}$, where q_i is the probability of including the population with parameter $\theta_{[i]}$ in the subset retained.

6. Different Variance Set-ups.

Case I: Variances equal with the common value known.

This is the case we have considered. All the k normal populations are assumed to have the same variance σ^2, which we have taken without loss of generality to be unity.

Case II: Variances known and unequal.

If the k normal populations have known variances σ_i^2 $(i=1,2,\ldots,k)$ and if the sample sizes n_i $(i=1,2,\ldots,k)$ have been so chosen as to make the variances of sample means equal, that is, $\frac{\sigma_i^2}{n_i}$ equals a known positive constant c $(i=1,2,\ldots,k)$ then again we can use the proposed procedure R with γ obtained from Table I of Bechhofer [1] and $d = c^{\frac{1}{2}}\nu$.

Case III: Variances equal with the common value unknown.

Let $\sigma_1^2 = \sigma_2^2 = \ldots = \sigma_k^2 = \sigma^2$ (say) where σ^2 is unknown. An unbiased estimate of σ^2 is given by

$$s^2 = \frac{1}{k(n-1)} \sum_{i=1}^{k} \sum_{j=1}^{n} (x_{ij} - \bar{x}_i)^2 \tag{6.1}$$

where $\bar{x}_i = \frac{1}{n}\sum_{j=1}^{n} x_{ij}$ and x_{ij} $(i=1,2,\ldots,k;\ j=1,2,\ldots,n)$ are independent observations from the k normal populations. Now $\frac{\nu^{\frac{1}{2}}s}{\sigma}$ with $\nu = k(n-1)$ is distributed as a chi-chance variable with ν degrees of freedom and is stochastically independent of $\bar{X}_i$ $(i=1,2,\ldots,k)$ and hence also of $W_i = |\bar{X}_i|$ $(i=1,2,\ldots,k)$. We can modify the procedure R defined by (2.2) as follows. Retain π_i in the selected subset if and only if

$$w_i \geq w_{[k]}-sd, \quad (i=1,2,\ldots,k), \tag{6.2}$$

where $d = d(n, k, P^*)$ is chosen as the smallest non-negative value satisfying (1.2) for all the parameter points $(\theta_{[1]}, \theta_{[2]},\ldots,\theta_{[k]}, \sigma^2)$. With $W_{(i)}$ defined as in section 3, we can write

$$P\{CS|R\} = P\{ \frac{n^{\frac{1}{2}}W_{(i)}}{\sigma} < \frac{n^{\frac{1}{2}}W_{(k)}}{\sigma} + \frac{\nu^{\frac{1}{2}}s}{\sigma}\,\frac{n^{\frac{1}{2}}d}{\nu^{\frac{1}{2}}},\ i=1,2,\ldots,k-1\}$$

$$= \int_0^\infty \int_0^\infty \prod_{i=1}^{k-1} \{F(u+\gamma v-\theta_i')-F(-u-\gamma v-\theta_i')\}\{f(u-\theta_k')+f(u+\theta_k')\}g_\nu(v)du dv \tag{6.3}$$

where $\gamma = n^{\frac{1}{2}}d/\nu^{\frac{1}{2}} = n^{\frac{1}{2}}d/\{k(n-1)\}^{\frac{1}{2}}$, $\theta_i' = n^{\frac{1}{2}}\theta_{[i]}/\sigma$ $(i=1,2,\ldots,k)$, F,f are respectively the standard normal c.d.f. and p.d.f. and g_ν is the chi-density with ν degrees of freedom. Now the entire discussion for minimizing $P\{CS|R\}$ in Section 3 goes through and consequently γ is obtained as a solution of

$$\int_0^\infty g_\nu(v)[\int_{-\infty}^\infty F^{k-1}(u+\gamma v)f(u)du]dv = P^*, \tag{6.4}$$

which is equivalent to equation (4.2) of Gupta and Sobel [7] where γ is tabulated for specified values of k, ν and P^*.

7. The Worst Population Problem.

In this section we shall consider the related problem of selecting a subset containing a "worst" population, where a worst population is one which has the smallest absolute value of the mean, namely, $\theta_{[1]}$.

Everything remains the same as in Section 1 except that a correct selection (CS) is now defined as any selection of a subset which contains at least one population with parameter value equal to $\theta_{[1]}$. Using this definition of a correct selection we require a procedure R' so as to satisfy (1.2) with R replaced by R'.

The following procedure R' is proposed. Retain π_i in the selected subset if and only if

$$w_i \leq w_{[1]}+d, \qquad (i=1,2,\ldots,k), \tag{7.1}$$

where the notation is same as in Section 2 and $d = d(n, k, P^*)$ is again a non-negative constant to be determined in advance of the experimentation subject to R' satisfying (1.2).

Again letting $W_{(i)}$ denote the statistic associated with the population with parameter $\theta_{[i]}$, we can write

$$P\{CS|R'\} = P\{W_{(1)} \leq W_{[k]}+d\} = P\{W_{(i)} > W_{(1)}-d,\ i=2,3,\ldots,k\}$$

$$\begin{aligned} = \int_d^\infty \prod_{i=2}^{k} \{F(n^{\frac{1}{2}}(-w+d+\theta_{[i]}))+F(n^{\frac{1}{2}}(-w+d-\theta_{[i]}))\} \\ n^{\frac{1}{2}}\{f(n^{\frac{1}{2}}(w-\theta_{[1]}))+f(n^{\frac{1}{2}}(w+\theta_{[1]}))\}dw \\ + n^{\frac{1}{2}} \int_0^d \{f(n^{\frac{1}{2}}(w-\theta_{[1]}))+f(n^{\frac{1}{2}}(w+\theta_{[1]}))\}dw. \end{aligned} \tag{7.2}$$

Here $F(n^{\frac{1}{2}}(-w+d+\theta_{[i]}))+F(n^{\frac{1}{2}}(-w+d-\theta_{[i]}))$ is a strictly increasing function of $\theta_{[i]}$ $(i=2,3,\ldots,k)$ for $w > d$ and hence as a first step for minimizing (7.2) we can take the configuration given by (3.2). With (3.2) and renaming $n^{\frac{1}{2}}w$ as u, $n^{\frac{1}{2}}\theta^*$ as θ and $n^{\frac{1}{2}}d$ as γ we can write

$$\inf_{\underset{\sim}{\theta} \in \Omega} P\{CS|R'\} = \inf_{\theta \geq 0} \left[\int_{\gamma}^{\infty} \{F(-u+\gamma+\theta)+F(-u+\gamma-\theta)\}^{k-1}\{f(u-\theta)+f(u+\theta)\}du + \int_{0}^{\gamma} \{f(u-\theta)+f(u+\theta)\}du \right]$$

$$(7.3) \qquad = \inf_{\theta \geq 0} \left[\int_{\gamma}^{\infty} \{F(-u+\gamma+\theta)+F(-u+\gamma-\theta)\}^{k-1} f(u-\theta)du + \int_{\gamma}^{\infty} \{F(-u+\gamma+\theta)+F(-u+\gamma-\theta)\}^{k-1} f(u+\theta)du + F(\gamma-\theta)+F(\gamma+\theta)-1 \right].$$

Now a proof similar to that of Theorem 3 shows that the quantity in the square brackets in (7.3) is a nonincreasing function of θ for fixed $\gamma \geqq 0$. Thus the least favorable configuration of the parameters for this problem is the same as the configuration (3.9). Hence,

$$\inf_{\underset{\sim}{\theta} \in \Omega} P\{CS|R'\} = \lim_{\theta \to \infty} \left[\int_{\gamma-\theta}^{\infty} \{F(-x+\gamma)+F(-x+\gamma-2\theta)\}^{k-1} f(x)dx + \int_{\gamma+\theta}^{\infty} \{F(-x+\gamma+2\theta)+F(-x+\gamma)\}^{k-1} f(x)dx + F(\gamma-\theta)+F(\gamma+\theta)-1 \right]$$

$$(7.4) \qquad = \int_{-\infty}^{\infty} F^{k-1}(x+\gamma)f(x)dx$$

which is the same as (3.10). The quantity $\gamma = n^{\frac{1}{2}}d$ is obtained as the solution of (3.11) and is tabulated by Bechhofer [1, Table I] with t=1.

Other discussions for this problem are similar to those of the original problem and hence will be omitted.

8. Directions of Certain Generalizations.

We now indicate two different directions for generalizing the problem considered in this chapter. One direction is to allow unequal number of observations from the k given normal populations with unknown means and the same unit variance. The other generalization deals with the problem of selecting a subset containing the t best populations, that is, the populations with the t largest absolute values of the means for $t \geq 1$. These problems are not treated here.

CHAPTER IV: SOME REMARKS ON SUBSET FORMULATION FOR POPULATIONS WITH MONOTONE LIKELIHOOD RATIO

When a ranking and selection procedure is based on a statistic that has the monotone likelihood ratio property certain simplifying results can be obtained. This has been demonstrated in the discussion of Theorem 1 and the properties considered in Section 7, Chapter II for the indifference zone formulation. We now give some results for the subset formulation.

Let $p(x_i, \theta_i)$ be a family of densities on the real line with monotone likelihood ratio in x_i $(i=1,2,\ldots,k)$. Suppose $X_1, X_2, \ldots, X_k$ are independently distributed with density $p(x_1, \theta_1), p(x_2, \theta_2), \ldots, p(x_k, \theta_k)$ respectively. Let the population with density $p(x_i, \theta_i)$ be denoted by π_i $(i=1,2,\ldots,k)$ and let $\theta_{[1]} \leq \theta_{[2]} \leq \ldots \leq \theta_{[k]}$ be the ordered θ-values. We shall consider the problem of selecting a non-empty subset containing the population that has the largest (smallest) parameter and call it problem 1 (problem 2). Any selection of a subset containing at least one population with parameter value equal to $\theta_{[k]}(\theta_{[1]})$ is regarded as a correct selection (CS) for problem 1 (problem 2). Consider the procedures $R_{(1)}$ and $R_{(2)}$ for problem 1 and procedures $R'_{(1)}$ and $R'_{(2)}$ for problem 2 defined respectively as follows. Retain π_i $(i=1,2,\ldots,k)$ in the selected subset if and only if

(1) $$x_i \geq x_{[k]} - d_1$$

(2) $$x_i \geq d_2 x_{[k]}$$

(3) $$x_i \leq x_{[1]} + d'_1$$

(4) $$x_i \leq x_{[1]}/d'_2$$

where $x_{[1]} \leq x_{[2]} \leq \ldots \leq x_{[k]}$ are the ordered values of $x_1, x_2, \ldots, x_k$

and $d_1 \geq 0$, $0 < d_2 \leq 1$, $d_1' \geq 0$ and $0 < d_2' \leq 1$ are the constants determined in advance of the experimentation for procedures $R_{(1)}$, $R_{(2)}$, $R_{(1)}'$ and $R_{(2)}'$ respectively. With these hypotheses we can state the following theorem.

Theorem 7:

$P\{CS|R_{(j)}\}(j=1,2)$ is a nonincreasing function of $\theta_{[\alpha]}(\alpha=1,2,\ldots,k-1)$ and a nondecreasing function of $\theta_{[k]}$. Also $P\{CS|R_{(j)}'\}(j=1,2)$ is a nondecreasing function of $\theta_{[\beta]}(\beta=2,3,\ldots,k)$ and a nonincreasing function of $\theta_{[1]}$.

Proof:

Let $X_{(i)}$ denote the statistic from the population with parameter $\theta_{[i]}(i=1,2,\ldots,k)$. Define random variables $\psi_j(j=1,2)$ and $\psi_j'(j=1,2)$ as follows:

(5) $$\psi_1 = \begin{cases} 1 \text{ if } X_{(k)} \geq X_{[k]} - d_1 \\ 0, \quad \text{otherwise.} \end{cases}$$

(6) $$\psi_2 = \begin{cases} 1 \text{ if } X_{(k)} \geq d_2 X_{[k]} \\ 0, \quad \text{otherwise} \end{cases}$$

(7) $$\psi_1' = \begin{cases} 1 \text{ if } X_{(1)} \leq X_{[1]} + d_1' \\ 0, \quad \text{otherwise} \end{cases}$$

(8) $$\psi_2' = \begin{cases} 1 \text{ if } X_{(1)} \leq X_{[1]}/d_2' \\ 0, \quad \text{otherwise} \end{cases}$$

Then $E(\psi_j) = P\{CS|R_{(j)}\}$ and $E(\psi_j') = P\{CS|R_{(j)}'\}$, $(j=1,2)$. It should be noted that each of ψ_1, ψ_2, ψ_1', and ψ_2' is a function of $X_{(i)}(i=1,2,\ldots,k)$. It is easy to see that for each $\alpha(\alpha=1,2,\ldots,k-1)$ ψ_j $(j=1,2)$ is a nonincreasing function of $X_{(\alpha)}$ holding all $X_{(i)}(i\neq\alpha)$ fixed and is a nondecreasing function of $X_{(k)}$ holding all $X_{(i)}(i \neq k)$ fixed. Similarly for each $\beta(\beta=2,3,\ldots,k)$ $\psi_j'(j=1,2)$ is a nondecreasing function of $X_{(\beta)}$ holding all $X_{(i)}(i \neq \beta)$ fixed and is a nonincreasing function of $X_{(1)}$ holding all $X_{(i)}(i \neq 1)$ fixed. Now Lemma 2 applies and the theorem is proved.

Replacing the monotone likelihood ratio requirement in the hypotheses of Theorem 7 by a less stringent requirement, namely, that $F(x,\theta_{[m]}) \leq F(x,\theta_{[i]})$ for all x and $\theta_{[m]} \geq \theta_{[i]}$, we can state a theorem about the monotonicity properties of the procedures $R_{(j)}$ and $R'_{(j)}$ (j=1,2) given by (1)-(4). It should be noted that the above monotonicity requirement of the c.d.f. is implied by the monotone likelihood ratio property.

Theorem 8:

Let $q_i(q_i')$ denote the probability of including the population with parameter $\theta_{[i]}$ in the subset selected by procedure $R_{(j)}$ ($R'_{(j)}$), j=1,2. Then $q_m \geq q_i$ and $q_m' \leq q_i'$ for $\theta_{[m]} \geq \theta_{[i]}$.

This theorem is the same as Theorem 6 when $R_{(1)}$ is the procedure employed. Proof for the cases when $R_{(2)}$, $R'_{(1)}$ and $R'_{(2)}$ are used is similar and shall be omitted.

TABLE I

Probability of selecting from a set of k normal populations with unit variance the one whose mean has the largest absolute value, for specified values of $\lambda = n^{\frac{1}{2}}\delta^*$.

λ	k = 2	k = 3	k = 4	k = 5	k = 6	k = 7	k = 8	k = 9	k = 10
0.0	.500000	.333333	.250000	.200000	.166667	.142857	.125000	.111111	.100000
0.1	.501589	.335169	.251836	.201772	.168359	.144469	.126535	.112575	.101399
0.2	.506324	.340644	.257317	.207068	.173420	.149291	.131131	.116962	.105593
0.3	.514111	.349668	.266368	.215826	.181801	.157287	.138760	.124249	.112567
0.4	.524798	.362094	.278865	.227946	.193422	.168394	.149374	.134404	.122298
0.5	.538178	.377719	.294635	.243287	.208172	.182524	.162905	.147373	.134747
0.6	.553998	.396293	.313466	.261675	.225907	.199562	.179262	.163088	.149863
0.7	.571965	.417526	.335105	.282897	.246454	.219369	.198334	.181460	.167579
0.8	.591760	.441093	.359267	.306713	.269613	.241776	.219984	.202378	.187805
0.9	.613041	.466642	.385638	.332854	.295153	.266592	.244049	.225707	.210431
1.0	.635460	.493808	.413886	.361025	.322822	.293598	.270343	.251289	.235323
1.1	.658667	.522216	.443661	.390917	.352344	.322553	.298655	.278939	.262319
1.2	.682321	.551490	.474609	.422203	.383427	.353195	.328750	.308448	.291234
1.3	.706101	.581264	.506372	.454551	.415764	.385242	.360374	.339584	.321857
1.4	.729707	.611189	.538600	.487626	.449040	.418401	.393251	.372092	.353950
1.5	.752871	.640937	.570955	.521094	.482934	.452367	.427093	.405700	.387259
1.6	.775357	.670206	.603116	.554634	.517130	.486832	.461604	.440122	.421508
1.7	.796965	.698729	.634786	.587936	.551317	.521488	.496481	.475062	.456411
1.8	.817531	.726273	.665696	.620713	.585197	.556034	.531422	.510224	.491673
1.9	.836931	.752638	.695606	.652699	.618489	.590180	.566134	.545308	.526997
2.0	.855072	.777667	.724311	.683658	.650937	.623653	.600333	.580028	.562092
2.1	.871898	.801236	.751639	.713384	.682307	.656204	.633757	.614109	.596674
2.2	.887380	.823258	.777454	.741703	.712399	.687608	.666162	.647295	.630478
2.3	.901519	.843680	.801656	.768475	.741041	.717670	.697334	.679355	.663260
2.4	.914336	.862479	.824177	.793596	.768096	.746227	.727089	.710086	.694800
2.5	.925872	.879661	.844982	.816992	.793462	.773147	.755272	.739316	.724910
2.6	.936185	.895254	.864063	.838623	.817067	.798336	.781765	.766904	.753433
2.7	.945344	.909309	.881440	.858479	.838872	.821728	.806481	.792745	.780244
2.8	.953423	.921892	.897157	.876576	.858869	.843293	.829368	.816768	.805256
2.9	.960507	.933083	.911273	.892954	.877077	.863027	.850405	.838933	.828412
3.0	.966680	.942972	.923867	.907672	.893537	.880957	.869599	.859232	.849690
3.1	.972025	.951653	.935027	.920809	.908313	.897129	.886984	.877686	.869095
3.2	.976628	.959227	.944852	.932453	.921484	.911613	.902616	.894339	.886663
3.3	.980568	.965793	.953444	.942705	.933143	.924492	.916573	.909257	.902450
3.4	.983922	.971450	.960910	.951671	.943393	.935865	.928944	.922526	.916533
3.5	.986760	.976295	.967356	.959460	.952343	.945838	.939832	.934242	.929006
3.6	.989150	.980420	.972886	.966182	.960105	.954524	.949349	.944514	.939972

TABLE I (continued)

λ	k = 2	k = 3	k = 4	k = 5	k = 6	k = 7	k = 8	k = 9	k = 10
3.7	.991151	.983910	.977600	.971946	.966791	.962035	.957608	.953457	.949545
3.8	.992816	.986845	.981594	.976856	.972513	.968488	.964726	.961188	.957843
3.9	.994196	.989300	.984957	.981012	.977377	.973994	.970820	.967825	.964985
4.0	.995333	.991342	.987770	.984507	.981485	.978661	.976001	.973484	.971089
4.1	.996265	.993029	.990111	.987428	.984932	.982590	.980377	.978275	.976271
4.2	.997025	.994416	.992046	.989854	.987806	.985876	.984047	.982304	.980638
4.3	.997641	.995550	.993636	.991857	.990187	.988607	.987105	.985670	.984295
4.4	.998139	.996471	.994935	.993500	.992147	.990862	.989637	.988464	.987336
4.5	.998538	.997216	.995990	.994839	.993750	.992712	.991719	.990766	.989847
4.6	.998857	.997815	.996842	.995925	.995053	.994220	.993421	.992651	.991908
4.7	.999111	.998293	.997526	.996799	.996106	.995442	.994802	.994185	.993588
4.8	.999312	.998673	.998072	.997499	.996951	.996425	.995917	.995425	.994949
4.9	.999470	.998974	.998505	.998057	.997626	.997212	.996810	.996422	.996044
5.0	.999593	.999211	.998847	.998498	.998162	.997837	.997522	.997217	.996919
5.1	.999689	.999396	.999115	.998845	.998584	.998332	.998086	.997847	.997614
5.2	.999764	.999540	.999324	.999117	.998915	.998720	.998530	.998344	.998163
5.3	.999822	.999651	.999487	.999328	.999174	.999023	.998877	.998733	.998593
5.4	.999866	.999737	.999612	.999491	.999374	.999259	.999147	.999037	.998929
5.5	.999899	.999802	.999708	.999617	.999528	.999440	.999355	.999271	.999189
5.6	.999925	.999852	.999782	.999713	.999646	.999580	.999515	.999452	.999389
5.7	.999944	.999890	.999838	.999786	.999736	.999686	.999638	.999590	.999543
5.8	.999959	.999919	.999880	.999841	.999804	.999767	.999731	.999695	.999659
5.9	.999970	.999940	.999911	.999883	.999855	.999828	.999801	.999774	.999748
6.0	.999978	.999956	.999935	.999914	.999894	.999873	.999853	.999834	.999814
6.1	.999984	.999968	.999953	.999937	.999922	.999907	.999893	.999878	.999864
6.2	.999988	.999977	.999966	.999954	.999943	.999933	.999922	.999911	.999901
6.3	.999992	.999983	.999975	.999967	.999959	.999951	.999943	.999936	.999928
6.4	.999994	.999988	.999982	.999976	.999971	.999965	.999959	.999954	.999948
6.5	.999996	.999991	.999987	.999983	.999979	.999975	.999971	.999967	.999963
6.6	.999997	.999994	.999991	.999988	.999985	.999982	.999979	.999976	.999973
6.7	.999998	.999996	.999994	.999991	.999989	.999987	.999985	.999983	.999981
6.8	.999998	.999997	.999995	.999994	.999993	.999991	.999990	.999988	.999987
6.9	.999999	.999998	.999997	.999996	.999995	.999994	.999993	.999992	.999991
7.0	.999999	.999999	.999998	.999997	.999996	.999996	.999995	.999994	.999993

TABLE II

Value of $\lambda = n^{\frac{1}{2}}\delta^*$ needed to determine the sample size n so that the probability is at least P* that the rule R_1 will lead to a correct selection.

k	P*=.5000	P*=.7500	P*=.9000	P*=.9500	P*=.9750	P*=.9900	P*=.9950	P*=.9990	P*=.9995	P*=.9999
2	0.000000	1.487464	2.288787	2.756050	3.162856	3.640308	3.968602	4.653308	4.922438	5.502111
3	1.022126	1.889770	2.632586	3.079897	3.471675	3.932213	4.248950	4.909823	5.169847	5.730795
4	1.280091	2.093848	2.819277	3.258250	3.642843	4.094815	4.405596	5.054111	5.309388	5.860497
5	1.437058	2.230361	2.946501	3.380284	3.760246	4.206648	4.513549	5.153992	5.406153	5.950761
6	1.549959	2.332450	3.042374	3.472441	3.849054	4.291419	4.595504	5.230075	5.479960	6.019792
7	1.638013	2.413649	3.118944	3.546153	3.920182	4.359428	4.661337	5.291355	5.539468	6.075567
8	1.710065	2.480841	3.182474	3.607385	3.979333	4.416067	4.716222	5.342558	5.589232	6.122290
9	1.770938	2.538004	3.236627	3.659634	4.029855	4.464503	4.763199	5.386467	5.631939	6.162446
10	1.823557	2.587648	3.283731	3.705121	4.073876	4.506753	4.804207	5.424859	5.669303	6.197623

TABLE III

Probability of selecting from a set of k normal populations with unit variance the one whose mean has the smallest absolute value, for specified values of $\lambda = n^{\frac{1}{2}}\delta^*$.

λ	k = 2	k = 3	k = 4	k = 5	k = 6	k = 7	k = 8	k = 9	k = 10
0.0	.500000	.333333	.250000	.200000	.166667	.142857	.125000	.111111	.100000
0.1	.501589	.334679	.251100	.200918	.167451	.143539	.125603	.111651	.100489
0.2	.506324	.338719	.254415	.203690	.169819	.145602	.127428	.113285	.101968
0.3	.514111	.345462	.259983	.208362	.173821	.149093	.130518	.116056	.104476
0.4	.524798	.354913	.267866	.215013	.179538	.154090	.134949	.120032	.108080
0.5	.538178	.367067	.278139	.223746	.187079	.160703	.140826	.125314	.112873
0.6	.553998	.381896	.290879	.234682	.196581	.169070	.148284	.132032	.118979
0.7	.571965	.399338	.306153	.247948	.208198	.179355	.157487	.140345	.126552
0.8	.591760	.419286	.324006	.263669	.222093	.191739	.168623	.150442	.135776
0.9	.613041	.441584	.344440	.281947	.238427	.206413	.181897	.162534	.146863
1.0	.635460	.466021	.367404	.302848	.257339	.223563	.197523	.176848	.160045
1.1	.658667	.492332	.392785	.326381	.278930	.243350	.215702	.193612	.175568
1.2	.682321	.520208	.420398	.352488	.303241	.265893	.236608	.213039	.193671
1.3	.706101	.549298	.449988	.381029	.330239	.291245	.260362	.235303	.214568
1.4	.729707	.579229	.481236	.411782	.359802	.319374	.287012	.260515	.238422
1.5	.752871	.609613	.513768	.444442	.391712	.350152	.316509	.288698	.265318
1.6	.775357	.640066	.547171	.478632	.425658	.383343	.348695	.319769	.295238
1.7	.796965	.670219	.581012	.513922	.461250	.418615	.383304	.353527	.328048
1.8	.817531	.699730	.614855	.549845	.498029	.455544	.419959	.389650	.363484
1.9	.836931	.728296	.648280	.585922	.535498	.493639	.458195	.427712	.401164
2.0	.855072	.755655	.680899	.621682	.573141	.532368	.497482	.467199	.440600
2.1	.871898	.781595	.712367	.656684	.610452	.571184	.537255	.507539	.481228
2.2	.887380	.805953	.742394	.690534	.646957	.609557	.576942	.548138	.522439
2.3	.901519	.828617	.770746	.722894	.682232	.646994	.615998	.588409	.563618
2.4	.914336	.849517	.797250	.753491	.715918	.683061	.653926	.627805	.604176
2.5	.925872	.868629	.821791	.782120	.747725	.717393	.690295	.665837	.643576
2.6	.936185	.885964	.844308	.808642	.777438	.749705	.724756	.702097	.681354
2.7	.945344	.901565	.864788	.832979	.804915	.779790	.757041	.736259	.717133
2.8	.953423	.915500	.883260	.855112	.830080	.807517	.786964	.768087	.750629
2.9	.960507	.927857	.899789	.875066	.852917	.832825	.814420	.797430	.781646
3.0	.966680	.938738	.914466	.892908	.873460	.855713	.839371	.824214	.810071
3.1	.972025	.948253	.927401	.908735	.891787	.876233	.861840	.848431	.835870
3.2	.976628	.956521	.938720	.922669	.908005	.894477	.881899	.870133	.859068
3.3	.980568	.963657	.948556	.934846	.922247	.910566	.899658	.889414	.879746
3.4	.983922	.969778	.957047	.945411	.934661	.924647	.915257	.906404	.898022
3.5	.986760	.974997	.964327	.954516	.945405	.936879	.928852	.921259	.914046
3.6	.989150	.979419	.970530	.962309	.954637	.947427	.940614	.934148	.927986

TABLE III (continued)

λ	k = 2	k = 3	k = 4	k = 5	k = 6	k = 7	k = 8	k = 9	k = 10
3.7	.991151	.983145	.975783	.968936	.962517	.956461	.950718	.945249	.940023
3.8	.992816	.986265	.980202	.974535	.969199	.964144	.959335	.954742	.950340
3.9	.994196	.988863	.983899	.979236	.974827	.970636	.966635	.962803	.959121
4.0	.995333	.991015	.986973	.983159	.979538	.976085	.972778	.969602	.966543
4.1	.996265	.992787	.989514	.986412	.983457	.980629	.977914	.975298	.972773
4.2	.997025	.994237	.991602	.989094	.986697	.984395	.982179	.980039	.977969
4.3	.997641	.995419	.993309	.991293	.989359	.987497	.985700	.983960	.982273
4.4	.998139	.996376	.994695	.993084	.991533	.990036	.988587	.987182	.985816
4.5	.998538	.997147	.995816	.994535	.993298	.992102	.990941	.989813	.988714
4.6	.998857	.997765	.996716	.995703	.994723	.993773	.992848	.991948	.991070
4.7	.999111	.998258	.997435	.996639	.995867	.995116	.994384	.993670	.992972
4.8	.999312	.998649	.998007	.997385	.996780	.996190	.995614	.995051	.994499
4.9	.999470	.998957	.998459	.997975	.997504	.997043	.996592	.996151	.995718
5.0	.999593	.999199	.998815	.998440	.998075	.997717	.997366	.997023	.996685
5.1	.999689	.999387	.999093	.998805	.998523	.998247	.997976	.997709	.997447
5.2	.999764	.999534	.999309	.999089	.998872	.998660	.998452	.998247	.998045
5.3	.999822	.999647	.999476	.999308	.999144	.998982	.998822	.998665	.998510
5.4	.999866	.999734	.999605	.999478	.999353	.999230	.999109	.998989	.998871
5.5	.999899	.999801	.999703	.999608	.999514	.999421	.999329	.999238	.999149
5.6	.999925	.999851	.999779	.999707	.999636	.999566	.999497	.999429	.999362
5.7	.999944	.999890	.999835	.999782	.999729	.999677	.999626	.999574	.999524
5.8	.999959	.999918	.999878	.999839	.999800	.999761	.999722	.999684	.999647
5.9	.999970	.999940	.999910	.999881	.999852	.999824	.999795	.999767	.999739
6.0	.999978	.999956	.999934	.999913	.999892	.999871	.999850	.999829	.999808
6.1	.999984	.999968	.999952	.999937	.999921	.999906	.999890	.999875	.999860
6.2	.999988	.999977	.999965	.999954	.999943	.999931	.999920	.999909	.999898
6.3	.999992	.999983	.999975	.999967	.999959	.999950	.999942	.999934	.999926
6.4	.999994	.999988	.999982	.999976	.999970	.999964	.999959	.999953	.999947
6.5	.999996	.999991	.999987	.999983	.999979	.999975	.999970	.999966	.999962
6.6	.999997	.999994	.999991	.999988	.999985	.999982	.999979	.999976	.999973
6.7	.999998	.999996	.999994	.999991	.999989	.999987	.999985	.999983	.999981
6.8	.999998	.999997	.999995	.999994	.999992	.999991	.999989	.999988	.999987
6.9	.999999	.999998	.999997	.999996	.999995	.999994	.999993	.999992	.999991
7.0	.999999	.999999	.999998	.999997	.999996	.999996	.999995	.999994	.999993

TABLE IV

Value of $\lambda = n^{\frac{1}{2}}\delta^*$ needed to determine the sample size n so that the probability is at least P* that the rule R_{k-1} will lead to a correct selection.

k	P*=.5000	P*=.7500	P*=.9000	P*=.9500	P*=.9750	P*=.9900	P*=.9950	P*=.9990	P*=.9995	P*=.9999
2	0.000000	1.487464	2.288787	2.756050	3.162856	3.640308	3.968602	4.653308	4.922438	5.502111
3	1.128021	1.978923	2.689456	3.119966	3.500064	3.950437	4.262106	4.916175	5.174540	5.733166
4	1.458080	2.226226	2.901354	3.315981	3.683977	4.121571	4.425145	5.063828	5.316652	5.864255
5	1.660837	2.388269	3.043291	3.448496	3.809133	4.238782	4.537235	5.166007	5.415208	5.955525
6	1.805293	2.507409	3.149065	3.547838	3.903368	4.327421	4.622222	5.243839	5.490396	6.025355
7	1.916499	2.600946	3.232894	3.626908	3.978611	4.398425	4.690436	5.306532	5.551033	6.081796
8	2.006351	2.677572	3.302055	3.692357	4.041048	4.457495	4.747274	5.358919	5.601751	6.129092
9	2.081404	2.742246	3.360756	3.748054	4.094288	4.507970	4.795905	5.403849	5.645285	6.169751
10	2.145637	2.798052	3.411641	3.796439	4.140616	4.551969	4.838343	5.443138	5.683381	6.205378

REFERENCES

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

[2] BECHHOFER, R. E., DUNNETT, C. W. and SOBEL, M. (1954). A two-stage multiple decision procedure for ranking means of normal populations with a common unknown variance. Biometrika 41 170-176.

[3] BECHHOFER, R. E. and SOBEL, M. (1954). A single-sample multiple decision procedure for ranking variances of normal populations. Ann. Math. Statist. 25 273-289.

[4] BLACKWELL, D. and GIRSHICK, M. A. (1954). Theory of Games and Statistical Decisions. John Wiley and Sons, New York.

[5] ELANDT, R. C. (1961). The folded normal distribution: two methods of estimating parameters from moments. Technometrics 3 551-562.

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

[7] GUPTA, S. S. and SOBEL, M. (1957). On a statistic which arises in selection and ranking problems. Ann. Math. Statist. 28 957-967.

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

[9] HALL, W. J. (1958). Most economical multiple-decision rules. Ann. Math. Statist. **29** 1079-1094.

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

[11] JOHNSON, N. L. (1962). The folded normal distribution III: accuracy of estimation by maximum likelihood. Technometrics **4** 249-258.

[12] KARLIN, S. (1957). Polya type distributions, II. Ann. Math. Statist. **28** 281-308.

[13] KARLIN, S. and TRUAX, D. (1960). Slippage problems. Ann. Math. Statist. **31** 296-324.

[14] LEHMANN, E. L. (1950). Notes on the theory of Estimation. Mimeo. Lecture notes, Univ. of California.

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

[16] LEONE, F. C., NELSON, L. S. and NOTTINGHAM, R. B. (1961). The folded normal distribution. Technometrics **3** 543-550.

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

[18] SAVAGE, I. R. (1956). Contributions to the theory of rank order statistics--the two-sample case. Ann. Math. Statist. **27** 590-615.

[19] TEICHROEW, D. (1955). Probabilities Associated with Order Statistics in Samples from Two Normal Populations with Equal Variance. Chemical Corps Engineering Agency, Army Chemical Center, Maryland.

[20] WALD, A. (1950). Statistical Decision Functions. John Wiley and Sons, New York.