February 1966

ON A GENERALIZED GOAL IN FIXED-SAMPLE RANKING AND SELECTION PROBLEMS*

Desu M. Mahamunulu

Technical Report No. 72

University of Minnesota
Minneapolis, Minnesota

*This research was supported by National Science Foundation Grant No. NSF-GP-3813.

TABLE OF CONTENTS

Introduction and Summary

During the last fifteen years considerable research has been devoted to developing multiple decision procedures for ranking and selection problems. These problems are concerned with ranking a subset of the k given populations or with selection of certain subsets from the given set of k populations; the particular goal of interest being defined in terms of the unknown ordered values of the population parameters. The parameter of interest may be the mean or the variance or any other explicit or implicit function of the given population. These procedures can be studied for their own interest and/or they can be regarded as alternatives to (classical) tests of homogeneity which are found to be unrealistic in some situations and inadequate for many purposes.

Some of the significant contributions towards developing single-sample procedures that deal with the above type of problems are due to Mosteller (1948), Paulson (1949, 1952), Bahadur (1950), Bahadur and Robbins (1950), Bahadur and Goodman (1952), Bechhofer (1954), Bechhofer and Sobel (1954), Bechhofer, Dunnett and Sobel (1954), Gupta (1956), Gupta and Sobel (1957, 1958, 1962), Seal (1955, 1958), Hall (1959), Lehmann (1961) and Sobel (1963). At present we have a vast literature on this type of problems.

Generally, in ranking and selection problems, populations with large or small parameter values are of interest. The decision procedures are usually designed to select subsets of fixed or random size, such that the selected subset includes the t best populations; here the t best populations are those with the t largest parameter values (or perhaps those with the t smallest parameter values). In this investigation we are concerned with the problem of selecting subsets of specified or fixed size, from a given set of k populations. We are interested only in single-sample procedures that achieve a particular goal of interest.

Part I of this investigation consists of three chapters and it deals with the selection of subsets of specified size s, where $s < k$. The goal of interest, which is called Goal I, is the selection of a subset of size s which contains at least c of the t best populations. The t best populations are those with largest values for the parameter, which is of interest. Two particular cases of Goal I are of special interest. They are -- Goal 1: To select a subset of size s which includes the t best populations, where $s \geq t$; Goal 2: To select a subset of size s which includes any s of the t best populations, where $s \leq t$. It should be noted that all the three goals coincide when $c = s = t$. Then the common goal is to select the t best populations without ordering, which has been extensively studied in relation to various populations. Goal 2 has been suggested by Sobel (see the footnote in Bechhofer (1954)) in relation to the means of normal populations; but no detailed investigation about the procedures that achieve this goal is available in the literature. By considering the complimentary subset that is not selected, the problem in relation to Goal I is related to the corresponding problem where the t best populations are those with smallest parameter values.

In chapter I we formulate the above problem along the lines of Bechhofer (1954) and Bechhofer and Sobel (1954) and a preference zone in the parameter space is pre-assigned. A single-sample procedure for selecting the subset of interest has been proposed. The common number of observations needed from each population for this procedure, so as to guarantee a pre-assigned probability of achieving the goal, is determined. In the general discussion the particular distributions which characterize the populations are not specified; we merely assume that the statistics on which the proposed procedure is based are such that their distribution functions form a stochastically increasing family, when indexed by the parameter of interest.

Using this assumption we determine the common number of observations required per population, as a function of the underlying distributions. These results are applied to specific populations in Chapter II.

Chapter III deals with some properties of the single-sample procedure proposed to achieve the above goal. In particular an unbiased property of the procedure with respect to the parameter values has been proved. Further, considering the class of impartial decision rules along the lines of Bahadur (1950), it has been shown that the proposed procedure is uniformly best in the class of impartial decision rules with respect to a particular loss function. This result has been proved under the assumption that the densities of the statistics, on which the decision rules are based, possess monotone likelihood ratio property.

Part II of this investigation deals with the problem of selecting a subset of fixed size, where the subset size is to be determined as a function of the size of the sample taken from each of the k given populations. The subset of interest is a subset which includes the t best populations. The solution to this problem is closely related to the solution of the problem of selecting a subset of specified size which is treated in part I. Here also a preference zone in the parameter space is pre-assigned. The subset size is chosen so that the probability that it will contain the t best populations is not less than a pre-assigned number P* for all parameter points in the preference zone; an application for this type of problem is given. Proofs of certain monotone properties of the sample size needed to achieve Goal I and its particular cases are included and these properties are used to solve this problem.

Part I

A Problem Dealing with the Selection of Subsets of Specified Size

Chapter I

Formulation of the Problem and the Solution

1.1 Statement of the problem.

We have at our disposal $k \geq 2$ populations $\prod_1, \prod_2, \ldots, \prod_k$ is characterized by a scalar measure θ. The population $\prod(\theta)$ generates independent random variables $X_1, X_2, \ldots$, each X having the same distribution function $P(X \leq x) = F(x|\theta)$ say, and a set of X's which have been generated by $\prod$ is called a sample from the population. That is, the k given populations $\prod_1, \prod_2, \ldots, \prod_k$ are such that $\prod_i$ is characterized by the distribution function $F(x|\theta_i)$, $i = 1,2,\ldots,k$. We assume that the functional form of F is known. We also assume that the values of the real valued parameters θ_i are unknown; but it is assumed that θ_i belong to a space Θ, where Θ is a finite or infinite interval. Let $\theta_{[1]} \leq \theta_{[2]} \leq \ldots \leq \theta_{[k]}$ be the ordered θ_i. We assume that it is not known which population is associated with $\theta_{[i]}$, $i = 1,2,\ldots,k$.

Let c, s and t be integers such that $\max(1,s+t+1-k) \leq c \leq \min(s,t)$; this implies that $\max(s,t) \leq k-1$. From the given set of k populations, the experimenter is interested in the goal of selecting a subset of fixed size s which should contain a certain subset of the t "best" populations. The t best populations are those t, whose parameter values are (say) the largest. Here we consider the following goal.

Goal I. To select a subset of size s which contains at least c of the t best populations.

The problem is to devise a procedure with small fixed sample size which guarantees a pre-assigned probability of achieving the experimenter's goal.

It should be noted (by considering the complimentary subset which is not selected) that the above problem in relation to Goal I is logically equivalent to the same problem in relation to the goal of selecting a subset of size (k-s) which includes at least (k-t)-(s-c) of the (k-t) populations whose parameter values are the smallest. In other words the solutions

to the above problem in relation to Goal I, for all admissible values of c, s and t (with fixed k) will provide solutions to the same problem in relation to Goal II - to select a subset of size s which contains at least c of those t populations whose parameter values are the smallest.

It should also be noted that Goal I reduces to the goal of selecting the t best populations (in an unordered manner) when c=s=t. We formulate the problem along the lines of Bechhofer (1954) and Bechhofer and Sobel (1954).

1.2 Formulation of the problem.

Let $\vec{\theta}$ denote the vector of ordered θ-values viz., $(\theta_{[1]}, \theta_{[2]}, \ldots, \theta_{[k]})$ and let Ω stand for the parameter space, which is the set of all admissible vectors $\vec{\theta}$. Further let $d(x,y)$ be a continuous non-negative real-valued function defined for $x \geq y$ where x and y are both real, such that $d(x,y)=0$ if and only if $x=y$. Further for fixed y, it is a strictly increasing function of x and for fixed x, it is a strictly decreasing function of y. To avoid trivialities, we also assume that $d(x,y)$ can take on indefinitely large values. We shall call such a function a distance measure. Let d^* be a specified positive number. The parameter space Ω is partitioned into a "preference zone" $\Omega(d^*)$ defined by

$$\Omega(d^*) = \{\vec{\theta} : d(\theta_{[k-t+1]}, \theta_{[k-t]}) \geq d^*\} \tag{1.2.1}$$

and its complement $\bar{\Omega}(d^*)$, the "indifference zone." The choice of the distance measure depends on the class of distribution functions $\mathcal{F} = \{F(\cdot|\theta), \theta \in \Theta\}$ under consideration in a specific example.

In addition to specifying d^*, the experimenter also specifies another positive number P^* where $P^* < 1$. Without loss of generality we can assume that $P^* > P(c,k,s,t)$ (see the remark below), where

$$P(c,k,s,t) = \binom{k}{s}^{-1} \sum_{i=c}^{\min(s,t)} \binom{t}{i}\binom{k-t}{s-i} . \tag{1.2.2}$$

For all points $\vec{\theta}$ in the preference zone the definition of a correct selection (CS) is the obvious one, namely that the selected set includes any subset, consisting of at least c of the t best populations (i.e., those populations with the parameters $\theta_{[k-t+1]}, \theta_{[k-t+2]}, \ldots, \theta_{[k]}$). Any natural generalization of this definition of a CS in the indifference zone will suffice but, since we are concerned with a CS only in the preference zone, we shall not specify any particular definition of a CS in the indifference zone.

After specifying d* and P*, the experimenter desires to have a fixed sample size procedure for which the probability of a CS satisfies the condition

$$P(CS|\vec{\theta}) \geq P^* \quad \text{for all} \quad \vec{\theta} \in \Omega(d^*) \quad . \tag{1.2.3}$$

Remark 1.2

If P* is smaller than the bound (1.2.2), we can satisfy the requirement (1.2.3) by a random selection of the subset without taking any observations, since the bound P(c,k,s,t) is the probability of a CS under a random selection of the subset. Thus to make the problem non-trivial we have set the bound on P*. Clearly we need $c \geq 1$ to have a non-trivial problem. It should also be noted that if $c = s+t-k$ then $P(c,k,s,t) = 1$ by (1.2.2) and hence for any P* the requirement is again satisfied by a random selection of the subset without taking any observations. Hence to make the problem non-trivial we consider only those values of c, s and t for which $c \geq s+t-k+1$; that is, only those values for which $P(c,k,s,t) < 1$.

1.3 Proposed procedure R_s.

Given independent random variables $\{X_{ij}\}$, $j = 1,2,\ldots,n$; $i = 1,2,\ldots,k$ from the k populations $\prod_i$, let $T_i = T(X_{i1},\ldots,X_{in})$, $i = 1,2,\ldots,k$, be independent random variables having density functions. Here n is some fixed positive integer. Let the distribution function of T_i be denoted by

$G_n(\cdot|\theta_i)$. The choice of the function T will depend upon particular cases. $T_1, T_2, \ldots, T_k$ will be statistics relevant to the estimation of $\theta_1, \theta_2, \ldots, \theta_k$ respectively. The existence of statistics T_i with the desired properties is a basic assumption and this assumption will have to be checked in any particular case. The proposed procedure is based on these statistics T_i.

Procedure R_s:

Let $T_{[1]} \leq T_{[2]} \leq \cdots \leq T_{[k]}$ be the ordered T_i. The set of populations corresponding to $T_{[k-s+1]}, T_{[k-s+2]}, \ldots, T_{[k]}$ is the set to be selected.

Once the common sample size n is determined, the procedure R_s is completely defined; our problem will be that of determining this sample size so that the probability requirement (1.2.3) is satisfied. It should be noted that the required n-value not only depends on the class of distribution functions $G_n(\cdot|\theta)$, but also on the distance measure used in defining the preference zone of the parameter space.

As to the existence of the required n-value one can argue heuristically as follows: if T is a consistent estimator of θ, then the largest T-values will come from the populations with largest θ-values with a probability that tends to one as n tends to infinity. Hence the probability of a CS, under the procedure R_s, will tend to one as n tends to infinity. In some of the particular cases considered, it is shown explicitly that this is the case.

Remark 1.3

In practice we do encounter situations in which two or more T_i may be equal, even when T is a continuous random variable. In such cases the equal T-values should be ranked by using a randomized procedure which assigns equal probability to each possible ordering of those values.

A brief outline of the rest of this chapter can now be given: After defining a stochastically increasing family of distribution functions, we give some known properties of such a family. We prove a new result concerning such a family, which is used to prove a theorem on the monotone

properties of the probability of a CS under the procedure R_s. In proving this theorem we make the assumptions that T_i are absolutely continuous random variables and that the family $\mathcal{G} = \{G_n(\cdot|\theta) : \theta \in \Theta\}$ of distribution functions is stochastically increasing for all values of n. This theorem is then used to determine the required sample size.

1.4 Stochastically increasing family of distribution functions.

Here we give the definition of a stochastically increasing family of distribution functions and some examples of such families. Later we make some remarks concerning the choice of the statistics T_i. Let Θ be an interval of the real line.

Definition: A family of distribution functions $\mathcal{F} = \{F(\cdot|\theta) = F_\theta(\cdot) : \theta \in \Theta\}$ on the real line is said to be stochastically increasing (SI) if

(1.4.1) $\theta < \theta' \Rightarrow F_{\theta'}(x) \leq F_\theta(x)$ for all x, with strict inequality holding for some x.

The family is said to be strictly stochastically increasing if

(1.4.2) $\theta < \theta' \Rightarrow F_{\theta'}(x) < F_\theta(x)$ for all x.

If the distribution functions of the random variables X and X' are $F_\theta(\cdot)$ and $F_{\theta'}(\cdot)$, respectively, which satisfy (1.4.1) then

(1.4.3) $P(X > x) \leq P(X' > x)$ for all x.

In this case the variable X' is said to be stochastically larger than X.

A set of necessary and sufficient conditions for (1.4.1) to hold for two given distribution functions is given in lemma 1.5.1.

One of the simplest examples of an SI family is any location parameter family, that is, a family $F_\theta(x)$ such that $-\infty < \theta < \infty$ and

(1.4.4) $F_\theta(x) = G(x-\theta)$ for all x,

where G is some distribution function. Another example is any scale parameter

family having the interval (0,a) as the support, where a is allowed to depend on θ or it may be infinite; i.e., a family $F_\theta(x)$ such that $\theta > 0$, $F_\theta(0) = 0$ and

(1.4.5) $$F_\theta(x) = G\left(\frac{x}{\theta}\right) \text{ for all } x,$$

where G is some distribution function. A third example is any family of distribution functions whose densities possess the monotone likelihood ratio property.

Remarks about the choice of T.

Whenever a sufficient statistic for θ exists, which has fixed dimensionality for all n, then the proper choice of T is some appropriate function of the sufficient statistic. The choice of T becomes a problem only when such a sufficient statistic does not exist. We are mainly concerned with a property of the family of distribution functions $\mathcal{G}$. We would like to choose the function T such that the induced family $\mathcal{G}$ is stochastically increasing for each value of n. A sufficient condition for this is that T possess the monotone likelihood ratio property.

Special remark.

There are cases of interest where the distribution function characterizing $\prod_i$ involves a nuisance parameter. The results to be proved will also apply to such cases, provided the distribution of T_i depends only on θ_i (not on the nuisance parameter) in addition to the properties mentioned in the previous paragraph. For the purpose of simplicity (with slight loss of generality) we have assumed that the function F involves a single unknown parameter θ, but we also give some examples which involve nuisance parameters.

1.5 Some properties of a stochastically increasing family of distribution functions.

We shall need some properties of a stochastically increasing family of distribution functions. First we give some known results (lemmas 1.5.1 and

1.5.2) without proof and then we use them to prove a new result (lemma 1.5.3).

Lemma 1.5.1

Let F_0 and F_1 be two cumulative distribution functions on the real line such that $F_1(x) \leq F_0(x)$ for all x. Then there exists two non-decreasing functions g_0 and g_1, and a random variable V, such that (a) $g_0(v) \leq g_1(v)$ for all real v and (b) the distribution functions of the variables $g_0(V)$ and $g_1(V)$ are F_0 and F_1, respectively; the converse is also true.

The proof is given in Lehmann (1959, p. 73).

As a consequence of the above lemma we get the result: "If F_0 and F_1 are two distribution functions on the real line such that $F_1(x) \leq F_0(x)$ for all x, then $E_0\psi(X) \leq E_1\psi(X)$ for any non-decreasing function ψ."

The proof of this result is simple; in fact it is a problem in Lehmann (1959, p. 112). This result can be generalized in the following manner.

Lemma 1.5.2

Let $F(x|\theta) = F_\theta(x)$ where $\theta \in \Theta$, be an SI family of distribution functions on the real line. If ψ is any non-decreasing (non-increasing) function of x, then $E_\theta\psi(X)$ is a non-decreasing (non-increasing) function of θ.

Lemma 1.5.3†

Let $F(x|\theta) = F_\theta(x)$ where $\theta \in \Theta$, be an SI family of distribution functions on the real line. Let $X_1, X_2, \ldots, X_k$ be independent random variables with distribution functions $F(x_1|\theta_1), F(x_2|\theta_2), \ldots, F(x_k|\theta_k)$, respectively. For any fixed i (i = 1,2,...,k), if $\psi = \psi(x_1,x_2,\ldots,x_k)$ is a non-decreasing (non-increasing) function of x_i when all x_j for $j \neq i$ are held fixed, then $E\psi(X_1,X_2,\ldots,X_k)$ is a non-decreasing (non-

† After obtaining this lemma, I learned that Alam and Rizvi (1965) have independently derived a similar lemma.

increasing) function of θ_i.

Proof:

$$(1.5.1) \qquad E\psi(X_1,X_2,\ldots,X_k) = \int \psi \prod_{i=1}^{k} dF(x_i|\theta_i)$$

$$= \int \Big[\int \psi dF(x_i|\theta_i) \Big] \prod_{\substack{j=1 \\ j\neq i}}^{k} dF(x_j|\theta_j).$$

Since ψ is a non-decreasing (non-increasing) function of x_i when all x_j for $j \neq i$ are held fixed, from the lemma 1.5.2 it follows that

$$(1.5.2) \qquad E\{\psi(X_1,X_2,\ldots,X_k)|x_1,x_2,\ldots,x_{i-1},x_{i+1},\ldots,x_k\} = \int \psi dF(x_i|\theta_i)$$

is a non-decreasing (non-increasing) function of θ_i. Since this holds for each value $(x_1,x_2,\ldots,x_{i-1},x_{i+1},\ldots,x_k)$, the right hand member of (1.5.1) and hence $E\psi$ is a non-decreasing (non-increasing) function of θ_i. Since this holds for each fixed i, the lemma follows.

This lemma is used in proving a theorem, dealing with some monotone properties of the probability of a CS, in the next section.

1.6 Probability of a correct selection and its infimum.

In this section we determine the infimum of the probability of a correct selection for Goal I when the procedure R_s is used. Using this infimum we shall determine (in the next section) the required common sample size.

Let Y_i be the statistic based on the sample from the population with the parameter $\theta_{[i]}$, i = 1,2,...,k. That is, the set $(Y_1, Y_2, \ldots, Y_k)$ is same as the set $(T_{j_1}, T_{j_2}, \ldots, T_{j_k})$ where $(j_1, j_2, \ldots, j_k)$ is some permutation of (1,2,...,k). Our procedure R_s is based on the statistics T_j and hence it is based on the statistics Y_i. We make the following assumptions.

Assumption 1.6.1: The statistics Y_i (i = 1,2,...,k) are absolutely continuous

random variables.

Assumption 1.6.2: The family of distribution functions $\mathcal{G} = \{G_n(\cdot|\theta) : \theta \in \Theta\}$ is an SI family for each positive integer n.

First we shall prove the following

Lemma 1.6.1

$$\{CS\} \equiv \{c^{\underline{th}} \text{ largest of } (Y_{k-t+1}, Y_{k-t+2}, \ldots, Y_k) > (s-c+1)^{\underline{st}} \text{ largest of } (Y_1, Y_2, \ldots, Y_{k-t})\}.$$

Proof:

$$\{CS\} \equiv \{\text{among the s largest of } (Y_1, Y_2, \ldots, Y_k) \text{ there are at least c of } (Y_{k-t+1}, Y_{k-t+2}, \ldots, Y_k)\}$$

$$\equiv \bigcup_{j=c}^{\min(s,t)} \{\text{among the s largest of } (Y_1, Y_2, \ldots, Y_k) \text{ there are exactly j of } (Y_{k-t+1}, Y_{k-t+2}, \ldots, Y_k)\}$$

$$\equiv \{\text{at most (s-c) of } (Y_1, Y_2, \ldots, Y_{k-t}) \text{ are greater than the } c^{\underline{th}} \text{ largest of } (Y_{k-t+1}, Y_{k-t+2}, \ldots, Y_k)\}$$

$$\equiv \{\text{at least (k-t-s+c) of } (Y_1, Y_2, \ldots, Y_{k-t}) \text{ are less than the } c^{\underline{th}} \text{ largest of } (Y_{k-t+1}, Y_{k-t+2}, \ldots, Y_k)\}$$

[illegible]

$$\equiv \{c^{\underline{th}} \text{ largest of } (Y_{k-t+1}, Y_{k-t+2}, \ldots, Y_k) > (s-c+1)^{\underline{st}} \text{ largest of } (Y_1, Y_2, \ldots, Y_{k-t})\}.$$

This completes the proof of the lemma.

From the lemma, it follows that the probability of a correct selection at the parameter point $\vec{\theta}$ is given by

$$(1.6.1) \quad P(CS|\vec{\theta}) = P[c^{\underline{th}} \text{ largest of } (Y_{k-t+1}, Y_{k-t+1}, \ldots, Y_k) > (s-c+1)^{\underline{st}} \text{ largest of } (Y_1, Y_2, \ldots, Y_{k-t})],$$

where $Y_1,\ldots,Y_k$ is a set of independent random variables such that the distribution function of Y_i is $G_n(\cdot|\theta_{[i]})$, $i = 1,2,\ldots,k$.

We shall now prove a theorem giving some monotone properties of $P(CS|\vec{\theta})$, which is a function of the ordered θ-values.

Theorem 1.6

Under the assumptions 1.6.1 and 1.6.2, the $P(CS|\vec{\theta})$ is a non-increasing function of $\theta_{[\alpha]}$ ($\alpha = 1,2,\ldots,k-t$) and a non-decreasing function of $\theta_{[\beta]}$ ($\beta = k-t+1, k-t+2, \ldots, k$).

Proof:

By (1.6.1) the set of points in R^k where a CS occurs is the set $\{(y_1, y_2, \ldots, y_k) : u < v\}$ where u and v are, respectively, the $(s-c+1)^{st}$ largest of $(y_1, y_2, \ldots, y_{k-t})$ and the c^{th} largest of $(y_{k-t+1}, y_{k-t+2}, \ldots, y_k)$. If Ψ is the indicator function of this set, then

$$P(CS|\vec{\theta}) = E\Psi(Y_1, Y_2, \ldots, Y_k). \tag{1.6.2}$$

It is easy to see that u is a non-decreasing function of y_α ($\alpha = 1,2,\ldots, k-t$) when all y_i for $i \neq \alpha$ are held fixed and also that v is a non-decreasing function of y_β ($\beta = k-t+1, k-t+2, \ldots, k$) when all y_m for $m \neq \beta$ are held fixed. Hence Ψ is a non-increasing function of y_α ($\alpha = 1,2,\ldots,k-t$) when all other y's are held fixed and it is a non-decreasing function of y_β ($\beta = k-t+1, k-t+2, \ldots, k$) when all other y's are held fixed. By applying the lemma 1.5.3 to the function Ψ we obtain the desired result.

This theorem represents a valuable tool in obtaining the infimum of $P(CS|\vec{\theta})$. It forms one of the key results of this investigation.

Remark 1.6

When the assumption 1.6.1 is not satisfied, we transform the statistics of discrete type into statistics of continuous type. Section 2.8 deals with such a transformation.

From the theorem it follows that for any subset ω of the parameter space Ω

which has the structure of an ordered subset of a cartesian product of k identical sets Θ.

$$\inf_{\vec{\theta} \in \omega} P(CS|\vec{\theta}) = \inf_{\vec{\theta} \in \omega(\theta,\theta_0)} P(CS|\vec{\theta}) \tag{1.6.3}$$

where $\omega(\theta,\theta_0)$ is that set of points $\vec{\theta} \in \omega$, for which

$$\theta_{[1]} = \theta_{[2]} = \cdots = \theta_{[k-t]} = \theta_0 \text{ (say)}, \quad \theta_{[k-t+1]} = \theta_{[k-t+2]} = \cdots = \theta_{[k]} = \theta \text{ (say)}. \tag{1.6.4}$$

Here θ and θ_0 are arbitrary values such that $\theta \geq \theta_0$ and both belong to Θ. A configuration of the parameters $\theta_1, \theta_2, \ldots, \theta_k$ for which (1.6.4) holds is, sometimes, called a generalized least favorable (GLF) configuration. The $P(CS|\vec{\theta})$ for the GLF configuration (1.6.4) is given by

$$P(\theta,\theta_0) = \int_{-\infty}^{\infty} U(x|\theta_0)\,dV(x|\theta) = \int_{-\infty}^{\infty} [1-V(x|\theta)]\, d\,U(x|\theta_0) \quad , \tag{1.6.5}$$

where $U(\cdot|\theta_0)$ is the c.d.f. of the $(s-c+1)^{\underline{st}}$ largest of $(k-t)$ independent random variables, each having the c.d.f. $G_n(\cdot|\theta_0)$ and $V(\cdot|\theta)$ is the c.d.f. of the $c^{\underline{th}}$ largest of t independent random variables, each having the c.d.f. $G_n(\cdot|\theta)$. That is

$$U(x|\theta_0) = \sum_{\alpha=0}^{s-c} \binom{k-t}{\alpha} G_n^{k-t-\alpha}(x|\theta_0)[1-G_n(x|\theta_0)]^{\alpha} = I[G_n(x|\theta_0);\ c', s-c+1] \tag{1.6.6}$$

and

$$V(x|\theta) = \sum_{\alpha=0}^{t-c} \binom{t}{\alpha} G_n^{\alpha}(x|\theta)[1-G_n(x|\theta)]^{t-\alpha} = I[G_n(x|\theta);\ t-c+1, c], \tag{1.6.7}$$

where

$$c' = k-t-s+c \text{ and } I(x;p,q) = I_x(p,q) = [\beta(p,q)]^{-1} \int_0^x t^{p-1}(1-t)^{q-1}dt. \tag{1.6.8}$$

Since $G_n(x|\theta_0)$ is a non-increasing function of θ_0 for each x, from (1.6.6) it follows that $U(x|\theta_0)$ is a non-increasing function of θ_0 for

each x. Thus $P(\theta,\theta_0)$ is a non-increasing function of θ_0 for fixed θ.

Infimum of $P(CS|\vec{\theta})$ over the entire parameter space Ω

From (1.6.3) and (1.6.5) we have

$$\inf_{\vec{\theta}\in\Omega} P(CS|\vec{\theta}) = \inf_{\{(\theta,\theta_0) \,:\, \theta,\theta_0\in\Theta,\ \theta\geq\theta_0\}} P(\theta,\theta_0). \tag{1.6.9}$$

Since $P(\theta,\theta_0)$ is a non-increasing function of θ_0 for fixed θ we have

$$\begin{aligned}\inf_{\vec{\theta}\in\Omega} P(CS|\vec{\theta}) &= \inf_{\theta\in\Theta} P(\theta,\theta) \\ &= \inf_{\theta\in\Theta} \int_{-\infty}^{\infty} I[G_n(x|\theta);\, c', s-c+1]\, dI[G_n(x|\theta);\, t-c+1, c] \\ &= \int_0^1 I_y(c', s-c+1)\, dI_y(t-c+1, c) = J(c,k,s,t) \text{ (say).}\end{aligned} \tag{1.6.10}$$

Lemma 1.6.2

$J(c,k,s,t) = P(c,k,s,t)$, where $P(c,k,s,t)$ is defined by (1.2.2).

Proof:

By expressing $I(y;c',s-c+1)$ as a finite series, we have

$$J(c,k,s,t) = \sum_{j=0}^{s-c} \frac{t!}{(t-c)!\,(c-1)!} \binom{k-t}{s-c-j} \int_0^1 y^{k-s+j}(1-y)^{s-j-1}\,dy \tag{1.6.11}$$

$$= \sum_{i=c}^{s} \frac{t!}{(t-c)!\,(c-1)!} \cdot \frac{(k-t)!}{(s-i)!\,(k-t-s+i)!} \cdot \frac{(k-s-c+i)!\,(s+c-i-1)!}{k!}$$

$$= \binom{k}{t}^{-1} \sum_{i=c}^{s} \binom{s+c-i-1}{c-1} \binom{k-s-c+i}{t-c}$$

$$= \binom{k}{t}^{-1} \sum_{j=c}^{s} \binom{j-1}{c-1} \binom{k-j}{t-c}.$$

Let X denote the number of red balls in a random sample of size s chosen, without replacement, from an urn containing k balls of which t are red. Also let Y denote the number of balls needed to be drawn without replacement from the above urn so as to include exactly c red balls in the sample. Now it is easy to see that

$$(1.6.12)\qquad P(c,k,s,t) = P(X \geq c) = P(Y \leq s)$$

$$= \sum_{i=c}^{s} \frac{\binom{t}{c-1}\binom{k-t}{i-c}}{\binom{k}{i-1}} \cdot \frac{t-c+1}{k-i+1}$$

$$= \sum_{i=c}^{s} \binom{i-1}{c-1}\binom{k-i}{t-c} \Big/ \binom{k}{t}.$$

The lemma follows from (1.6.11) and (1.6.12).

Using the lemma, from (1.6.10) we obtain

$$(1.6.13)\qquad \inf_{\vec{\theta} \in \Omega} P(CS|\vec{\theta}) = P(c,k,s,t).$$

Infimum of $P(CS|\vec{\theta})$ over the preference zone $\Omega(d^*)$ [See (1.2.1)]

From (1.6.3) and (1.6.5), for any distance measure d we have

$$(1.6.14)\qquad \inf_{\vec{\theta} \in \Omega(d^*)} P(CS|\vec{\theta}) = \inf_{\{(\theta,\theta_0) \,:\, \theta,\theta_0 \in \Theta, d(\theta,\theta_0) \geq d^*\}} P(\theta,\theta_0).$$

From the monotone properties of the distance measure d (see section 1.2) and of the function $P(\theta,\theta_0)$, it follows that for fixed θ

$$(1.6.15)\qquad \inf_{\theta_0, d(\theta,\theta_0) \geq d^*} P(\theta,\theta_0) = P(\theta,\theta') = Q(\theta,n) \text{ (say)},$$

where θ' is that function of θ determined by $d(\theta,\theta') = d^*$. Hence

$$(1.6.16)\qquad \inf_{\vec{\theta} \in \Omega(d^*)} P(CS|\vec{\theta}) = \inf_{\theta \in \Theta} Q(\theta,n).$$

Using (1.6.6) and (1.6.7) in the first expression for $P(\theta,\theta')$ (see 1.6.5) we obtain

(1.6.17) $Q(\theta,n)$

$$= \frac{t!}{(t-c)!\,(c-1)!} \sum_{\alpha=0}^{s-c} \binom{k-t}{\alpha} \int_{-\infty}^{\infty} G_n^{k-t-\alpha}(x|\theta')[1-G_n(x|\theta')]^{\alpha} G_n^{t-c}(x|\theta)[1-G_n(x|\theta)]^{c-1} dG_n(x|\theta)$$

$$= \int_{-\infty}^{\infty} I[G_n(x|\theta');\ c', s-c+1]\, dI[G_n(x|\theta);\ t-c+1, c].$$

Using the second expression for $\bar{P}(\theta,\theta')$ in (1.6.5) we obtain

(1.6.18) $Q(\theta,n)$

$$= \frac{(k-t)!}{(s-c)!\,(c'-1)!} \sum_{\alpha=0}^{t-c} \binom{t}{\alpha} \int_{-\infty}^{\infty} G_n^{\alpha}(x|\theta)[1-G_n(x|\theta)]^{t-\alpha} G_n^{c'-1}(x|\theta')[1-G_n(x|\theta')]^{s-c} dG_n(x|\theta')$$

$$= \int_{-\infty}^{\infty} \{1-I[G_n(x|\theta);\ t-c+1, c]\}\, dI[G_n(x|\theta');\ c', s-c+1].$$

The infimum of $Q(\theta,n)$ over admissible values of θ is not easy to obtain in general. In each particular case we need special analysis to obtain this infimum. But when θ happens to be either a location parameter or a scale parameter for the family of distribution functions $\mathcal{G}$, this infimum can be obtained "automatically," i.e., without any further analysis by adopting a suitable definition of the distance measure. In each of the other particular cases considered in chapter II we determine this infimum explicitly.

1.7 Determination of the required sample size.

The required sample size is the smallest value of n for which

$$\inf_{\vec{\theta}\in\Omega(d^*)} P(CS|\vec{\theta}) = \inf_{\theta\in\Theta} Q(\theta,n) \geq P^*, \tag{1.7.1}$$

where $Q(\theta,n)$ is given by (1.6.17) or (1.6.18). Let us denote the infimum of $Q(\theta,n)$ by $H(n;d^*)$. If H is a non-decreasing function of n, then the required sample size is the smallest integer not less than the solution of the equation

$$H(n;d^*) = P^*. \tag{1.7.2}$$

In such a case the required sample size is unique. Further if the limit of $H(n)$, as $n \to \infty$, is one then a solution for (1.7.2) exists for any specified $P^* < 1$.

Remarks on the need and definition of the preference zone

Now we can answer the question - why we restrict our attention to the preference zone in writing the probability requirement (1.2.3)?

If there were no such restriction, then the sample size necessary is the smallest integer value of n for which

$$\inf_{\vec{\theta}\in\Omega} P(CS|\vec{\theta}) \geq P^*. \tag{1.7.3}$$

We have shown that the infimum of the $P(CS|\vec{\theta})$ over Ω is $P(c,k,s,t)$, which is the lower bound for P^* regardless of the sample size. Thus without the restriction to the preference zone, we cannot achieve our goal however large our sample may be.

The choice of the preference zone is equivalent to the choice of the d-function. This choice is governed by the behavior of $P(\theta,\theta')$ as a function of $\theta = \theta_{[k-t+1]}$ and $\theta' = \theta_{[k-t]}$. The behavior of $P(\theta,\theta')$ depends on the form of $G_n(\cdot|\theta)$. It should be noted that in some problems, it is sufficient to define the preference zone through one restriction such as $d(\theta_{[k-t+1]}, \theta_{[k-t]}) \geq d^*$, whereas in other problems it may be desirable to

introduce more than one restriction. One such example is the problem where θ is the mean of a Poisson population (Sobel 1963).

The particular definition of the distance measure given in any specific case enables us to determine explicitly the infimum of $P(\theta,\theta')$. In some cases obtaining this infimum may not be a simple matter and may even have to be obtained by numerical methods. One such example is the problem where θ is the probability of success for a Bernoulli variable; this problem, for the case $c = s = t = 1$, is considered by Sobel and Huyett (1957).

We shall now see how the equation (1.7.1) simplifies in the cases when θ is either a location or a scale parameter for the family $\mathcal{G}$.

Case (i) θ is a location parameter for the family $\mathcal{G}$

In this case we have that for all x

(1.7.4) $$G_n(x|\theta) = G_n(x-\theta), \text{ where } G_n(x) = G_n(x|0).$$

Using (1.6.7) and (1.6.8) in (1.6.6) and transforming the variable of integration, we have

(1.7.5) $$P(\theta,\theta_0) = P_n(d) = \int_{-\infty}^{\infty} I[G_n(x+d);\ k-t-s+c, s-c+1]\, dI[G_n(x);\ t-c+1, c]$$

where $d = \theta-\theta_0$. Since $P_n(d)$ depends on θ,θ_0 only through d, we define the "natural" distance measure for such a problem as

(1.7.6) $$d(a,b) = a-b.$$

It is easy to see that

(1.7.7) $$\inf_{\vec{\theta} \in \Omega(d^*)} P(CS|\vec{\theta}) = \inf_{d \geq d^*} P_n(d) = P_n(d^*) = H_L(n;d^*) \text{ (say)}.$$

Hence the equation (1.7.1) reduces to

(1.7.8) $$H_L(n;d^*) \geq P^*,$$

where $H_L(\cdot;\cdot)$ can be expressed in any one of the following equivalent forms.

(1.7.9)

$$H_L(n;d^*) = \frac{t!}{(t-c)!\,(c-1)!}\sum_{\alpha=0}^{s-c}\binom{k-t}{\alpha}\int_{-\infty}^{\infty} G_n^{k-t-\alpha}(x+d^*)[1-G_n(x+d^*)]^{\alpha}G_n^{t-c}(x)[1-G_n(x)]^{c-1}dG_n(x),$$

$$= \int_{-\infty}^{\infty} I[G_n(x+d^*);\, c', s-c+1]\,dI[G_n(x);\, t-c+1, c]$$

$$= \frac{(k-t)!}{(s-c)!\,(c'-1)!}\sum_{\alpha=0}^{t-c}\binom{t}{\alpha}\int_{-\infty}^{\infty} G_n^{\alpha}(x-d^*)[1-G_n(x-d^*)]^{t-\alpha}G_n^{c'-1}(x)[1-G_n(x)]^{s-c}dG_n(x)$$

$$= \int_{-\infty}^{\infty}\{1-I[G_n(x-d^*);\, t-c+1, c]\}\,dI[G_n(x);\, c', s-c+1].$$

Case (ii) θ is a scale parameter for the family $\mathcal{G}$

In this case we have that for all x

(1.7.10) $$G_n(x|\theta) = G_n\left(\frac{x}{\theta}\right), \text{ where } G_n(x) = G_n(x|1) \text{ and } G_n(0) = 0.$$

By transforming the variable of integration we obtain

(1.7.11) $$P(\theta,\theta_0) = P_n(d_1) = \int_0^{\infty} I[G_n(xd_1);\, k-t-s+c, s-c+1]\,dI[G_n(x);\, t-c+1, c]$$

where $d_1 = \theta/\theta_0$. Here we define the distance measure as

(1.7.12) $$d(a,b) = a/b.$$

Now it is easy to see

(1.7.13) $$\inf_{\vec{\theta}\in\Omega(d^*)} P(CS|\vec{\theta}) = \inf_{d_1\geq d^*} P_n(d_1) = P_n(d^*) = H_S(n;d^*) \text{ (say)}.$$

Hence the equation (1.7.1)' reduces to

(1.7.14) $$H_S(n;d^*) \geq P^*.$$

We can obtain the various (equivalent) expressions for H_S from those of H_L by changing $x+d^*$ to xd^*, $x-d^*$ to x/d^* and changing the lower limit of integration from $-\infty$ to 0. That is

$$(1.7.15)\quad H_S(n;d^*)= \frac{t!}{(t-c)!\,(c-1)!}\sum_{\alpha=0}^{s-c}\binom{k-t}{\alpha}\int_0^{\infty} G_n^{k-t-\alpha}(xd^*)[1-G(xd^*)]^{\alpha}G_n^{t-c}(x)[1-G_n(x)]^{c-1}dG_n(x)$$

$$= \int_0^{\infty} I[G_n(xd^*);\ c', s-c+1]\,dI[G_n(x);\ t-c+1, c]$$

$$= \frac{(k-t)!}{(s-c)!\,(c'-1)!}\sum_{\alpha=0}^{t-c}\binom{t}{\alpha}\int_0^{\infty} G_n^{\alpha}\left(\frac{x}{d^*}\right)\left[1-G_n\left(\frac{x}{d^*}\right)\right]^{t-\alpha} G_n^{c'-1}(x)[1-G_n(x)]^{s-c}dG_n(x)$$

$$= \int_0^{\infty}\{1-I[G_n\left(\frac{x}{d^*}\right);\ t-c+1, c]\}\,dI[G_n(x);\ c', s-c+1].$$

In the sequel we shall write $H_L(n)$, $H_S(n)$ for $H_L(n;d^*)$, $H_S(n;d^*)$.

1.8 <u>An approximation to the solution of the equation (1.7.2).</u>

We now make the following assumptions:

(i) $\inf_{\theta \in \Theta} Q(\theta,n) = Q(\theta_1,n)$,

(ii) The inverse function G_n^{-1} exists. That is, given any y such that $0 < y < 1$, there exists a unique x such that $G_n(x|\theta) = y$. We shall denote such an x-value by $G_n^{-1}(y|\theta)$.

Now from (1.6.17)

$$(1.8.1)\quad H(n) = \inf_{\theta\in\Theta} Q(\theta,n) = Q(\theta_1,n)$$

$$= \int_{-\infty}^{\infty} I[G_n(x|\theta_1');\ c', s-c+1]\,dI[G_n(x|\theta_1);\ t-c+1, c] = \int_0^1 b(y)\,dI[y;\ t-c+1, c],$$

where $b(y)$ and θ_1' are defined by

$$(1.8.2)\quad b(y) = I[a(y,n);\ c', s-c+1] \text{ and } d(\theta_1,\theta_1') = d^*.$$

The function $a(y,n)$ is given by the relation

$$(1.8.3)\quad a(y,n) = G_{n_1}[G_{n_1}^{-1}(y|\theta_1')|\theta_1'].$$

We can regard H(n) as the expectation of the function b(Y) of the random variable Y which has the beta distribution with parameters t-c+1 and c.

Replacing Y by EY in b(Y), we obtain the approximation

$$(1.8.4) \qquad H(n;d^*) \approx b(EY) = b\left(\frac{t-c+1}{t+1}\right) .$$

Thus an approximation to the solution of (1.7.2) is the solution of

$$(1.8.5) \qquad a\left(\frac{t-c+1}{t+1},n\right) = a^*,$$

where a^* is determined by the relation

$$(1.8.6) \qquad I_{a^*}(k-t-s+c,s-c+1) = P^*.$$

1.9 Particular cases of goal I which are of special interest.

Two particular cases of goal I, corresponding to c = t when s ≧ t and c = s when s ≦ t, are of special interest. These are the following goals.

Goal 1: Selection of a subset of size s which includes the t best populations, where s ≧ t.

Goal 2: Selection of a subset of size s which includes any s of the t best populations, where s ≦ t.

It should be noted that these two goals coincide when s = t. Then the common goal is the selection of the t best populations (without ordering). The solution to the (basic) problem in relation to goal 1 is used to obtain a solution to another problem, which is considered in part II of this investigation. The solutions to the problems, when the above goals are of interest, have been mentioned earlier by the author in an abstract (1965).

We shall now give the final results for these particular cases since we will be using them later.

Goal 1:

Here the lower bound for P^* is $\binom{k-t}{s-t}/\binom{k}{s}$. Selection of a subset which includes the t best populations (those with parameter values $\theta_{[k-t+1]}, \theta_{[k-t+2]}, \ldots, \theta_{[k]}$) is a correct selection. Now the sample size needed to achieve this goal, when the procedure R_s is used, is the smallest value of n for which

(1.9.1) $$\inf_{\theta\in\Theta} Q_1(\theta,n) \geq P^*$$

where

(1.9.2) $$Q_1(\theta,n) = \binom{k-t}{k-s} \int_{-\infty}^{\infty} [1-G_n(x|\theta)]^t [1-G_n(x|\theta')]^{s-t} d[G_n^{k-s}(x|\theta')]$$

$$= \int_{-\infty}^{\infty} \{1-I[G_n(x|\theta);1,t]\}\, dI[G_n(x|\theta');k-s,s-t+1].$$

Here θ', as a function of θ, is determined by $d(\theta,\theta') = d^*$.

Goal 2:

In this case the lower bound to P* is $\binom{t}{s}/\binom{k}{s}$. Selecting any subset of size s of the t best populations constitutes a correct selection. The sample size necessary is the smallest value of n for which

(1.9.3) $$\inf_{\theta\in\Theta} Q_2(\theta,n) \geq P^*$$

where

(1.9.4) $$Q_2(\theta,n) = \frac{t!}{s!(t-s-1)!} \int_{-\infty}^{\infty} G_n^{k-t}(x|\theta') G_n^{t-s}(x|\theta)[1-G_n(x|\theta)]^{s-1} dG_n(x|\theta),$$

$$= \int_{-\infty}^{\infty} I[G_n(x|\theta');k-t,1]\, dI[G_n(x|\theta);t-s+1,s].$$

Here also θ' is determined by the relation $d(\theta,\theta') = d^*$.

It is easy to see that Goal I is less "stringent" than both Goal 1 and Goal 2. So one expects that, for fixed c,k,t,P* and d*, the sample size necessary to achieve Goal I will be smaller than the sample size necessary to achieve Goal 1 (if $s \geq t$) or Goal 2 (if $s \leq t$). We now prove a general result from which this result follows immediately. Let n(c,s) denote the sample size necessary to achieve Goal I.

Theorem 1.9.1

For fixed k,t,s,P*,d* and for any distance measure

(1.9.5) $$n(c+1,s) \geq n(c,s),$$

provided $c+1 \leq \min(s,t)$, i.e., provided Goal I is meaningful with c replaced by c + 1.

Proof

In order to prove the theorem, it is sufficient to prove that, for fixed (but arbitrary) values of n and θ, the Q function satisfies the inequality

(1.9.6) $\quad Q(c,s) \geq Q(c+1,s).$

Here $Q(c,s)$ is the function $Q(\theta,n)$ which is the $P(CS|\vec{\theta})$ at the GLF configuration.

(1.9.7) $\quad \theta_{[1]} = \theta_{[2]} = \ldots = \theta_{[k-t]} = \theta'; \theta_{[k-t+1]} = \theta_{[k-t+2]} = \ldots = \theta_{[k]} = \theta,$

where $d(\theta,\theta') = d^*$.

Let $Y_1, Y_2, \ldots, Y_{k-t}$ be independent random variables each with the c.d.f. $G_n(.|\theta')$ and $Y_{k-t+1}, Y_{k-t+2}, \ldots, Y_k$ be independent random variables each with the c.d.f. $G_n(.|\theta)$. Further let the two sets of variables be independent of each other. Then

(1.9.8) $\quad Q(c,s) = Q(\theta,n)$

$$= P[c^{\underline{th}} \text{ largest of } (Y_{k-t+1},\ldots,Y_k) > (s-c+1)^{\underline{st}} \text{ largest of } (Y_1,\ldots,Y_{k-t})].$$

It is easy to see that

(1.9.9) $\quad [(c+1)^{\underline{st}} \text{ largest of } (Y_{k-t+1},\ldots,Y_k) > (s-c)^{\underline{th}} \text{ largest of } (Y_1,\ldots,Y_{k-t})]$

$$\Rightarrow [c^{\underline{th}} \text{ largest of } (Y_{k-t+1},\ldots,Y_k) > (s-c+1)^{\underline{st}} \text{ largest of } (Y_1,\ldots,Y_{k-t})].$$

Hence, from (1.9.9) and (1.9.8), we obtain

$$Q(c+1,s) \leq Q(c,s).$$

This completes the proof of the theorem.

From (1.9.6), since $c \leq t \leq s$ for Goal 1 and $c \leq s \leq t$ for Goal 2 we obtain

(1.9.10) $\quad Q_1(\theta,n) = Q(t,s) \leq Q(c,s)$ when $s \geq t$,

and

(1.9.11) $Q_2(\theta,n) = Q(s,s) \leq Q(c,s)$ when $s \leq t$.

Thus from (1.9.10) and (1.9.11), we have the following .

Corollary 1.9.1

For fixed c,k,t,P*,d* and for any distance measure

(i) $n(c,s) \leq n_1(s)$

(ii) $n(c,s) \leq n_2(s)$

where $n_i(s)$ is the sample size necessary to achieve Goal i (i = 1,2)...

Some more results of this type will be proved in part II of this investigation.

1.10 A sufficient condition for the existence of the required sample size.

We know that the required common sample size is the smallest value of n for which

(1.10.1) $$\inf_{\theta \in \Theta} Q(\theta,n) \geq P^*,$$

where $Q(\theta,n)$ is $P(CS|\vec{\theta})$ at the GLF configuration given by (1.9.7). The solution of the above equation exists provided the left side of (1.10.1) tends to 1 as $n \to \infty$. We shall now find a sufficient condition for the same. We make the assumption that the infimum of $Q(\theta,n)$ is its value at θ_1. Then (1.10.1) will reduce to

(1.10.2) $$Q(\theta_1,n) \geq P^* .$$

Now

(1.10.3) $Q(\theta_1,n)$

$= P[c^{\underline{th}}$ largest of $(Y_{k-t+1},\ldots,Y_k) > (s-c+1)^{\underline{st}}$ largest of $(Y_1,\ldots,Y_{k-t})]$

$\geq P[\min(Y_{k-t+1},\ldots,Y_k) > \max(Y_1,\ldots,Y_{k-t})]$,

where $Y_1,\ldots,Y_{k-t}$ are independent random variables each with the c.d.f. $G_n(\cdot|\theta_1')$ and $Y_{k-t+1},\ldots,Y_k$ are independent random variables each with

the c.d.f. $G_n(\cdot|\theta_1)$. The two sets of variables are independent sets.

Here θ_1' is determined by the relation $d(\theta_1,\theta_1') = d^*$. From (1.10.3), we have

$$(1.10.4)\quad 1-Q(\theta_1,n) \leq 1 - P[\min(Y_{k-t+1},\ldots,Y_k) > \max(Y_1,\ldots,Y_{k-t})]$$

$$= 1 - P[\bigcap_{i,j} \{Y_i > Y_j\} \quad \begin{matrix} i = k-t+1,\ldots,k \\ j = 1,2,\ldots,k-t \end{matrix}]$$

$$= P[\bigcup_{i,j} \{Y_i < Y_j\}]$$

$$\leq \sum_{i,j} P[Y_i < Y_j] = t(k-t)\, P[Y_k < Y_1].$$

It is easy to see that

$$(1.10.5)\quad \lim_{n\to\infty} P[Y_k < Y_1] = 0 \Rightarrow \lim_{n\to\infty} 1-Q(\theta_1,n) = 0 \Rightarrow \lim_{n\to\infty} Q(\theta_1,n) = 1.$$

Hence a sufficient condition for the existence of the required sample size is

$$(1.10.6)\quad \lim_{n\to\infty} P[Y_k < Y_1] = 0.$$

It may be interesting to find a sufficient condition for (1.10.6) to be true. One such condition is given below. Now

$$(1.10.7)\quad P[Y_k < Y_1] = P[Z < -a],$$

where $Z = \dfrac{(Y_k-Y_1) - E(Y_k-Y_1)}{\sqrt{Var(Y_k-Y_1)}}$ and $a = \dfrac{E(Y_k-Y_1)}{\sqrt{Var(Y_k-Y_1)}}$.

Since $\theta_1 > \theta_1'$, we have $G_n(\cdot|\theta_1) \leq G_n(\cdot|\theta_1') \Rightarrow E(Y_k) \geq E(Y_1)$. Thus a is non-negative. Now by Chebyschev's inequality we have from (1.10.7),

$$(1.10.8)\quad P[Y_k < Y_1] = P[Z < -a] \leq P[|Z| > a] \leq \frac{1}{a^2}.$$

Thus

$$(1.10.9)\quad \lim_{n\to\infty} \frac{1}{a^2} = 0 \Rightarrow \lim_{n\to\infty} P[Y_k < Y_1] = 0.$$

i.e. A sufficient condition for (1.10.6) to be true is

(1.10.10) $$\lim_{n\to\infty} \frac{\mu_2(\theta_1) + \mu_2(\theta_1')}{[\mu_1'(\theta_1) - \mu_1'(\theta_1')]^2} = 0$$

where $\mu_1'(\theta)$ and $\mu_2(\theta)$ are the mean and variance of the distribution with c.d.f. $G_n(.|\theta)$.

Chapter II

Applications to Specific Distributions

2.1 Summary.

In this chapter we consider the problem in relation to various specific families of distributions such as Normal, Gamma, Rectangular, Cauchy, and Poisson. That is, assuming that the distributions (which characterize the populations) belong to a specific family we obtain the equation, whose solution gives the required sample size. For the case of normal distributions two tables giving the values of $\lambda(= d^*\sqrt{n}/\sigma)$ have been prepared and are given at the end. From these values one can obtain the required sample size. For the cases in which the statistics (on which the procedure is based) are asymptotically normal, we give an approximation to the infimum of PCS. This approximation has been used to obtain an approximation to the required sample size. In the first seven sections we consider continuous distributions. In the last section we consider the problem in relation to Poisson populations.

2.2 Normal populations with unknown means and common known variance.

Here we assume that the populations under consideration are normal so that

$$(2.2.1) \qquad F(x|\theta) = \Phi[(x-\theta)/\sigma] = \int_{-\infty}^{(x-\theta)/\sigma} \phi(y)dy$$

where $\phi(\cdot)$ and $\Phi(\cdot)$ are respectively, the density and the distribution functions of the standard normal distribution. We further assume that the variances are all equal and the common value σ^2 is known. We base our procedure on sample means, i.e.,

$$(2.2.2) \qquad T_i = \frac{1}{n}\sum_{j=1}^{n} X_{ij} \qquad (i = 1,2,\ldots,k).$$

Now, for each i $(i = 1,2,\ldots,k)$ we have

(2.2.3) $G_n(x|\theta_i) = P[T_i \leq x] = \Phi[(x-\theta_i)\sqrt{n}/\sigma]$.

It is easy to see that θ is a pure location parameter for $G_n(\cdot|\theta)$ and we use the distance measure defined by

(2.2.4) $d(x,y) = x - y$.

In this case $H_L(n)$, as defined by (1.7.9), reduces to

$$(2.2.5)\quad H(\lambda) = H_L(n) = \int_{-\infty}^{\infty} I[\Phi(x+\lambda);\ c', s-c+1]\, dI[\Phi(x);\ t-c+1, c]$$

$$= \int_0^1 I[a(u,\lambda);\ c', s-c+1]\, dI[u;\ t-c+1, c],$$

where

(2.2.6) $a(u,\lambda) = \Phi[\Phi^{-1}(u)+\lambda]$, $\lambda = (d^*\sqrt{n})/\sigma$ and $c' = k-t-s+c$.

Here $\Phi^{-1}(\cdot)$ is the inverse function corresponding to the function $\Phi(\cdot)$. As n increases λ increases, so that $H(\lambda)$ increases with n. Further we have

$$(2.2.7)\quad \lim_{n\to\infty} H(\lambda) = \int_0^1 I_1(c', s-c+1)\, dI_u(t-c+1, c) = 1.$$

Thus the required sample size is the smallest integer greater than or equal to

(2.2.8) $n_0 = (\lambda\sigma/d^*)^2$,

where λ is the root of the equation

(2.2.9) $H(\lambda) = P^*$.

Since $H(\lambda) \to 1$ as $n \to \infty$, the solution of (2.2.9) exists for any $P^* < 1$ and it is unique since H is an increasing function λ.

<u>An approximation to the solution of the equation (2.2.9)</u>.

An approximation to the solution of the equation (2.2.9) can be

obtained from the results of section 1.8. In this case the approximation is

$$(2.2.10) \qquad \lambda_1 = \Phi^{-1}(a^*) - \Phi^{-1}\left(\frac{t-c+1}{t+1}\right) ,$$

where a^* is given by the equation

$$(2.2.11) \qquad I_{a^*}(c', s-c+1) = P^* .$$

Now we prove a result, which will be used to show that λ_1 is smaller than the solution of the equation (2.2.9) in some cases.

Lemma 2.2.1

For each value of λ and for $c' = 1$,

$$(2.2.12) \qquad H(\lambda) < I\left[a\left(\frac{t-c+1}{t+1}, \lambda\right); c', s-c+1\right].$$

Proof:

The required result can be viewed as

$$(2.2.13) \qquad EI[a(U,\lambda); \ 1, s-c+1] < I[a(EU,\lambda); \ 1, s-c+1],$$

where U is a beta random variable with the parameters t-c+1 and c. To prove (2.2.13), it is sufficient to show that, for fixed but arbitrary value of λ, $b(u) = I[a(u,\lambda); \ 1, s-c+1]$ is a strictly concave function of u on (0,1). To show that b(u) is a concave function on (0,1), it is sufficient to show that $\frac{d^2b(u)}{du^2}$ is negative on (0,1). Let us denote $a(u,\lambda)$ by $a(u)$. On differentiating b(u) we obtain

$$(2.2.14) \qquad \frac{db(u)}{du} = \frac{(k-t)!}{(s-c)!}[1-a(u)]^{s-c} \frac{da(u)}{du} ,$$

and

$$(2.2.15) \qquad \frac{d^2b(u)}{du^2} = \frac{(k-t)!}{(s-c)!}\left[\{1-a(u)\}^{s-c} \frac{d^2a(u)}{du^2} - (s-c)\{1-a(u)\}^{s-c-1}\left(\frac{da(u)}{du}\right)^2\right] .$$

Further we have

$$(2.2.16)\quad \frac{da(u)}{du} = \phi\{\Phi^{-1}(u)+\lambda\}/\phi\ \{\Phi^{-1}(u)\},$$

and

$$(2.2.17)\quad \frac{d^2a(u)}{du^2} = -[\lambda\,\phi\{\Phi^{-1}(u)+\lambda\}]/[\,\phi\{\Phi^{-1}(u)\}]^2 < 0, \text{ for } 0 \leq u < 1.$$

Also for $0 < u < 1$ we have $0 < a(u) < 1$ and this implies that

$$(2.2.18)\quad (s-c)\{1-a(u)\}^{s-c-1}\left(\frac{da(u)}{du}\right)^2 \geq 0, \text{ since } s \geq c.$$

Using (2.2.17) and (2.2.18) in (2.2.15), we obtain

$$(2.2.19)\quad \frac{d^2b(u)}{du^2} < 0\ , \quad \text{for } 0 < u < 1.$$

This completes the proof of the lemma.

From the lemma, whenever $c' = 1$, we have

$$(2.2.20)\quad H(\lambda_1) \leq I\left[a\left(\frac{t-c+1}{t+1}\ ,\lambda_1\right)\ ;\ 1,s-c+1\right] = P^*.$$

Since $H(\lambda)$ is an increasing function of λ, from (2.2.20) it follows that the solution of (2.2.9) is larger than λ_1, when $c' = 1$.

Recently in his thesis, Milton (1965) gave a table of the values of $H(\lambda)$. In his notation the equation (2.2.9) becomes

$$(2.2.21)\quad P(c;s,k-t,t,\lambda) = P^*.$$

That table gives the values of $P(c;s,k-t,t,\lambda)$ to 6 decimals, for $1 \leq t \leq k-t \leq 7$ and $t = 1$, $k-t = 8(1)12$; $s = 1(1)[(k-t)/2]$, $c = 1(1)t$; and $\lambda = 0(\cdot2)1,1.5,2,3$. Using suitable interpolation one can get the solution of (2.2.21).

Sample size determination for goals 1 and 2

As these two goals are of special interest we give the equations that determine the sample sizes necessary to achieve these goals. It may be noted that there is a relationship between the sample sizes necessary to achieve these two goals.

Goal 1

The sample size necessary to achieve this goal is the smallest integer greater than or equal to

$$n_1 = (\lambda\sigma/d^*)^2, \tag{2.2.22}$$

where λ is the solution of the equation

$$H_1(\lambda) = P^*. \tag{2.2.23}$$

Here $H_1(\lambda)$ is the value of $H(\lambda)$ as given by (2.2.5), when $c = t$. That is

$$\begin{aligned} H_1(\lambda) &= \int_{-\infty}^{\infty} I[\Phi(x+\lambda);k-s,s-t+1]\, dI[\Phi(x);1,t] \\ &= \int_{-\infty}^{\infty} \{1-I[\Phi(x-\lambda);1,t]\} dI[\Phi(x);k-s,s-t+1] \\ &= \frac{(k-t)!}{(s-t)!(k-s-1)!} \int_{-\infty}^{\infty} \Phi^t(x+\lambda)[1-\Phi(x)]^{k-s-1}\Phi^{s-t}(x)\varphi(x)dx. \end{aligned} \tag{2.2.24}$$

In obtaining the last expression from the second one we used the relations $\varphi(x) = \varphi(-x)$, $\Phi(x) = 1 - \Phi(-x)$ and $I_x(p,q) = 1 - I_{1-x}(q,p)$. This problem for the special case $k = 3$, $t = 1$ and $s = 2$ was considered earlier by the author and the result was mentioned in an abstract (1964).

Goal 2

Here the sample size necessary to achieve this goal is the smallest integer greater than or equal to

$$n_2 = (\lambda\sigma/d^*)^2, \tag{2.2.25}$$

where λ is the solution of the equation

$$H_2(\lambda) = P^*. \tag{2.2.26}$$

Here $H_2(\lambda)$ is the value of $H(\lambda)$ (as given by 2.2.5) when $c = s$. That is

$$(2.2.27)\quad H_2(\lambda) = \int_{-\infty}^{\infty} I[\Phi(x+\lambda);k-t,1]\,dI[\Phi(x);t-s+1,s]$$
$$= \frac{t!}{(s-1)!(t-s)!} \int_{-\infty}^{\infty} \Phi^{k-t}(x+\lambda)[1-\Phi(x)]^{s-1}\Phi^{t-s}(x)\ \phi(x)dx.$$

This goal was suggested by Sobel (see the footnote on page 22 of Bechhofer 1954) but no details were given.

Relationship between the sample sizes necessary to achieve the goals 1 and 2.

Comparing the third expression for H_1 and the second expression for H_2, it is easy to see that equivalent expressions for H_2 can be obtained from the equivalent expressions for H_1 by changing s to k-s and t to k-t. Thus if $\lambda_1(s,t)$ and $\lambda_2(s,t)$ are, respectively, the solutions of the equations (2.2.23) and (2.2.26), then

$$(2.2.28)\quad \lambda_2(s,t) = \lambda_1(k-s,k-t).$$

If $n_{10}(s,t)$ and $n_{20}(s,t)$ are respectively, the sample sizes necessary to achieve the goals 1 and 2, then

$$(2.2.29)\quad n_i(s,t) \leq n_{i0}(s,t) < n_i(s,t) + 1\ ,$$

where

$$(2.2.30)\quad n_i(s,t) = \{\lambda_i(s,t)\sigma/d^*\}^2,\ i = 1,2.$$

Further from (2.2.28), (2.2.29) and (2.2.30), we obtain

$$(2.2.31)\quad n_1(k-s,k-t) \leq n_{20}(s,t) < n_1(k-s,k-t) + 1.$$

Thus for a given set of k and P* values, a table of λ_1-values alone, considering all admissible s and t combinations, will provide solutions to the equation (2.2.26).

Table 1

Table 1 gives the values of λ_1 for some values of k,s,t and P*. These

values are obtained in the following manner:

From the third expression for $H_1(\lambda)$, we obtain

(2.2.32) $H_1(\lambda_1)$

$$= \frac{(k-t)!}{(s-t)!(k-s-1)!} \int_{-\infty}^{\infty} \Phi^t(x+\lambda_1)[1-\Phi(x)]^{k-s-1}[1-\{1-\Phi(x)\}]^{s-t}\phi(x)dx$$

$$= \frac{(k-t)!}{(s-t)!(k-s-1)!} \sum_{j=0}^{s-t} \binom{s-t}{j}(-1)^j \int_{-\infty}^{\infty} \Phi^t(x+\lambda_1)[1-\Phi(x)]^{k-s-1+j}\phi(x)dx$$

$$= \frac{(k-t)!}{(s-t)!(k-s-1)!} \sum_{j=0}^{s-t} \binom{s-t}{j}(-1)^j \frac{1}{k-s+j} P(\lambda_1|k-s+j, k-s+t+j), \text{ (say)}.$$

The function $P(x|r,k)$ has been tabulated by Teichroew (1955) at x-values increasing by .01. Using these tables $H_1(\lambda_1)$ has been calculated over suitable range of λ_1-values. The λ_1-values corresponding to given P* values have been obtained by linear interpolation. This table gives the λ_1-values for t = 1(1)k-2, s = t+1(1)k-1, k = 3(1)5 and P* = .9995, .999, .995, .99(.01).95, .90, .80. This table incidentally gives λ_2-values for some combinations of k,t,s, in view of the relation (2.2.28).

Remark on accuracy of λ_1-values

For a particular combination of k, s, t and P* let λ_{11} be the λ_1-value from table 1, truncated after two decimals (without rounding the second decimal) and $\lambda_{12} = \lambda_{11} + .01$. Then the required λ_1-value satisfies the inequality

$$\lambda_{11} < \lambda_1 < \lambda_{12} ,$$

so that

$$n_{11} < n_1 < n_{12} ,$$

where

$$n_{1i} = (\lambda_{1i}\, \sigma/d^*)^2 \qquad (i = 1,2).$$

If we take n_{12} for n_1, an upper bound on the error is

$$(\lambda_{12}^2 - \lambda_{11}^2)(\sigma/d^*)^2 = .01(2\lambda_{11} + .01)(\sigma/d^*)^2.$$

Table 2

This table gives λ-values which provide solutions to sample size determination problem in relation to Goal I. From table D of Milton (1965), these values have been obtained by linear interpolation. The table gives the values for c = 1, t = s = 2, k = 4(1)6 and c = 1, t = 2, s = 3, k = 6 and P* = .9995, .999, .995, .99(.01).95, .90, .80.

An illustration

Suppose we have 4 populations. Let P* = 0.99. We want to select two populations which includes the best. From table 1, for k = 4, t = 1 s = 2 and P* = .99 we have

$$(n_1^{\frac{1}{2}}\, d^*)/\sigma = 2.809.$$

Now if $d^* = \sigma$, then $n_1 = 7.890$. That is, we need 8 observations from each population to achieve our goal when the procedure R_s is used.

Suppose we are interested in choosing any two of the three best. From the relation (2.2.8), it follows that we need 8 observations from each population.

If we are interested in choosing two which include at least one of the two best, from Table 2

$$n_0 = (1.52)^2 = 2.3104.$$

That is, we need 3 observations from each population to achieve this goal.

Some remarks concerning different variance models

So far in our discussion we have assumed that the variances of the k normal populations are equal and the common variance value is known. The natural question is how to deal with the cases where this assumption is not satisfied. Now we indicate some methods of dealing with such cases.

Case 1. Variances known and unequal

Let σ_i^2 be the variance of $\prod_i$, i = 1,2,...,k. Intuitively one might decide to choose the sample sizes n_i so that the variances of the sample means are equal (or approximately equal). Using this intuitive idea we can

proceed as follows: obtain the λ-value with assumption of common variance, in accordance with the goal of interest. Then n_i $(i = 1,2,\ldots,k)$ is chosen as the integer greater than or equal to $(\lambda\sigma_i/d^*)^2$. It can be shown such a choice of n_i will ensure that the $P(CS|\vec{\theta})$ is not less than P^*.

The question as to how to solve the problem with common n under different variances, is not treated here.

Case 2. Variances are equal and the common value is unknown

Here we redefine our distance measure as

$$d(a,b) = (a - b)/\sigma \quad , \tag{2.2.33}$$

where σ^2 is the common unknown variance. With this modification the enitre discussion can be carried over so that the sample size needed is the smallest integer greater than or equal to

$$n_0 = (\lambda/d^*)^2 \quad , \tag{2.2.34}$$

where λ is the solution of (2.2.9) or (2.2.23) or (2.2.26) according as the goal of interest is Goal I or Goal 1 or Goal 2.

If one insists on the distance measure (2.2.4) instead of (2.2.33), then this solution does not hold. Then it will be necessary to consider a two-stage or sequential procedure to provide a solution to the problem.

2.3 Examples in which the statistics T_i have gamma distribution.

This section deals with the examples in which the statistics used have gamma distribution with unknown scale parameter and known shape parameter.

Definition: A variable X is said to have the gamma distribution $\gamma(\alpha,\beta)$ with parameters α and β, if the probability density of X is given by

$$g(x|\alpha,\beta) = \begin{cases} [\beta^{\alpha}\Gamma(\alpha)]^{-1} x^{\alpha-1} e^{-x/\beta} & \text{for } x > 0, \\ 0 & \text{for } x \leq 0. \end{cases} \tag{2.3.1}$$

Here α and β are positive constants; β is the scale parameter and α is called the shape parameter. The distribution $\gamma(\alpha,\beta)$ is sometimes referred to as a Type III distribution. The class of distributions $\{\gamma(\alpha,\beta): \beta > 0\}$

for a fixed α, is an SI family. We shall denote the c.d.f. of the distribution $\gamma(\alpha,\beta)$, by $G(\cdot|\alpha,\beta)$.

(i) Normal populations with unknown variances

Let $\prod_i$ be characterized by the normal distribution with mean μ_i and variance θ_i $(i = 1,2,\ldots,k)$. Here θ's are unknown and μ's may be known or unknown. We assume that Goal II is of interest. We shall use statistics T_{i1} or T_{i2} where

$$(2.3.2) \qquad T_{i1} = \sum_{j=1}^{n} (X_{ij} - \mu_i)^2, \quad T_{i2} = \sum_{j=1}^{n} (X_{ij} - \bar{X}_i)^2,$$

according as the means are known or unknown. The statistic T_{i1} is sufficient for θ_i when μ_i is known, whereas T_{i2} is a function of the sufficient statistic $(\sum_{j=1}^{n} X_{ij}, \sum_{j=1}^{n} X_{ij}^2)$ when μ_i is unknown. Further T_{im} is distributed as $\theta_i \chi^2_{2\nu_m}$ where $\nu_1 = n/2$ and $\nu_2 = (n-1)/2$; i.e., the distribution of T_{im} is $\gamma(\nu_m, 2\theta_i)$. Thus here

$$(2.3.3) \qquad G_n(\cdot|\theta) = G(\cdot|\nu_m, 2\theta),$$

when the statistics T_{im} are used. When the means are unknown this example is the one in which nuisance parameters are present. Since θ is the scale parameter for the family $\mathcal{G}$, we use the distance measure defined by

$$(2.3.4) \qquad d(x,y) = x/y.$$

Here the selected subset is the set of populations which correspond to the s smallest T-values. As mentioned in section 1.1, we know that the sample size necessary to achieve Goal II for given values of c,k,s,t,P* and d* is same as the sample size necessary to achieve Goal I with the parameters $c' = (k-t)-(s-c)$, k, k-s, k-t, P* and d*. That is, the sample size necessary is ν_0 or ν_0+1 according as the means are known or unknown, where ν_0 is the smallest (integer) value of ν for which

(2.3.5) $$H_S^{II}(\nu) \geq P^*.$$

Here $H_S^{II}(\nu)$ is the value of $H_S(n)$ as given by (1.7.15) where c,s,t and $G_n(\cdot)$ are to be replaced, respectively, by c', k-s, k-t and $G_\nu(\cdot)$. In other words

(2.3.6) $$H_S^{II}(\nu) = \int_0^\infty I[G_\nu(xd^*);c,t-c+1]\,dI[G_\nu(x);s-c+1,c'],$$

where $G_\nu(\cdot) = G(\cdot|\nu_m,2)$ when the statistics T_{im} (m = 1,2) are used. If $H_S^{II}(\nu)$ is a non-decreasing function of ν, we can replace the inequality in (2.3.5) by an equality.

This problem for the case c = s = t was considered by Bechhofer and Sobel (1954).

(ii) Life testing with (negative) exponential distributions

Suppose we have a random sample of n_i items from $\prod_i$ which are put on a life test (i = 1,2,...,k). Let the life distribution of items from $\prod_i$ be the (negative)exponential distribution with the probability density

(2.3.7) $$f(x|\theta_i,A_i) = \begin{cases} \theta^{-1} \exp[-(x-A_i)/\theta_i] & \text{for } A_i < x < \infty \\ 0 & \text{for } A_i \geq x \end{cases}.$$

Here $\theta_i > 0$ and $A_i \geq 0$. Suppose we stop testing the items from $\prod_i$ after obtaining the first r(≥ 1) failures. On the basis of this information the experimenter is interested in achieving Goal I.

Let the r ordered failure times of the items from $\prod_i$ be $X_{i1} < X_{i2} < \ldots < X_{ir}$ (i = 1,2,...,k). We use the statistics T_{i1} or T_{i2} according as A_i's are known or unknown. A_i's are nuisance parameters when they are unknown. The statistics are defined as

(2.3.8) $$\begin{cases} T_{i1} = \sum_{j=1}^{r} (X_{ij} - A_i) + (n_i - r)(X_{ir} - A_i), \\ T_{i2} = \sum_{j=1}^{r} (X_{ij} - X_{i1}) + (n_i - r)(X_{ir} - X_{i1}). \end{cases}$$

Epstein and Sobel (1954) proved that the distribution of $T_{i\alpha}$ is $\gamma(\nu_\alpha, \theta_i)$, where $\nu_1 = r$ and $\nu_2 = r-1$. The problem here is the determination of r, so as to achieve Goal I.

(iii) Double exponential distributions (Laplace distributions)

Let us assume that the population $\prod_i$ is characterized by the double exponential distribution with the probability density

$$(2.3.9) \qquad f(x|\theta_i) = (2\theta_i)^{-1} \exp[-|x|/\theta_i].$$

We use the statistics $T_i = \sum_{j=1}^{n} |X_{ij}|$; T_i is sufficient for $\theta_i (i = 1,2,\ldots,k)$. Further T_i has the distribution $\gamma(n,\theta_i)$. Here also Goal I is of interest.

(iv) Gamma distributions with unknown scale parameters and common known α

In some experimental situations one is sampling from gamma populations with unknown scale parameter and common known α. The scale parameters are of interest. For example such a thing arises when one is observing life distributions of structures. It has been shown by Birnbaum and Saunders (1958), that the life length of certain structures under a particular load pattern follows a gamma distribution with known α and unknown scale parameter. Then for each i, X_{ij} has the distribution $\gamma(\alpha,\theta_i)(j = 1,2,\ldots,n)$. We use the statistics $T_i = \sum_{j=1}^{n} X_{ij}$, where T_i has the distribution $\gamma(n\alpha,\theta_i)$ $(i = 1,2,\ldots,k)$. Here again Goal I is of interest.

In each of the cases (ii) through (iv), the statistics T_i have gamma distributions and the parameters of interest are the scale parameters. We have to find ν_0, the smallest (positive integer) value of ν for which

$$(2.3.10) \qquad H_S(\nu) \geq P^*,$$

Here $H_S(\nu)$ is $H_S(n)$ as given by (1.7.15), where $G_n(\cdot)$ is to be replaced by $G(\cdot|\nu,1)$. Now the required r-value in case (ii) is ν_0 or ν_0+1 according as A's are known or unknown. In case (iii) the required sample size is ν_0 and in case (iv) the required sample size is $\left[\frac{\nu_0}{\alpha}\right] + 1$ ([x] denotes

the integral part of x).

Again it may be pointed out that when $H_S(\nu)$ is a non-decreasing function of ν, we can replace the inequality in (2.3.10) by an equality. We now show that the sufficient condition for the existence of the solution of (2.3.10), as found in section 1.10, is satisfied here. We give the proof of this in relation to cases (ii) through (iv). Since the proof is true for all values of c,s,t for given k,P* and d*, a similar result holds for the case (i).

Using the notation of section 1.10, we have to show that

$$(2.3.11) \qquad P[Y_k < Y_1] \to 0 \qquad \text{as } \nu \to \infty.$$

Since Y_1 and Y_k are independent random variables with the distributions $\gamma(\nu,1)$ and $\gamma(\nu,d^*)$ we have

$$E(Y_k - Y_1) = \nu(d^* - 1),$$

$$\text{Var}\,(Y_k - Y_1) = \text{Var}\, Y_k + \text{Var}\, Y_1 = \nu(d^{*2} + 1).$$

Now

$$\frac{\text{Var}(Y_k - Y_1)}{\{E(Y_k - Y_1)\}^2} = \frac{\nu(d^{*2}+1)}{\nu^2(d^*-1)^2} \to 0, \qquad \text{as } \nu \to \infty.$$

This implies that (2.3.11) is true, which implies that

$$\lim_{\nu \to \infty} H_S(\nu) = 1.$$

Thus the solution of (2.3.10) exists for any $P^* < 1$.

When c = s = t = 1, we have

$$H_S(\nu) = \int_0^\infty G_\nu^{k-1}(xd^*)\, dG_\nu(x),$$

where $G_\nu(\cdot)$ is the c.d.f. of $\gamma(\nu,1)$. Tables prepared by Gupta (1963) can be used to find ν values for given values of k,P* and d*.

Large sample approximation to the infimum of the PCS

Now we consider an approximation to the infimum of the PCS assuming

that the sample size n is sufficiently large (in case (ii) we assume that r is large). It is known that the family of distributions $\gamma(\alpha,\beta)$ satisfy the addition theorem for independent random variables, namely

$$\gamma(\alpha_1,\beta) \ * \ \gamma(\alpha_2,\beta) \ = \ \gamma(\alpha_1 + \alpha_2,\beta)$$

where * stands for convolution. Thus in each of the above four cases, T_i can be viewed as a sum of independent and identically distributed random variables each having gamma distribution. By lemma 5e.1 of Rao(1952), it follows that $Z_i = b(\nu)[\log_e T_i/\{a(\nu)\theta_i\}]$ is asymptotically distributed as a standard normal variable. Here $a(\nu)$ and $b(\nu)$ are suitable functions of ν.

Expressing the PCS in terms of the variables Z_i, and using their asymptotic distributions we obtain

$$\inf P_I \approx \int_{-\infty}^{\infty} I[\Phi(x+\lambda);c',s-c+1] \, dI[\Phi(x);t-c+1,c], \tag{2.3.12}$$

where

$$\lambda = b(\nu) \log_e d^*,$$

and P_I is the PCS for Goal I. Also we have

$$H_S^{II}(\nu) \approx \int_{-\infty}^{\infty} I[\Phi(x+\lambda);c,t-c+1] \, dI[\Phi(x);s-c+1,c']$$

$$= \int_{-\infty}^{\infty} I[\Phi(x+\lambda);c',s-c+1] \, dI[\Phi(x);t-c+1,c].$$

In obtaining the second expression from the first, we used the well known results $\Phi(x) = 1-\Phi(-x)$ and $I_x(p,q) = 1-I_{1-x}(q,p)$.

Thus an approximation to the solution of (2.3.5) and also to the solution of (2.3.10) is the smallest integer greater than

$$\nu_0' = b^{-1}[\lambda(c,k,s,t)/\log_e d^*],$$

where $b^{-1}(\cdot)$ is the inverse to $b(\cdot)$ and $\lambda(c,k,s,t)$ is given by

$$\int_{-\infty}^{\infty} I[\Phi(x+\lambda);c',s-c+1] \, dI[\Phi(x);t-c+1,c] = P^*.$$

These λ-values can be obtained from tables 1 and 2. Approximating ν_0 by

$[\nu_0'] + 1$, one can obtain an approximation to the required sample size (the r-value in case (ii)).

2.4 Uniform distributions.

In this section we consider the problem in relation to two types of uniform distributions viz., (a) those involving one parameter, (b) those involving two parameters. It may be noted that these distributions are non-regular. In this respect these special cases differ from those considered so far.

(a) One-parameter uniform distributions

Here we assume that the population $\prod_i$ is characterized by the uniform distribution over the interval $(0,\theta_i)$, $i = 1,2,\ldots,k$. We base our procedure on the statistics T_i where $T_i = \max_j X_{ij}$. It may be noted these statistics are sufficient for $\theta_1,\theta_2,\ldots,\theta_k$. The density function of T_i is

$$(2.4.1) \qquad g_n(y|\theta_i) = \begin{cases} ny^{n-1}/\theta_i^n & \text{for } 0 < y < \theta_i \\ 0 & \text{otherwise.} \end{cases}$$

If $G_n(\cdot|\theta_i)$ is the distribution function of T_i, then the class of distribution functions $\mathcal{G} = \{G_n(\cdot|\theta);\theta > 0\}$ is a scale parameter family such that $G_n(0|\theta) = 0$. Hence it is an SI family. Since θ is a scale parameter for the family $\mathcal{G}$, we define the distance measure as

$$(2.4.2) \qquad d(x,y) = x/y.$$

Here the infimum of the PCS is $H_S(n)$. The value of $H_S(n)$ can be obtained from (1.7.15) by taking $G_n(y)$ to be the distribution function corresponding to the density $g_n(y|1)$ (which is given by 2.4.1). Denoting y^n by u and $(d^*)^n$ by d_n^*, from (1.7.15) we have

$$(2.4.3) \qquad H_S(n) = \int_0^1 \{1-I[\left(\frac{u}{d_n^*}\right);t-c+1,c]\}dI[u;c',s-c+1].$$

Since $d^* > 1$, d_n^* increases with n. So $1-I\left[\left(\frac{u}{d_n^*}\right);t-c+1,c\right]$

increases with n for each u. Hence $H_S(n)$ is an increasing function of n. Thus the required sample size is the smallest integer greater than or equal to the solution of the equation

$$(2.4.4) \qquad \int_0^1 \{1-I[(\frac{u}{d_n^*});t-c+1,c]\}\, dI[u;c',s-c+1] = P^*.$$

Also since $d^* > 1$, $\lim_{n\to\infty} \frac{u}{d_n^*} = 0$, so that

$$(2.4.5) \qquad \lim_{n\to\infty} H_S(n) = \int_0^1 dI_u(c',s-c+1) = 1.$$

Thus the solution of (2.4.4) and hence the required sample size exists for any specified $P^* < 1$. It is unique since $H_S(n)$ is an increasing function of n.

This problem with $c = s = t$ has been considered by Barr and Rizvi(1964).

(b) Two-parameter uniform distributions

Now we assume that $\prod_i$ is characterized by the uniform distribution over $(\mu_i, \mu_i + \theta_i)$ $(i = 1,2,\ldots,k)$. We consider the following different cases.

Case (i) μ's are equal with the common value known or unknown and the populations are ranked according to θ-values

Whether the common μ-value is known or unknown, we can express the PCS in terms of $U_i = T_i - \mu$, where the statistics T_i are defined as in (a). Then the problem reduces to the one-parameter case considered above.

Case (ii) θ-values are equal and the populations are ranked according to μ-values

Whether the common θ-values is known or unknown we use the minimum of the sample as our statistic. The probability density function of T_i is

$$(2.4.6) \qquad g_n(y|\mu_i,\theta)' = \begin{cases} \frac{n}{\theta}\,[1-(\frac{y-\mu_i}{\theta})]^{n-1} & \text{for } \mu_i < y < \mu_i + \theta \\ 0 & \text{otherwise.} \end{cases}$$

If $G_n(\cdot|\mu_i,\theta)$ is the distribution function of T_i, then the family $\mathcal{G} = \{G_n(\cdot|\mu,\theta) : -\infty < \mu < +\infty\}$ is a location parameter family for fixed θ. Hence it is an SI family. We take the distance measure given by

(2.2.4), when θ is known. In the case of known θ-value, we make a scale transformation on the statistics T_i and use the statistics $Z_i = T_i/\theta$, so that the new location parameters are $\Phi_i = \mu_i/\theta$ and we can define the preference zone in relation to the new parameters. If the θ-value is unknown we express the PCS in terms of the random variables Z_i and assume

$$d(x,y) = \frac{x-y}{\theta} . \tag{2.4.7}$$

With this modification we can solve the problem. This type of modification in the definition of the distance measure has been suggested in case of normal populations with unknown common variance, which are ranked according to means.

Let δ stand for (d^*/θ) or d^* according as θ is known or unknown. Setting $G_n(\cdot) = G_n(\cdot|0,1)$ (the distribution function corresponding to $g_n(\cdot|0,1)$ of (2.4.6)) in the second expression of (1.7.9), we obtain for $\delta < 1$.

$$H_L(n) = \int_0^{1-\delta} I[1-(1-y-\delta)^n; c', s-c+1]\, dI[1-(1-y)^n; t-c+1, c] + \int_{1-\delta}^{1} dI[1-(1-y)^n; t-c+1, c], \tag{2.4.8}$$

and for $\delta \geq 1$

$$H_L(n) = \int_0^1 dI[1-(1-y)^n; t-c+1, c] = \int_0^1 dI[u; t-c+1, c] = 1 . \tag{2.4.9}$$

The required sample size is the smallest integer value of n for which

$$H_L(n) \geq P^* . \tag{2.4.10}$$

It is evident that we have to consider only those values of d^* for which $\delta < 1$, to have a non-trivial problem.

Now we shall show that $H_L(n)$ tends to one as $n \to \infty$. It is sufficient to show that (see section 1.10)

$$P(Y_k < Y_1) \to 0 \quad , \quad \text{as } n \to \infty. \tag{2.4.11}$$

We shall show that (1.10.10) is satisfied here. Now

$$\frac{\mu_2(\theta^0) + \mu_2(\theta_0')}{\mu_1'(\theta^0) - \mu_1(\theta_0')^2} = \frac{2n}{(d^*)^2(n+1)^2(n+2)} \to 0, \text{ as } n \to \infty.$$

This implies that (2.4.11) is true. Hence the required sample size exists for any specified $P^* < 1$.

Case (iii) μ-values are unknown and the populations are ranked according to θ-values

Here we use the statistics $T_i = \max_j X_{ij} - \min_j X_{ij}$. The distribution of T_i is independent of μ_i. The probability density of T_i is

$$(2.4.12) \quad g_n(y|\theta_i) = \begin{cases} \frac{n(n-1)}{\theta_i} \left(\frac{y}{\theta_i}\right)^{n-2} \left(1 - \frac{y}{\theta_i}\right) & \text{for } 0 < y < \theta_i \\ 0 & \text{otherwise.} \end{cases}$$

If $G_n(\cdot|\theta_i)$ is the c.d.f. of T_i, then the family $\mathcal{G} = \{G_n(\cdot|\theta); \theta > 0\}$ is a scale parameter family with $G_n(0|\theta) = 0$; hence it is an SI family. Here we take the ratio as the distance measure (see (2.3.4)).

Now the required sample size is the smallest integer value of n for which

$$(2.4.13) \quad H_S(n) \geq P^*.$$

Replacing $G_n(\cdot)$ by $I_y(n-1,2)$, from (1.7.15) we have

$$(2.4.14) \quad H_S(n) = \int_0^1 1 - I\left[I_{\frac{y}{d^*}}(n-1,2); t-c+1, c\right] \, dI[I_y(n-1,2); c', s-c+1].$$

We shall now show that

$$(2.4.15) \quad \lim_{n\to\infty} H_S(n) = 1.$$

By the sufficient condition derived in section 1.10, we have to show that

$$P(Y_k < Y_1) \to 0 \quad , \quad \text{as } n \to \infty.$$

Here Y_1 and (Y_k/d^*) are independent and each has the beta distribution with the parameters n - 1 and 2. We show that the sufficient condition for (2.4.15)

to hold, namely, (1.10.10) is true here. Here $\theta^0 = d^*$ and $\theta_o' = 1$,

$$\mu_1'(d^*) = d^*\mu_1'(1) \text{ and } \mu_2(d^*) = (d^*)^2\mu_2(1).$$

(2.4.16) Thus
$$\frac{\mu_2(d^*)+\mu_2(1)}{[\mu_1'(d^*)-\mu_1'(1)]^2} = \frac{1+d^{*2}}{(d^*-1)^2}\,\frac{\mu_2(1)}{[\mu_1'(1)]^2} .$$

Since $G_n(x|1) = I_x(n-1,2)$, we have

$$\mu_1'(1) = \frac{n-1}{n+1} \text{ and } \mu_2(1) = \frac{2(n-1)}{(n+1)^2(n+2)} .$$

Now the right side expression of (2.4.16) becomes

$$\frac{1+d^{*2}}{(d^*-1)^2} \cdot \frac{2(n-1)}{(n+1)^2(n+2)} \cdot \left(\frac{n+1}{n-1}\right)^2 \to 0 \text{ as } n \to \infty,$$

which implies that

$$\lim_{n\to\infty} P[Y_k \leqslant Y_1] = 0 \Rightarrow H_S(n) \to 1, \quad \text{as } n \to \infty.$$

This means the required sample size exists for any specified $P^* < 1$.

2.5 Normal populations with common known variance and ranked according to the absolute values of the means

Here we assume that $\prod_i$ is characterized by the normal distribution with mean μ_i and variance σ^2, which is known. The parameters θ_i of interest are defined by

(2.5.1) $$\theta_i = |\mu_i| , \quad i = 1,2,\ldots,k.$$

The problem for the case c = s = t has been considered by Rizvi (1963).

We use the absolute values of the sample means as our statistics. That is, $T_i = |\bar{X}_i|$, $i = 1,2,\ldots,k$. It has been shown by Rizvi (1963), that the probability density of T_i has a strict monotone likelihood ratio. For convenience we assume that the common variance σ^2 is unity. the probability density of T_i is

$$(2.5.2)\quad g_n(x|\theta_i) = \begin{cases} n^{\frac{1}{2}}[\phi\{n^{\frac{1}{2}}(x-\theta_i)\} + \phi\{n^{\frac{1}{2}}(x+\theta_i)\}] & \text{for } x > 0 \\ 0 & \text{for } x \leq 0, \end{cases}$$

and its distribution function is

$$(2.5.3)\quad G_n(x|\theta_i) = \begin{cases} \Phi[n^{\frac{1}{2}}(x-\theta_i)] - \Phi[n^{\frac{1}{2}}(-x-\theta_i)] & \text{for } x > 0 \\ 0 & \text{for } x \leq 0. \end{cases}$$

Here Φ and ϕ stand for the c.d.f. and p.d.f. of the standard normal distribution.

Here θ is neither a pure scale parameter nor a pure location parameter for the family $\mathcal{G}$. We define our distance measure as the difference (see (2.2.4)). Denoting $k - t - s + c$ by c' we have

$$(2.5.4)\quad Q(\theta,n) = \int_0^\infty I[G_n(x|\theta-d^*);c',s-c+1]\, dI[G_n(x|\theta);t-c+1,c].$$

Denoting $n^{\frac{1}{2}}\theta$ and $n^{\frac{1}{2}}d^*$, respectively by α and β, we obtain

$$(2.5.5)\quad Q(\alpha,\beta) = Q(\theta,n) = A\int_0^\infty I[G(x|\alpha-\beta);c',s-c+1]\, G^{t-c}(x|\alpha)[1-G(x|\alpha)]^{c-1} g(x|\alpha)dx$$
$$= A\, J(\alpha,\beta) \text{ say,}$$

where

$$(2.5.6)\quad G(x|\theta) = \Phi(x-\theta) - \Phi(-x-\theta),\ g(x|\theta) = \phi(x-\theta) + \phi(x+\theta),$$
$$A = t!/[(c-1)!(t-c)!].$$

Now we have to find the infimum of Q over the possible values of θ, or equivalently over possible values of α, for fixed value of β. Both θ and $\theta-d^*$ belong to $\Theta = [0,\infty)$. Hence θ varies over $[d^*,\infty)$. In other words α varies over $[\beta,\infty)$. The following result helps us to find the required infimum.

It is to be noted that here we need the second stage minimization in obtaining the infimum of the PCS. In the previous examples this is not the case. In this respect this example differs from the previous ones.

Lemma 2.5.1

$J(\alpha,\beta)$ is an increasing function of α, for fixed β.

Proof:

Substituting for $g(x|\alpha)$ and splitting the integral into two integrals we obtain

$$(2.5.7)\quad J(\alpha,\beta) = \int_0^\infty I[G(x|\alpha-\beta);c',s-c+1)]G^{t-c}(x|\alpha)[1-G(x|\alpha)]^{c-1}\varphi(x-\alpha)dx + \int_0^\infty I[G(x|\alpha-\beta);c',s-c+1)]G^{t-c}(x|\alpha)[1-G(x|\alpha)]^{c-1}\varphi(x+\alpha)dx.$$

By substituting u for $x - \alpha$ in the first integral and u for $x + \alpha$ in the second integral, we obtain

$$(2.5.8)\quad J(\alpha,\beta) = \int_{-\alpha}^\infty I[G(u+\alpha|\alpha-\beta);c',s-c+1)]G^{t-c}(u+\alpha|\alpha)[1-G(u+\alpha|\alpha)]^{c-1}\varphi(u)du + \int_\alpha^\infty I[G(u-\alpha|\alpha-\beta);c',s-c+1)]G^{t-c}(u-\alpha|\alpha)[1-G(u-\alpha|\alpha)]^{c-1}\varphi(u)du.$$

Using the relationship between Φ and $G(\cdot|\theta)$ and denoting $(k-t)!/[(s-c)!(c'-1)!]$ by B, from (2.5.8) we obtain on differentiation

$$(2.5.9)\quad \frac{\partial J(\alpha,\beta)}{\partial\alpha}$$

$$= 2(t-c)\int_{-\alpha}^\infty I[G(u+\alpha|\alpha-\beta);c',s-c+1]G^{t-c-1}(u+\alpha|\alpha)[1-G(u+\alpha|\alpha)]^{c-1}\varphi(u+2\alpha)\varphi(u)du$$

$$- 2(c-1)\int_{-\alpha}^\infty I[G(u+\alpha|\alpha-\beta);c',s-c+1]G^{t-c}(u+\alpha|\alpha)[1-G(u+\alpha|\alpha)]^{c-2}\varphi(u+2\alpha)\varphi(u)du$$

$$+ 2B\int_{-\alpha}^\infty G^{c'-1}(u+\alpha|\alpha-\beta)[1-G(u+\alpha|\alpha-\beta)]^{s-c}G^{t-c}(u+\alpha|\alpha)[1-G(u+\alpha|\alpha)]^{c-1}\varphi(u+2\alpha-\beta)\varphi(u)du$$

$$- 2(t-c)\int_\alpha^\infty I[G(u-\alpha|\alpha-\beta);c',s-c+1]G^{t-c-1}(u-\alpha|\alpha)[1-G(u-\alpha|\alpha)]^{c-1}\varphi(u-2\alpha)\varphi(u)du$$

$$+ 2(c-1)\int_\alpha^\infty I[G(u-\alpha|\alpha-\beta);c',s-c+1]G^{t-c}(u-\alpha|\alpha)[1-G(u-\alpha|\alpha)]^{c-2}\varphi(u-2\alpha)\varphi(u)du$$

$$- 2B\int_\alpha^\infty G^{c'-1}(u-\alpha|\alpha-\beta)[1-G(u-\alpha|\alpha-\beta)]^{s-c}G^{t-c}(u-\alpha|\alpha)[1-G(u-\alpha|\alpha)]^{c-1}\varphi(u-2\alpha+\beta)\varphi(u)du.$$

By substituting x for $u + \alpha$ in the first three integrals and x for $u - \alpha$ in the second three integrals of (2.5.9), it is easily seen that the

first and fourth integrals add to zero. Further the second and fifth integrals also add to zero. Hence we have

$$(2.5.10)\quad \frac{\partial J(\alpha,\beta)}{\partial \alpha}$$

$$= 2B \int_0^\infty G^{c'-1}(x|\alpha-\beta)[1-G(x|\alpha-\beta)]^{s-c} G^{t-c}(x|\alpha)[1-G(x|\alpha)]^{c-1}$$

$$[\varphi(x+\alpha-\beta)\varphi(x-\alpha) - \varphi(x-\alpha+\beta)\varphi(x+\alpha)]dx.$$

Since $-x \leq x$ and $\alpha - \beta < \alpha$, using the monotone likelihood ratio property of $\varphi(x|\theta) \Rightarrow \varphi(x-\theta)$ we have, for $y > 0$ and $\beta > 0$,

$$\varphi(-x-\alpha+\beta)\,\varphi(x-\alpha) > \varphi(-x-\alpha)\,\varphi(x-\alpha+\beta),$$

i.e., $\varphi(x+\alpha-\beta)\,\varphi(x-\alpha) > \varphi(x+\alpha)\,\varphi(x-\alpha+\beta)$.

Hence, from (2.5.10), it follows that $\frac{\partial J(\alpha,\beta)}{\partial \alpha}$ is positive for fixed value of β. In other words $J(\alpha,\beta)$ is an increasing function of α, for fixed β.

Thus, by the lemma 2.5.1, we have

$$(2.5.11)\quad \inf_{\alpha \geq \beta} Q(\alpha,\beta) = A \inf_{\alpha \geq \beta} J(\alpha,\beta) = A\,J(\beta,\beta) = H(n), \text{ say.}$$

Hence the required sample size is the smallest integer value of n for which

$$(2.5.12)\quad H(n) \geq P^*,$$

where $H(n) = Q(\beta,\beta)$ and can be obtained from (2.5.5).

We shall now show that $H(n)$ is an increasing function of n and it tends to one as n tends to infinity.

Lemma 2.5.2

$H(n)$ is an increasing function of n.

Proof:

From (2.5.5) we obtain

$$(2.5.13)\quad H(n) = \int_0^\infty I[G(x|0);c',s-c+1]\, dI[G(x|\beta);t-c+1,c]$$

$$= \int_0^\infty \{1-I[G(x|\beta);t-c+1,c]\}\, dI[G(x|0);c',s-c+1].$$

It is sufficient to show that for each x, $D(x,\beta)$ is an increasing function of n where

$$(2.5.14)\quad D(x,\beta) = 1 - I[G(x|\beta);t-c+1,c].$$

Now, using (2.5.6), we obtain

$$(2.5.15)\quad \frac{\partial D(x,\beta)}{\partial \beta} = -A\, G^{t-c}(x|\beta)[1-G(x|\beta)]^{c-1}[-\varphi(x-\beta) + \varphi(x+\beta)]$$

$$= A\, G^{t-c}(x|\beta)[1-G(x|\beta)]^{c-1}[\varphi(x-\beta) - \varphi(x+\beta)].$$

Since $x > 0$ and $\beta > 0$, it is easy to see that

$$\varphi(x - \beta) > \varphi(x+\beta),$$

and hence $\frac{\partial D(x,\beta)}{\partial \beta}$ is positive. Thus D is an increasing function of β (i.e., an increasing function of n). This completes the proof of the lemma.

It is easy to see that for each $x > 0$,

$$\lim_{n\to\infty} D(x,\beta) = 1 - \lim_{n\to\infty} I[G(x|\beta);t-c+1,c] = 1,$$

since $\lim_{n\to\infty} G(x|\beta) = 0$. Thus

$$\lim_{n\to\infty} H(n) = \int_0^{\infty} dI[G(x|0);c',s-c+1] = \int_0^1 dI_y[c',s-c+1] = 1.$$

Thus the required sample size is the smallest integer greater than or equal to

$$(2.5.16)\qquad n_0 = (\beta/d^*)^2$$

where β is the solution of the equation

$$(2.5.17)\quad \int_0^{\infty} I[1-G(x|\beta);c,t-c+1]\, dIG(x|0);c',s-c+1] = P^*.$$

When $c = s = t$, this equation reduced to (3.13) of Rizvi (1963). The existence and uniqueness of the solution of (2.5.17), follows from the lemma 2.5.2

When the common variance is σ^2, it is easily seen that the sample size is the smallest integer greater than or equal to

$$n_0 = (\beta\sigma/d^*)^2,$$

where β is the solution of (2.5.17).

The remarks of section 2.2 concerning the solution to the problem under different variance models are also applicable to this case.

2.6 Cauchy populations

Now we consider the problem in relation to two types of Cauchy distributions namely (i) those involving a single parameter, (ii) those involving two parameters.

(i) One parameter Cauchy distributions

Here we assume that the distribution function which characterizes $\prod_i$ is given by

$$(2.6.1) \quad F(x|\theta_i) = \frac{1}{2} + \frac{1}{\pi} \arctan(x-\theta_i),$$

so that the corresponding probability density is

$$(2.6.2) \quad f(x|\theta_i) = \frac{1}{\pi} \frac{1}{1+(x-\theta_i)^2} .$$

It is known that there does not exist a sufficient statistic of fixed-dimensions for the location parameter θ of Cauchy distribution (Koopman, 1936). Also it is known that the sample median is a consistent estimator of θ. We base our procedure on the statistics T_i, where T_i is the median of sample from $\prod_i$. For convenience, we assume that the common sample size is odd so that $n = 2m + 1$ (say).

We now show that the class of distribution functions of the sample median T, indexed by θ, is an SI family. In fact we prove a slightly more general result.

Lemma 2.6.1

Let U_r be the r^{th} order statistic in a random sample of size n from the distribution function $F(x|\theta)$, where $\{F(x|\theta); \theta \in \Theta\}$ is an SI family. The class of distribution functions of U_r, when indexed by θ, is an SI family.

Proof:

The distribution function of U_r is

$$(2.6.3)\quad H(x|\theta) = \sum_{j=r}^{n} \binom{n}{j} F^j(x|\theta)[1-F(x|\theta)]^{n-j}$$

$$= I[F(x|\theta);r,n-r+1].$$

Since for each x, F is a non-increasing function of θ, so is the function H. In other words $H(\cdot|\theta)$ constitutes an SI family of distribution functions.

Here $F(x|\theta)$ as defined by (2.6.1) is a location parameter family and hence it constitutes an SI family. Thus the distribution functions of the median T of a sample from $F(x|\theta)$ constitute an SI family. Now the c.d.f. of T is

$$(2.6.4)\quad G(x|\theta) = I[F(x|\theta);\ m+1,\ m+1],$$

and its probability density function is

$$(2.6.5)\quad g(x|\theta) = \binom{2m+1}{m} F^m(x|\theta)[1-F(x|\theta)]^m f(x|\theta).$$

Lemma 2.6.2

The densities $g(x|\theta)$ do not possess monotone likelihood ratio in x.

Proof:

Differentiating $g(x|\theta)$ with respect to x, we obtain after some simplification

$$(2.6.6)\quad \frac{\partial g(x|\theta)}{\partial x} = g(x|\theta)f(x|\theta)\ \left[\frac{m}{F(x|\theta)} - \frac{m}{1-F(x|\theta)} - 2\pi(x-\theta)\right].$$

Now differentiating both sides of (2.6.6) with respect to θ, we obtain after some simplification

$$(2.6.7)\quad \frac{\partial^2 g(x|\theta)}{\partial\theta\partial x} = \frac{1}{g(x|\theta)} \frac{\partial g(x|\theta)}{\partial x}\ \left[\frac{\partial g(x|\theta)}{\partial\theta} + 2\pi(x-\theta)g(x|\theta)f(x|\theta)\right]$$

$$+ g(x|\theta)\ f(x|\theta)\ \left[\frac{mf(x|\theta)}{F^2(x|\theta)} + \frac{mf(x|\theta)}{\{1-F(x|\theta)\}^2} + 2\pi\right],$$

$$(2.6.8)\quad \text{i.e.,}\quad g(x|\theta)\ \frac{\partial^2 g(x|\theta)}{\partial\theta\partial x} - \frac{\partial g(x|\theta)}{\partial x}\ \frac{\partial g(x|\theta)}{\partial\theta} = 2\pi(x-\theta)g(x|\theta)f(x|\theta)\ \frac{\partial g(x|\theta)}{\partial x}$$

$$+ g^2(x|\theta)f(x|\theta)\ \left\{\frac{mf(x|\theta)}{f^2(x|\theta)} + \frac{mf(x|\theta)}{[1-F(x|\theta)]^2} + 2\pi\right\}.$$

A necessary and sufficient condition for $g(x|\theta)$ to possess monotone likelihood ratio in x (Lehmann 1959, p. 111, problem 6) is that

$$(2.6.9)\quad g(x|\theta)\,\frac{\partial^2 g(x|\theta)}{\partial\theta\partial y} - \frac{\partial g(x|\theta)}{\partial\theta}\,\frac{\partial g(x|\theta)}{\partial x} \geq 0 \text{ for all } x \text{ and } \theta.$$

Using (2.6.6) and (2.6.8) this condition reduces to that for all x and θ

$$(2.6.10)\quad 2\pi(x-\theta)\; g^2(x|\theta)\; f^2(x|\theta)\left\{\frac{m}{F(x|\theta)} - \frac{m}{1-F(x|\theta)} - 2\pi(x-\theta)\right\}$$
$$+\; g^2(x|\theta)\; f(x|\theta)\left\{\frac{mf(x|\theta)}{F^2(x|\theta)} + \frac{mf(x|\theta)}{[1-F(x|\theta)]^2} + 2\pi\right\} \geq 0.$$

Thus it is necessary and sufficient that for all values of $u = x - \theta$

$$(2.6.11')\qquad g^2(u)\; f^2(u)\; h(u) \geq 0,$$

where
$$h(u) = 2\pi u\left\{\frac{m}{F(u)} - \frac{m}{1-F(u)} - 2\pi u\right\} + \frac{m}{F^2(u)} + \frac{m}{[1-F(u)]^2} + 2\pi^2(1+u^2)$$
$$= m\left\{\frac{2\pi u[1-2F(u)]}{F(u)[1-F(u)]} + \frac{1}{F^2(u)} + \frac{1}{[1-F(u)]^2}\right\} + 2\pi^2(1-u^2).$$

In order to prove the lemma it is sufficient to exhibit at least one value of u for which (2.6.11) is not true. Let $u = \sqrt{3}$ so that $\arctan\sqrt{3} = \frac{\pi}{3}$ and $F(\sqrt{3}) = \frac{5}{6}$. Now $h(\sqrt{3}) = m\left\{\frac{(-4\pi/\sqrt{3})}{\frac{5}{36}} + 26.\frac{36}{25}\right\} - 4\pi^2$

$$= \frac{36m}{25}\left\{26 - \frac{20\pi}{3}\right\} - 4\pi^2 < 0.$$

This completes the proof of the lemma.

In all the previous examples, the distribution function of T is not only SI, but it is also true that the density of T has the monotone likelihood ratio. Here the density of T does not possess the monotone likelihood ratio property although the distribution function of T is stochastically increasing. Hence this example is different from the others considered so far. Since θ is the location parameter for $G(\cdot|\theta)$, we define the distance measure as the difference (see (2.2.4)). Let m_0 be the smallest integer value of m for which

(2.6.12) $$H_L(m) \geq P^*,$$

where H_L is given by (1.7.9) in which $G_n(\cdot)$ is to be taken as $G(\cdot|0)$ of (2.6.4). The required sample size is $2m_0 + 1$.

Now we shall prove that the limit of $H_L(m)$ as $m \to \infty$ is 1, so that the required m-value exists for any $P^* < 1$. (It would be interesting to show that $H_L(m)$ is a non-decreasing function of m.) In view of the discussion in section 1.10 it is sufficient (in the notation of that section) to show that

(2.6.13) $$\lim_{m\to\infty} P[Y_k < Y_1] = 0.$$

Here Y_1 and Y_k are medians in samples of size $2m + 1$ from $F(\cdot|0)$ and $F(\cdot|d^*)$ respectively. We know that if Y_n is the median of a random sample ot size n, from the one-dimensional distribution with c.d.f. $F(x|\theta)$ and p.d.f. $f(x|\theta)$, then

$$\{2f(\zeta|\theta)\sqrt{n}\}^{-1} (Y_n - \zeta) \xrightarrow{L} N(0,1),$$

where ζ is the population median. (Cramér 1946, p. 369). Now the median of the distribution $F(\cdot|\theta)$ as defined (2.6.1) is θ. Thus

(2.6.14) $$P[Y_k < Y_1] = P[\tfrac{2}{\pi}\sqrt{2m+1}\ (Y_k - d^* + d^*) < \tfrac{2}{\pi}\sqrt{2m+1}\ Y_1]$$

$$= P[U_m + \tfrac{2}{\pi}\sqrt{2m+1}\ d^* < V_m] \text{ say.}$$

Here U_m and V_m are independent sequences of random variables, each having standard normal as the limiting distribution. Now given $\epsilon > 0$, arbitrarily small and fixed, we can find a number a such that

(2.6.15) $$1 - \Phi(a/\sqrt{2}) \leq \epsilon.$$

When m is sufficiently large, we have

(2.6.16) $$\tfrac{2}{\pi}\sqrt{2m+1}\ d^* > a,$$

so that

$$(2.6.17)\quad 0 \leq P[V_m - U_m > \frac{2}{\pi}\sqrt{2m+1}\ d^*] \leq P[V_m - U_m > a].$$

Thus

$$(2.6.18)\quad 0 \leq \lim_{m\to\infty} P[V_m - U_m > \frac{2}{\pi}\sqrt{2m+1}\ d^*] \leq \lim_{m\to\infty} P[V_m - U_m > a]$$

$$= P[V - U > a] = 1 - \Phi(a/\sqrt{2}) \leq \epsilon .$$

Since (2.6.18) is true for every $\epsilon > 0$, we have

$$(2.6.19)\quad \lim_{m\to\infty} P[V_m - U_m > \frac{2}{\pi}\sqrt{2m+1}\ d^*] = 0.$$

From (2.6.19) and (2.6.14) we obtain (2.6.13).

An approximation to the infimum of $P(CS|\vec{\theta})$

Using the limiting distribution of the sample median (as stated above) we will obtain an approximation to the infimum of $P(CS|\vec{\theta})$. Here the infimum is $H_L(m)$ which is the PCS at the point $\vec{\theta}$ for which

$$\theta_{[1]} = \cdots = \theta_{[k-t]} = 0;\ \theta_{[k-t+1]} = \theta_{[k-t+2]} = \cdots = \theta_{[k]} = d^*.$$

Thus, denoting $2d^*\sqrt{2m+1}/\pi$ by λ, we have

$$(2.6.20)\quad H_L(m) \approx \int_{-\infty}^{\infty} I_{\Phi(x+\lambda)}(c', s-c+1)\ dI_{\Phi(x)}(t-c+1, c).$$

Now an approximation to m_0 is the smallest integer greater than

$$(2.6.21)\quad m_0' = [(\frac{\lambda\pi}{2d^*})^2 - 1]/2,$$

where λ is the solution of (2.2.9). The numerical value of λ for certain values of c, k, s, t and P* can be obtained from the tables 1 and 2.

(b) Two parameter Cauchy populations with common scale parameter

Let the distribution function characterizing $\prod_i$ be

$$(2.6.22)\quad F(x|\alpha_i,\beta) = \frac{1}{2} + \frac{1}{\pi}\arctan(\frac{x-\alpha_i}{\beta}),$$

so that the corresponding probability density is

$$(2.6.23) \qquad f(x|\alpha_i,\beta) = \frac{1}{\pi} \frac{\beta}{\beta^2+(x-\alpha_i)^2} .$$

The experimenter is interested in the location parameters. In the earlier discussion we have assumed that scale parameters are all equal and that the common value is one; we now relax this assumption. Here we consider the following two cases. If the common scale parameter value is known, by making an appropriate scale transformation on the statistics we can reduce the problem to the one discussed earlier.

Suppose the common scale parameter β is unknown. By defining the preference zone in a slightly different manner we can solve the problem in the same way as we have done in an earlier section, say for example in section 2.2. We define here the distance measure as

$$(2.6.24) \qquad d(x,y) = (x-y)/\beta .$$

Again the required sample is $n = 2m_0+1$, where m_0 is the smallest integer for which (2.6.12) is true.

2.7 Laplace distributions with common scale parameter.

Suppose that the distribution characterizing the population $\prod_i$ is the Laplace distribution with the probability density

$$f(x|\theta_i,\beta) = \frac{1}{2} e^{-|x-\theta_i|/\beta} ,$$

where β is known or unknown. The experimenter is interested in the location parameters. We use the sample median as our statistic; this is the maximum likelihood estimate of θ_i. We assume that we take an odd number of observations from each of the k given populations. The class of distribution functions of the sample median, indexed by θ, is an SI family since θ is a pure location parameter for the Laplace distribution.

Now the determination of the required sample size can be carried out exactly in the same way as we did in the case of the Cauchy distributions in the previous section.

2.8 Some remarks in relation to applications to discrete distributions.

The examples, so far considered, have the common feature that $F(x|\theta)$ is the c.d.f. of an absolutely continuous distribution. In the next section we consider the problem assuming that $\prod_i$ is the Poisson population with parameter θ_1 $(i = 1,2,\ldots,k)$. In this section we make some general remarks relating to ranking and selection problems dealing with discrete distributions.

Since the distribution characterizing $\prod_i$ is a discrete distribution, so will be the distribution of the statistic T_i. In view of this fact in writing the probability of a CS for the goal of interest under the procedure R_s, we have to take into consideration the possibility of multiple ties in certain places. This results in clumsy and cumbersome expressions for the PCS. For these reasons, following Sobel (1963), we introduce a statistic T_i' with a continuous distribution corresponding to each discrete-valued statistic T_i; this transforms the problem into one dealing with continuous statistics and we can use the solution to the problem, which has been obtained in Chapter I.

Let X be a discrete-valued random variable. We shall assume that X takes on non-negative integer values. Let the probability function of X be given by

$$(2.8.1) \qquad P(X = x) = f(x|\theta), \quad x = 0, 1, 2, \ldots .$$

Corresponding to the distribution of X, we define a continuous distribution with the probability density defined by

$$(2.8.2) \qquad g(y|\theta) = \begin{cases} f([y]|\theta & \text{for } 0 \leq y < \infty \\ 0 & \text{otherwise,} \end{cases}$$

where [y] is the largest integer $\leq y$. It is easily seen that the corresponding (cumulative) distribution function is

$$(2.8.3)\quad G(y|\theta) = \sum_{j=0}^{[y]-1} f(j|\theta) + (y-[y])\, f([y]|\theta).$$

Let Y denote a random variable having the distribution defined by the probability density $g(\cdot|\theta)$ given by (2.8.2) and let $U = U(0,1)$ denote a random variable with uniform distribution on (0,1) and is independent of the random variable X.

Lemma 2.8.1

The relation between the random variables X and Y is given by

$$Y = X + U ,$$

where $U = U[0,1]$.

Proof:

$$(2.8.4)\quad P(X+U \leq y) = \sum_{j=0}^{[y]-1} P(X+U \leq y|X = j)P(X = j) + P(X+U \leq y|X = [y])P(X = [y])$$

$$= \sum_{j=0}^{[y]-1} P(U \leq y - j)f(j|\theta) + P(U \leq y - [y])f([y]|\theta)$$

$$= \sum_{j=0}^{[y]-1} f(j|\theta) + (y - [y])\, f([y]|\theta) ,$$

for $j \leq [y] - 1 \leq y - 1 \Rightarrow y - j \geq 1 \Rightarrow P(U \leq y - j) = 1$ and since $y - [y] \leq 1$ we have $P(U \leq y - [y]) = y - [y]$.

i.e., $P(X + U \leq y) = G(y|\theta) = P(Y \leq y)$.

Lemma 2.8.2

If $f(x|\theta)$ has monotone likelihood ratio in x, then $g(y|\theta)$ has monotone likelihood ratio in y.

Proof:

Let $x_i = [y_i]$ $(i = 1,2)$; then $y_1 \leq y_2 \not\Rightarrow x_i \leq x_2$. Since f has m.l.r property, for $\theta_1 \leq \theta_2$ and $y_1 \leq y_2$ we have $x_1 \leq x_2$ and

$$(2.8.5)\quad g(y_1|\theta_1)\, g(y_2|\theta_2) = f(x_1|\theta_1)\, f(x_2|\theta_2)$$

$$\geq f(x_1|\theta_2)\, f(x_2|\theta_1) = g(y_1|\theta_2)\, g(y_2|\theta_1).$$

This proves the lemma.

The above lemma shows that the m.l.r. property is preserved when we transform the discrete distributions into absolutely continuous distributions in the manner described above. Now we will show that if the discrete distributions is an SI family, (without possessing m.l.r. property) the transformation into absolutely continuous distributions preserve this property also.

Lemma 2.8.3

Let $\mathcal{F}$ be an SI family of distributions defined by probability mass functions $f(x|\theta)$ where $\theta \in \Theta$. Then the class $\mathcal{G}$ of the corresponding absolutely continuous distributions defined by (2.8.2), is also an SI family.

Proof:

By the definition of an SI family we have

$$F(x|\theta_1) \geq F(x|\theta_2) \tag{2.8.6}$$

for all $x \in R$ and $\theta_1 < \theta_2$, where $F(\cdot|\theta_i)$ is the c.d.f. of the distribution defined by the probability mass function $f(\cdot|\theta_i)$ $(i = 1,2)$.

Let X_i be the random variables associated with the distribution defined by $F(\cdot|\theta_i)$ $(i = 1,2)$. Then the random variables associated with the corresponding continuous distributions defined by $G(\cdot|\theta_1)$ and $G(\cdot|\theta_2)$ are $X_1 + U_1$ and $X_2 + U_2$, where U_i is a uniform random variable on $(0,1)$ which is independent of X_i $(i = 1,2)$. Thus

$$\begin{aligned} G(y|\theta_1) &= P(X_1 + U_1 \leq y) = P(X_1 \leq y - U_1) \\ &= \int_0^1 F(y - u|\theta_1)du, \\ &\geq \int_0^1 F(y - u|\theta_2)du, \quad \text{from (2.8.6)} \\ &= G(y|\theta_2). \end{aligned} \tag{2.8.7}$$

Hence the lemma.

2.9 Poisson populations.

Here we assume that $\prod_i$ is characterized by the Poisson distribution with parameter θ_i $(i = 1,2,\ldots,k)$. It has been pointed out by Sobel (1963) that in the case of the goal of choosing the 'best' population, the solution based only on the parameter differences or only on the parameter ratios does not exist. He obtained a solution based on the simultaneous consideration of differences and ratios. Here also the same is true. Thus following Sobel (1963), we define the preference zone $\Omega(d^*,r^*)$ as

$$(2.9.1) \qquad \Omega(d^*,r^*) = \{\vec{\theta};\theta_{[k-t+1]} - \theta_{[k-t]} \geq d^* \text{ and } (\theta_{[k-t+1]}/\theta_{[k-t+1]}) \geq r^*\}.$$

Here $r^*(>1)$ and $d^*(>0)$ are specified numbers.

We use the statistics T_i where

$$(2.9.2) \qquad T_i = \sum_{j=1}^{n} X_{ij}, \qquad (i = 1,2,\ldots,k).$$

It is well known that T_i is distributed as a Poisson variable with the parameter $n\theta_i = \psi_i (i = 1,2,\ldots,k)$.

In view of the remarks of section 2.8, after computing the statistics T_i from the random samples, we transform them into observations on the variables Y_i by adding to each T_i a random observation from the uniform distribution over $[0,1]$. We apply the procedure R_s to Y_i's. Here the probability density of Y_i is

$$(2.9.3) \qquad g(y|\psi_i) = \begin{cases} e^{-\psi_i} \psi_i^{[y]} / ([y])! & \text{for } 0 < y < \infty \\ 0 & \text{otherwise.} \end{cases}$$

We shall denote the c.d.f. of Y_i by $G(y|\psi_i)$. Let us note the fact that the distribution $G(\cdot|\psi)$ forms an SI family, when indexed by ψ. Now theorem 1.6.1 is applicable. So in finding the infimum of the PCS over $\Omega(d^*,r^*)$, it is sufficient to confine our attention to the points in the GLF configuration and find the infimum of the PCS over such points. The points $\vec{\theta}$, which are in GLF configuration, are those points for which

$$(2.9.4)\quad \begin{cases} \theta_{[1]} = \theta_{[2]} = \cdots = \theta_{[k-t]} = \theta' \text{ (say)} \\ \theta_{[k-t+1]} = \theta_{[k-t+2]} = \cdots = \theta_{[k]} = \theta \text{ (say)}. \end{cases}$$

Let $\psi' = n\theta'$, $\psi = n\theta$, and let the PCS at the GLF configuration be denoted by $I(\psi,\psi')$. This corresponds to $P(\theta,\theta_0)$ of the discussion in section 1.6. Now

$$(2.9.5)\quad \frac{(c'-1)!(s-c)!}{(k-t)!}\ I(\psi,\psi') = \sum_{\alpha=0}^{t-c} \binom{t}{\alpha} J(\alpha;\psi,\psi') \text{ (say)},$$

where

$$(2.9.6)\quad J(\alpha;\psi,\psi') = \int_0^\infty G^{\alpha}(x|\psi)[1-G(x|\psi)]^{t-\alpha}\ G^{c'-1}(x|\psi')[1-G(x|\psi')]^{s-c}\ g(x|\psi')dx.$$

We denote the entire sum on the right side of (2.9.5) by $J(\psi,\psi')$. Now we have to find the infimum of $I(\psi,\psi')$, equivalently the infimum of $J(\psi,\psi')$ over the pairs (ψ,ψ') where

$$(2.9.7)\quad \psi/\psi' \geq r^* \quad \text{and} \quad \psi \leq \psi' \geq nd^*.$$

Remark

Here is another example where we need minimization of the PCS at GLF configuration. It may be noted that θ is neither a location parameter nor a scale parameter for the distribution $G(\cdot|n\theta)$.

Now we give some known results concerning $g(\cdot|\psi)$ and $G(\cdot|\psi)$ which are used in proving a theorem which in turn used to obtain the infimum of $J(\psi,\psi')$.

We know that $G(\cdot|\psi)$ is given by

$$(2.9.8)\quad G(y|\psi) = \sum_{j=0}^{[y]-1} f(j|\psi) + (y - [y])\ f([y]|\psi),$$

where

$$(2.9.9)\quad f(j|\psi) = e^{-\psi}\ \psi^j/j!\ .$$

Now

$$(2.9.10)\quad \frac{d}{d\psi}\ [g(y|\psi)] = g(y-1|\psi) - g(y|\psi) = g(y|\psi)\{\frac{[y]}{\psi} - 1\},$$

$$(2.9.11)\quad \frac{d}{d\psi}\,[G(y|\psi)] = (y - [y] - 1)\, g(y - 1|\psi) - (y - [y])\, g(y|\psi)$$

$$= g(y|\psi)\,\{(y - [y] - 1)\,\frac{[y]}{\psi} - (y - [y])\}\,.$$

Further

$$(2.9.12)\quad \Delta\, G(y|\psi) = G(y|\psi) - G(y - 1|\psi) = -\frac{d}{d\psi}\,[G(y|\psi)]\,.$$

Theorem 2.9.1

For $k \geq 2$, letting $\psi = a\psi' + b$ where $\psi' > 0$, $a \geq 1$ and $b \geq 0$ we have

$$(2.9.13)\quad \frac{d}{d\psi^r}\,[J(\psi,\psi')] \begin{cases} \geq 0 & \text{if } a \geq 1 \text{ and } b = 0 \quad \text{case 1} \\ \leq 0 & \text{if } a = 1 \text{ and } b \geq 0 \quad \text{case 2.} \end{cases}$$

The strict inequality holds when $a > 1$ in case 1 and when $b > 0$ in case 2.

Proof:

Let $D(\psi')$ stand for the derivative of J with respect to ψ'. Then

$$(2.9.14)\quad D(\psi') = \sum_{\alpha=0}^{t-c} \binom{t}{\alpha}\, \frac{d}{d\psi^r}\,[J(\alpha; a\psi' + b, \psi')].$$

Now

$$(2.9.15)\quad \frac{d}{d\psi^r}\,[J(\alpha; a\psi' + b, \psi')] =$$

$$- a\alpha \int_0^\infty G^{\alpha-1}(x|\psi)[1-G(x|\psi)]^{t-\alpha} G^{c'-1}(x|\psi')[1-G(x|\psi')]^{s-c} g(x|\psi')[\Delta G(x|\psi)]dx$$

$$+ a(t-\alpha)\int_0^\infty G^{\alpha}(x|\psi)[1-G(x|\psi)]^{t-\alpha-1} G^{c'-1}(x|\psi')[1-G(x|\psi')]^{s-c} g(x|\psi')[\Delta G(x|\psi)]dx$$

$$- (c'-1) \int_0^\infty G^{\alpha}(x|\psi)[1-G(x|\psi)]^{t-\alpha} G^{c'-2}(x|\psi')[1-G(x|\psi')]^{s-c} g(x|\psi')[\Delta G(x|\psi')]dx$$

$$+ (s-c) \int_0^\infty G^{\alpha}(x|\psi)[1-G(x|\psi)]^{t-\alpha} G^{c'-1}(x|\psi')[1-G(x|\psi')]^{s-c-1} g(x|\psi')[\Delta G(x|\psi')]dx$$

$$+ \int_0^\infty G^{\alpha}(x|\psi)[1-G(x|\psi)]^{t-\alpha} G^{c'-1}(x|\psi')[1-G(x|\psi')]^{s-c} [g(x-1|\psi') - g(x|\psi')]dx\,.$$

Denoting the last integral on the right side of (2.9.15) by I_4 we have

$$(2.9.16)\quad I_4 = \int_0^\infty G^{\alpha}(x|\psi)[1-G(x|\psi)]^{t-\alpha} G^{c'-1}(x|\psi')[1-G(x|\psi')]^{s-c} \frac{d}{dx}\,[-\Delta G(x|\psi')]dx\,.$$

Integration by parts gives

$$(2.9.17)\quad I_4 = \alpha \int_0^\infty [\Delta G(x|\psi')]G^{\alpha-1}(x|\psi)[1-G(x|\psi)]^{t-\alpha}G^{c'-1}(x|\psi')[1-G(x|\psi')]^{s-c}g(x|\psi)dx$$

$$- (t-\alpha) \int_0^\infty [\Delta G(x|\psi')]G^{\alpha}(x|\psi)[1-G(x|\psi)]^{t-\alpha-1}G^{c'-1}(x|\psi')[1-G(x|\psi')]^{s-c}g(x|\psi)dx$$

$$+ (c'-1) \int_0^\infty [\Delta G(x|\psi')]G^{\alpha}(x|\psi)[1-G(x|\psi)]^{t-\alpha}G^{c'-2}(x|\psi')[1-G(x|\psi')]^{s-c}g(x|\psi')dx$$

$$- (s-c) \int_0^\infty [\Delta G(x|\psi')]G^{\alpha}(x|\psi)[1-G(x|\psi)]^{t-\alpha}G^{c'-1}(x|\psi')[1-G(x|\psi')]^{s-c-1}g(x|\psi')dx.$$

Using (2.9.17) in (2.9.15), we obtain, after cancellation

$$(2.9.18)\quad \frac{d}{d\psi'} J(\alpha;a\psi' + b,\psi')$$

$$= (t-\alpha) \int_0^\infty G^{\alpha}(x|\psi)[1-G(x|\psi)]^{t-\alpha-1}\, G^{c'-1}(x|\psi')[1-G(x|\psi')]^{s-c}$$

$$g(x|\psi)g(x|\psi')\{(a-1)(x-[x]) + \frac{[x]b}{\psi\psi'}\ (x-[x]-1)\}dx$$

$$+\alpha \int_0^\infty G^{\alpha-1}(x|\psi)[1-G(x|\psi)]^{t-\alpha}\, G^{c'-1}(x|\psi')[1-G(x|\psi')]^{s-c}$$

$$g(x|\psi)g(x|\psi')\{(a-1)(x-[x]) + \frac{[x]b}{\psi\psi'}\ (x-[x]-1)\}dx\ .$$

Since $x \geq [x] > x - 1$, it follows that for $a \geq 1$ and $b = 0$ the integrals of (2.9.18) are non-negative and that for $a = 1$ and $b \geq 0$ the integrals are non-positive. To prove strict positiveness or negativeness of the integrals we first note the fact that $\psi' > 0$ so that $\psi > 0$. Further both the densities are non -degenerate. Since $x - \{x\} = 0$ only at integers and $[x] = 0$ only for $0 \leq x < 1$, it follows that for $a > 1$ and $b = 0$ the term $(a-1)(x-[x])$ is strictly positive and that for $b > 0$ and $a = 1$

$$[x]b\ (x - [x] - 1) < 0.$$

Hence

$$\frac{d}{d\psi'} J(\alpha;a\psi' + b,\psi') \quad \begin{cases} \geq 0 & \text{for } a \geq 1 \text{ and } b = 0 \\ \leq 0 & \text{for } a = 1 \text{ and } b \geq 0\ , \end{cases}$$

and

$$\frac{d}{d\psi'} J(\alpha; a\psi' + b, \psi') \quad \begin{cases} > 0 & \text{for } a > 1 \text{ and } b = 0 \\ < 0 & \text{for } a = 1 \text{ and } b > 0. \end{cases}$$

From (2.9.14) the required result follows.

From the second part of the theorem, we can find the pair (ψ,ψ') satisfying (2.9.7) for which $J(\psi,\psi')$ is minimum. In other words we can find the pair (θ,θ') such that

$$\theta/\theta' \geq r^* \quad \text{and} \quad \theta - \theta' \geq d^* \,, \tag{2.9.19}$$

for which $P(CS|GLF)$ is minimum and hence we can find the required sample size.

We can rewrite (2.9.5) as

$$I(\psi,\psi') = \int_0^\infty [1 - I_{G(x|\psi)}(t-c+1,c)] \, dI_{G(x|\psi')}(c',s-c+1). \tag{2.9.20}$$

Let us note the fact that the region (2.9.19) is such that by decreasing θ (with θ' fixed) we can change at least one of the inequalities in (2.9.19) to an equality. By the monotone (decreasing) nature of G and the form of the integral I, as given in (2.9.20), it is easy to see that any such decrease in θ will not increase I. Hence we can restrict our attention to points $\vec{\theta} = (\theta,\theta')$ on at least one of the two lines L_1, L_2 given by

$$\begin{aligned} L_1 &: \ \theta/\theta' = r^* \qquad (r^* > 1), \\ L_2 &: \ \theta-\theta' = d^* \qquad (\delta^* > 0). \end{aligned} \tag{2.9.21}$$

By the second part of the theorem 2.9.1, I is strictly increasing in θ' on L_1 and it is strictly decreasing in θ' on L_2. Hence I is minimum at the point where the two lines meet, i.e., at

$$\theta_0' = \frac{d^*}{r^*-1}, \quad \theta^0 = \frac{r^*d^*}{r^*-1}. \tag{2.9.22}$$

Thus denoting $n\theta^0$ and $n\theta_0'$ by ψ_0 and ψ_0' we have

$$\inf P(CS|\vec{\theta}) = I(\psi_0,\psi_0').$$

Hence the required sample size is the smallest integer value of n for which

$$I(n\theta^0, n\theta_0') \geq P^* .$$

(2.9.23) i.e., $$\int_0^\infty [1 - I_{G(x|\psi_0)}(t-c+1, c)] \, dI_{G(x|\psi_0')}(c', s-c+1) \geq P^* .$$

When $s = t = c = 1$, our goal reduces to the goal of selceting the 'best' population which is considered by Sobel (1963).

We shall now show that

(2.9.24) $$\lim_{n\to\infty} I(\psi_0, \psi_0') = 1.$$

From the discussion in section 1.10, it is sufficient to show that

(2.9.25) $$\frac{\text{Var } Y_k + \text{Var } Y_1}{(EY_k - EY_1)^2} \to 0, \quad \text{as} \quad n \to \infty.$$

Here $E(Y_1) = \psi_0' + \frac{1}{2}$, $E(Y_k) = \psi_0 + \frac{1}{2}$,

and $\text{Var}(Y_1) = \psi_0' + \frac{1}{12}$, $\text{Var}(Y_k) = \psi_0 + \frac{1}{12}$.

Thus

$$\frac{\text{Var } Y_k + \text{Var } Y_1}{\{E(Y_k - Y_1)\}^2} = \frac{\psi_0 + \psi_0' + (1/6)}{(\psi_0 - \psi_0')^2} = \frac{\theta^0 + \theta_0'}{n(\theta^0 - \theta_0')^2} + \frac{1}{6n^2(\theta^0 - \theta_0')^2} .$$

Hence (2.9.25) is true which implies (2.9.24). In other words the required sample size exists for any $P^* < 1$.

Chapter III

Some Properties of the Procedure R_s

In this chapter we prove some properties of the procedure R_s. For convenience, in this discussion we assume that the labels on the populations are such that the parameter associated with $\prod_i$ is θ_i where $\theta_1 \leq \theta_2 \leq \ldots \leq \theta_k$.

3.1. Multivariate unbiasedness of the Procedure R_s.

Let $\alpha_1, \alpha_2, \ldots, \alpha_s, \alpha_{s+1} \equiv k$ and β be integers such that

$$(3.1.1) \qquad 1 \leq \alpha_1 < \alpha_2 < \ldots < \alpha_{i-1} < \beta < \alpha_i < \ldots < \alpha_s \leq k ,$$

for some i where $2 \leq i \leq s+1$. Let I be the set of integers $1,2,\ldots,k$ and J be the set $I - \{\alpha_1,\ldots,\alpha_{i-1},\beta,\alpha_i,\ldots,\alpha_s\}$. Further let $P(\alpha_1,\alpha_2,\ldots,\alpha_s)$ denote the probability of selecting the populations $\prod_{\alpha_1}, \prod_{\alpha_2}, \ldots, \prod_{\alpha_s}$ under the procedure R_s. Since the procedure R_s is based on the statistics T_i, the probabilities of interest are functions of the distribution functions $G_n(\cdot|\theta_i)$. In this discussion the sample size n is some fixed (positive) integer and so we drop the subscript n of $G_n(\cdot|\theta)$. The results to be proved are based on one of the following assumptions.

Assumption 3.1.1: The family $\mathcal{G} = \{G_n(\cdot|\theta) : \theta \in \Theta\}$ is a stochastically increasing family for each value of n.

Assumption 3.1.1': The family $\mathcal{G}$ is a strictly stochastically increasing family for each value of n.

Lemma 3.1.1

Under the assumption 3.1.1 we have

$$(3.1.2) \qquad P(\alpha_1, \ldots, \alpha_{i-1}, \alpha_i, \alpha_{i+1}, \ldots, \alpha_s) \geq P(\alpha_1, \ldots, \alpha_{i-1}, \beta, \alpha_{i+1}, \ldots, \alpha_s).$$

If the assumption 3.1.1' holds and $\theta_\beta < \theta_{\alpha_i}$, the inequality in (3.1.2) is a strict inequality.

Proof:

First we consider the case $s \leq k-1$.

Let $U = \max_{\alpha \in J} T_\alpha$ and $H(u)$ be its c.d.f. Using assumption 3.1.1,

$$(3.1.3)\quad \beta < \alpha_i \Rightarrow \theta_\beta \leq \theta_{\alpha_i} \Rightarrow G(u|\theta_\beta) \geq G(u|\theta_{\alpha_i}) \text{ for all real } u,$$
$$\Rightarrow 1 - G(u|\theta_\beta) \leq 1 - G(u|\theta_{\alpha_i}) \text{ for all real } u.$$

Further

$$(3.1.4)\quad P(\alpha_1,\ldots,\alpha_{i-1},\alpha_i,\alpha_{i+1},\ldots,\alpha_s)$$
$$= P[\min_{1\leq m\leq s} T_{\alpha_m} > \max(T_\beta, U)]$$
$$= \int_{-\infty}^{\infty} H(u)G(u|\theta_\beta)d[1 - \prod_{m=1}^{s}\{1 - G(u|\theta_{\alpha_m})\}]$$
$$\geq \int_{-\infty}^{\infty} H(u)G(u|\theta_{\alpha_i})d[1 - \prod_{m=1}^{s}\{1 - G(u|\theta_{\alpha_m})\}] \quad \text{by (3.1.3)}$$
$$= 1 - \int_{-\infty}^{\infty}[1 - \prod_{m=1}^{s}\{1 - G(u|\theta_{\alpha_m})\}]d[H(u)G(u|\theta_{\alpha_i})]$$
$$= \int_{-\infty}^{\infty}\prod_{m=1}^{s}\{1 - G(u|\theta_{\alpha_m})\}d[H(u)G(u|\theta_{\alpha_i})]$$
$$\geq \int_{-\infty}^{\infty}[\prod_{\substack{m=1\\ m\neq i}}^{s}\{1 - G(u|\theta_{\alpha_m})\}]\{1 - G(u|\theta_\beta)\}d[H(u)G(u|\theta_{\alpha_i})] \quad \text{by (3.1.3)}$$
$$= P(\alpha_1,\ldots,\alpha_{i-1},\beta,\alpha_{i+1},\ldots,\alpha_s).$$

When $s = k-1$, the proof is similar to the above with $H(u)$ replaced by 1 and hence omitted. If $\theta_\beta < \theta_{\alpha_i}$ and if the assumption 3.1.1' holds, then the inequalities in (3.1.3) will be strict inequalities and consequently we get a strict inequality in the final result. This completes the proof of the lemma.

Remark 3.1.1

If $\alpha_1 \geq 2$ and $1 \geq \beta < \alpha_1$, by a reasoning similar to the above we have

$$P(\alpha_1, \alpha_2, \ldots, \alpha_s) \geq P(\beta, \alpha_2, \ldots, \alpha_s).$$

Let $\alpha_1, \alpha_2, \ldots, \alpha_s$ and $\beta_1, \beta_2, \ldots, \beta_s$ be two subsets of I such that

$\alpha_1 < \alpha_2 < \ldots < \alpha_s$, $\beta_1 < \beta_2 < \ldots < \beta_s$ and $\alpha_i \geq \beta_i$, $i = 1,2,\ldots,s$, with $\alpha_i > \beta_i$ for at least one value of i. Then the subset of populations $\prod_{\beta_1}, \prod_{\beta_2}, \ldots, \prod_{\beta_s}$ is said to be inferior to the subset of populations $\prod_{\alpha_1}, \prod_{\alpha_2}, \ldots, \prod_{\alpha_s}$. Now a repeated application of the lemma gives us the following

Theorem 3.1.1

Under the assumption 3.1.1

$$(3.1.5) \qquad P(\alpha_1, \alpha_2, \ldots, \alpha_s) \geq P(\beta_1, \beta_2, \ldots, \beta_s).$$

Remark 3.1.2

If $\theta_{\alpha_i} > \theta_{\beta_i}$ for at least one i and if the assumption 3.1.1' holds then strict inequality holds in (3.1.5).

Thus the probability of selecting any subset $\mathcal{G}$ of s populations is not less than the probability of selecting any subset of size s, which is inferior to $\mathcal{G}$. In this sense, the procedure R_s possesses the property of multivariate unbiasedness.

Let $q(\alpha)$ denote the probability of including the population $\prod_\alpha$ in the selected subset under the procedure R_s. As a direct consequence of the above theorem we get the following property of unbiasedness.

Corollary 3.1.1

Under the assumption 3.1.1

$$(3.1.6) \qquad q(\alpha) \leq q(\beta) \quad \text{for} \quad \alpha < \beta.$$

We now give an independent proof of this result.

Proof:

Let $q(\alpha,\bar{\beta})$ denote the probability of including $\prod_\alpha$ and excluding

$\prod_\beta$ in the selected subset under the procedure R_s. Further let V be the random variable corresponding to the $(s-1)^{\underline{st}}$ largest of $T_\gamma (\gamma = 1,2,\ldots,k;\ \gamma \neq \alpha,\beta)$ and H(v) be its c.d.f. Now

$$(3.1.7)\quad q(\alpha) - q(\beta) = q(\alpha,\bar{\beta}) - q(\bar{\alpha},\beta)$$

$$= P[T_\beta < V < T_\alpha] - P[T_\alpha < V < T_\beta]$$

$$= \int_{-\infty}^{\infty} G(v|\theta_\beta)[1-G(v|\theta_\alpha)]dH(v) - \int_{-\infty}^{\infty} G(v|\theta_\alpha)[1-G(v|\theta_\beta)]dH(v)$$

$$= \int_{-\infty}^{\infty} [G(v|\theta_\beta) - G(v|\theta_\alpha)]dH(v).$$

Since $\alpha < \beta$, we have

$$(3.1.8)\qquad \theta_\alpha \leq \theta_\beta \Rightarrow G(v|\theta_\beta) \leq G(v|\theta_\alpha) \text{ for all real } v.$$

Hence $q(\alpha) - q(\beta) \leq 0$.

Remark 3.1.2

Also if $\theta_\alpha < \theta_\beta$ and if the assumption 3.1.1' is satisfied, then strict inequality holds in (3.1.8) and hence in the result (3.1.6).

3.2 An optimal property.

In this section we prove that the procedure R_s is the uniformly best decision rule among the impartial decision rules for the loss function

(3.2.1) $$W = c - \eta ,$$

where η denote the number of populations from the set of the t 'best' populations, that are included in the selected subset of size s.

Before proving this result we specify our assumptions and define some terms which will be used later.

We assume that a random sample of size n is available from each of the k given populations. That is, we are given independent random variables $\{X_{ij}\}$, $i = 1,2,\ldots,k; j = 1,2,\ldots,n$, from the k populations $\prod_i$. Let

(3.2.2) $$T_i = T(X_{i1}, X_{i2}, \ldots, X_{in}), \quad i = 1,2,\ldots,k,$$

where $T_1, T_2, \ldots, T_k$ is an independent set, and T_i has the probability density $g(\cdot|\theta_i) = g_i(\cdot)$ such that $\theta_1 \leq \theta_2 \leq \ldots \leq \theta_k$. Further we assume that the densities $g(\cdot|\theta)$ possess monotone likelihood ratio property. Let $T_{[1]} < T_{[2]} < \ldots < T_{[k]}$ be the ordered T_i's.

Following Bahadur (1950) we now define a class of decision rules which shall be called the class of non-randomized impartial decision rules. This class of decision rules are based on the statistics $\{T_i\}$ and it is denoted by D(T).

Definition 3.2.1. A variable Y is said to be an indicator variable corresponding to the event E if Y = 1 when E happens and Y = 0 otherwise.

Definition 3.2.2. A decision rule $\delta = \delta(\{T_i\})$ is said to be an impartial non-randomized decision rule if δ defines non-negative random variables $\lambda_j(T_{[1]}, T_{[2]}, \ldots, T_{[k]})$, $j = 1,2,\ldots,k$ (depending only on the ordered T_i's) such that λ_j is the indicator variable corresponding to the event that the population which gave $T_{[j]}$ is selected under the decision rule δ and $\sum_{j=1}^{k} \lambda_j = s$. (In other words the impartial decision rules are those rules which are

invariant with respect to relabeling of the populations.)

Let A_{ij} be the event $\{T_i = T_{[j]}\}$ and a_{ij} be the indicator variable corresponding to the event A_{ij}. Since T_i's have a joint distribution which is absolutely continuous, the sets A_{ij} are well defined with probability one. Further we have $\sum_{i=1}^{k} a_{ij} = 1$ for every j and $\sum_{j=1}^{k} a_{ij} = 1$ for every i, with probability one.

We now give, without proof, a lemma due to Bahadur (1950) which will be used subsequently.

Lemma 3.2.1

For any non-negative random variable $\lambda = \lambda(T_{[1]}, T_{[2]}, \ldots, T_{[k]})$ and any $p,q,m = 1,2,\ldots,k$ with $p \leq q$, we have

$$(3.2.3) \qquad \sum_{i=m}^{k} E(\lambda a_{ip}) \leq \sum_{i=m}^{k} E(\lambda a_{iq}).$$

A direct consequence of this lemma is the following result.

Corollary 3.2.1

Under the assumptions of lemma 3.2.1, we have

$$(3.2.4) \qquad \sum_{i=1}^{m} E(\lambda a_{ip}) \geq \sum_{i=1}^{m} E(\lambda a_{iq}).$$

In the sequel we further assume that the parameter space Ω is defined by

$$(3.2.5) \qquad \Omega = \{\vec{\theta} : \theta_{k-t+1} > \theta_{k-t}\}.$$

Let b be any monotone non-decreasing function of θ so that $b_i = b(\theta_i)$, $i = 1,2,\ldots,k$. For $i < j$, we have $b_i \leq b_j$, since $\theta_i \leq \theta_j$. Any decision rule $\delta \epsilon D(T)$ defines a vector random variable $\lambda(\delta) = [\lambda_1(\delta), \lambda_2(\delta), \ldots, \lambda_k(\delta)]$ and this vector in turn defines another vector random variable $p(\delta) = [p_1(\delta), p_2(\delta), \ldots, p_k(\delta)]$ where

$$(3.2.6) \qquad p_i(\delta) = \sum_{j=1}^{k} \lambda_j(\delta) a_{ij}, \qquad i = 1,2,\ldots,k.$$

Here $p_i(\delta)$ can be interpreted as an indicator variable corresponding to the

event that the population $\prod_i$ is selected under the decision rule δ. It is easy to see that $\sum_{i=1}^{k} p_i(\delta) = s$. For any $\delta \epsilon D(T)$, let us define

$$(3.2.7) \qquad \rho(\delta|\vec{\theta}) = E[\sum_{i=1}^{k} b_i p_i(\delta)|\vec{\theta}].$$

Theorem 3.2.1

For every $\vec{\theta}\epsilon\Omega$

$$\rho(\delta^0|\vec{\theta}) = \sup_{\delta\epsilon D(T)} \rho(\delta|\vec{\theta}),$$

and

$$\rho(\delta_0|\vec{\theta}) = \inf_{\delta\epsilon D(T)} \rho(\delta|\vec{\theta}),$$

where δ^0 and δ_0 are those decision rules belonging to $D(T)$ which correspond to the λ-vectors $(0,0,\ldots,0,1,\ldots,1)$ and $(1,1,\ldots,1,0,\ldots,0)$.

Proof:

Now for any $\delta\epsilon D(T)$ and for any $\vec{\theta}\epsilon\Omega$, we have

$$(3.2.8) \qquad \rho(\delta|\vec{\theta}) = E[\sum_{i=1}^{k} b_i p_i(\delta)|\vec{\theta}]$$

$$= E[\sum_{i=1}^{k} b_i \{\sum_{j=1}^{k} \lambda_j(\delta) a_{ij}\}|\vec{\theta}] = \sum_{i,j=1}^{k} b_i E[\lambda_j(\delta) a_{ij}|\vec{\theta}].$$

We know that $b_1 \leq b_2 \leq \ldots \leq b_k$ and we write

$$(3.2.9) \qquad b_i = b_1 + c_1 + c_2 + \ldots + c_i, \quad c_i \geq 0, \quad i = 1,2,\ldots,k.$$

Let $\lambda(\delta)$ be such that $\lambda_{j_1}(\delta) = \ldots = \lambda_{j_s}(\delta) = 1$, where $j_1 < \ldots < j_s$. Further let the other components of $\lambda(\delta)$ be zero. Let us note that $\lambda_{k-s+1}(\delta^0) = \ldots = \lambda_k(\delta^0) = 1$ and the remaining components of $\lambda(\delta^0)$ are zero. Also $\lambda_1(\delta_0) = \ldots = \lambda_s(\delta_0) = 1$ and the remaining components of $\lambda(\delta_0)$ are zero. Now

$$(3.2.10) \qquad \alpha \leq j_\alpha \leq k - s + \alpha \quad \text{for} \quad \alpha = 1,2,\ldots,s,$$

with at least one of the inequalities being strict. It is easily seen that

$$\sum_{i,j=1}^{k} b_i E\{\lambda_j(\delta) a_{ij}\} = \sum_{i,j=1}^{k} (b_1+c_1+\ldots+c_i)\, E\{\lambda_j(\delta) a_{ij}\}$$

$$= b_1 s + \sum_{m=1}^{k} \sum_{\alpha=1}^{s} \Big[\sum_{i=m}^{k} E\{a_{i j_\alpha}\} \Big] c_m$$

$$\leq b_1 s + \sum_{m=1}^{k} \sum_{\alpha=1}^{s} \Big[\sum_{i=m}^{k} E\{a_{i,k-s+\alpha}\} \Big] c_m$$

$$= \sum_{i,j=1}^{k} b_i E\{\lambda_j(\delta^0) a_{ij}\}.$$

In obtaining the inequality we made use of (3.2.10) and the lemma 3.2.1.

(3.2.11) i.e., $\rho(\delta|\vec{\theta}) \leq \rho(\delta^0|\vec{\theta})$ for all $\vec{\theta}\in\Omega$.

Further by a similar argument we have

$$\Sigma\, b_i E\{\lambda_j(\delta_0) a_{ij}\} \leq \Sigma\, b_i E\{\lambda_j(\delta) a_{ij}\}.$$

(3.2.12) i.e., $\rho(\delta|\vec{\theta}) \geq \rho(\delta_0|\vec{\theta})$ for all $\vec{\theta}\in\Omega$.

From (3.2.11) and (3.2.12) we obtain, for every $\vec{\theta}\in\Omega$

$$\sup_{\delta\in D(T)} \rho(\delta|\vec{\theta}) = \rho(\delta^0|\vec{\theta}),$$

$$\inf_{\delta\in D(T)} \rho(\delta|\vec{\theta}) = \rho(\delta_0|\vec{\theta}).$$

The class of decision rules D(T) consists of $\xi = \binom{k}{s}$ members and let us denote them by $\delta_1,\delta_2,\ldots,\delta_\xi$. Let $D^*(T) \supset D(T)$ be the class of impartial decision rules based on $T = (T_1,T_2,\ldots T_k)$ such that any typical member δ^* of $D^*(T)$ can be represented by a vector with ξ components viz., $(\varphi_1,\varphi_2,\ldots,\varphi_\xi)$; here φ_α is the probability of choosing the non-randomized decision rule $\delta_\alpha\in D(T)$ and $\sum_{\alpha=1}^{\xi} \varphi_\alpha = 1$. Now it is easy to see that for any $\vec{\theta}\in\Omega$ and for any $\delta^*\in D^*(T)$ from (3.2.11) and (3.2.12)

(3.2.13) $$\rho(\delta^0|\vec{\theta}) \geq \rho(\delta^*|\vec{\theta}),$$

and

(3.2.14) $$\rho(\delta_0|\vec{\theta}) \leq \rho(\delta^*|\vec{\theta}).$$

Hence for any $\vec{\theta}\in\Omega$

$$\sup_{\delta^*\in D^*(T)} \rho(\delta^*|\vec{\theta}) = \rho(\delta^0|\vec{\theta}), \tag{3.2.15}$$

$$\inf_{\delta^*\in D^*(T)} \rho(\delta^*|\vec{\theta}) = \rho(\delta_0|\vec{\theta}). \tag{3.2.16}$$

Let us consider the problem of choosing a subset of size s from a set of k given populations. We confine our attention to the class of impartial decision rules D*(T) defined above. The problem is to find the uniformly best decision rule in the class D*(T) when the loss function is defined to be

$$W = c - \eta, \tag{3.2.17}$$

where η is the number of populations, from the set of the t best out of k given populations, that enter into the selected subset. Here the t best populations are those with the parameters $\theta_{k-t+1}, \theta_{k-t+2}, \ldots, \theta_k$. Let us define the b-function as follows:

$$b(\theta_i) = \begin{cases} 1 & \text{for } i = k-t+1, k-t+2, \ldots, k \\ 0 & \text{otherwise.} \end{cases} \tag{3.2.18}$$

Then for any $\delta^*\in D^*(T)$, using the definition (3.2.18) of b-function,

$$E\eta(\delta^*) = E\{\sum_{i=1}^{k} b_i p_i(\delta^*)\} = \rho(\delta^*). \tag{3.2.19}$$

By the previous discussion we have, for any $\vec{\theta}\in\Omega$ and for any $\delta^*\in D^*(T)$,

$$\sup_{\delta^*\in D^*(T)} E\eta(\delta^*|\vec{\theta}) = \eta(\delta^0|\vec{\theta}), \tag{3.2.20}$$

so that

$$\inf_{\delta^*\in D^*(T)} EW(\delta^*|\vec{\theta}) = EW(\delta^0|\vec{\theta}). \tag{3.2.21}$$

Hence δ^0 is the uniformly best decision rule in the class D*(T) with respect to the loss function (3.2.17). The decision rule δ^0 is the procedure R_s.

Part II

A Problem Dealing with the Selection of Subsets Where the Subset Size is a Function of the Common Sample Size

Chapter IV

Selection of Subsets of Fixed Size Depending on the Sample Size.

4.1 Introduction.

In part I we considered the problem of selecting a subset, of specified size s, from a given set of k populations. There specifying the subset size s, we have solved the problem of sample size determination so that the PCS under the procedure R_s meets certain requirement . Here we consider a related problem where the sample size is fixed in advance and the subset size s is to be determined in relation to the common sample size.

4.2 Statement of the problem.

The problem can be stated as follows: "Samples of size n* are available from each of k given populations. The experimenter is interested in choosing a subset of fixed size s (the value of s to be determined) which contains the t best of the k given populations. He desires to have a procedure which will tell him how large a subset and which subset he has to choose so that the probability of a correct selection (i.e., the subset selected by the procedure contains the t best populations) is not less than a preassigned number P*, for all parameter points in the preference zone".

4.3 Solution to the problem..

Here the preference zone of the parameter space is same as the one given by (1.2.1). We propose the procedure R_s of section 1.3, where the subset size s is suitably chosen so as to satisfy the probability requirement (1.2.3). The choice of s, the subset size, is to be made as follows: When the procedure R_s is used, let $n_1(s) = n(s|k,t,P^*,d^*)$ be the minimum sample size necessary to achieve Goal 1. From a tabular solution to the sample size determination problem in relation to Goal 1, we know values of $n_1(s)$ for given values of k,t,d*,P* and various s values. From such a table we can find an integer s (with $t \leqq s \leqq k$) such that the given sample size n* satisfies the relation

(4.3.1) $$n_1(s-1) > n^* \geq n_1(s);$$

this value of s is the required subset size when one is using the procedure R_s.

In giving the above solution we have tacitly assumed that, for given values k,t,d* and P*

(4.3.2) $$n_1(s-1) > n_1(s).$$

Later in this chapter we prove this result.

The above problem may arise in several ways of which one is the situation considered in the following section.

4.4 Relaxation of the goal of choosing the t best of k given populations.

Suppose an experimenter is interested in the goal of selecting the t best of k given populations (those with the largest parameter values) in the framework of Bechhofer (1954) with fixed t,k,d*,P* and a distance measure. It may happen either due to economic reasons or because the observations were already taken, that the procedure R_s with s = t requires too many observations for him. Then he may be willing to relax his goal. In other words he may be willing to change his goal to another goal, which he can achieve with lesser sample size. It is possible to think of several different ways of relaxing the goal.

One way is to relax the goal to that of Goal 1, namely, to choose a subset of size s which includes the t best. Then the above formulation of the problem (where $s \geq t$ is to be determined as a function of n^*, t, k, d^* and P^*) is appropriate; here n* is to be interpreted as the largest sample size that the experimenter can obtain or the size of the sample he has already obtained.

As pointed out by Sobel (see the footnote on page 22 of Bechhofer 1954) the experimenter may choose a second method of relaxing his goal, namely to choose a subset of size s and assert that they form a subset of the t best populations (i.e., Goal 2). Then a formulation with $s < t$, similar to the above

one is appropriate. Here again we determine s as a function of n*,t,k,d* and P*; n* has the same interpretation as before.

A third method of relaxation is to choose Goal I viz., selecting a subset of size s and asserting that the selected subset includes at least c of the t best. Here we have to find a pair (c,s) such that $n(c,s) \leq n^*$, $n(c,s-1) > n^*$ and $n(c+1,s) > n^*$; n^* has the same interpretation as before and $n(c,s)$ is the minimum sample size necessary to achieve Goal I under R_s.

In all these situations a particular form of the above formulation of the problem is appropriate. Which method of relaxation is the best among the several possibilities depends on various factors and this problem is not treated here.

In the above discussion we have implicitly assumed that each one of the sample sizes necessary to achieve Goal I, Goal 1 and Goal 2 is smaller than the sample size necessary to achieve the goal of selecting the t best. We now prove the relationships between the sample sizes necessary to achieve the various goals that are related to the goal of choosing the t best.

In the sequel we use the notation of section 1.9.

Theorem 4.4.1

For given k,t,P*,d* and the distance measure d(x,y),

$$(4.4.1) \qquad n(c,s-1) \geq n(c,s) \geq n(c-1,s-1) \geq n(c-1,s),$$

where $\max(1,s+1+t-k) \leq c-1$ and $c \leq \min(s-1,t)$.

Proof:

As in case of theorem 1.9.1, here also it is sufficient to prove that for fixed (but arbitrary) values of n and θ

$$(4.4.2) \qquad Q(c,s-1) \leq Q(c,s) \leq Q(c-1,s-1) \leq Q(c-1,s),$$

where

$$(4.4.3) \quad Q(c,s) = P[c^{\underline{th}} \text{ largest of } (Y_{k-t+1},\ldots,Y_k) > (s-c+1)^{\underline{st}} \text{ largest of } (Y_1,\ldots,Y_{k-t})].$$

It is easy to see that

(4.4.4) $[c^{\underline{th}}$ largest of $(Y_{k-t+1},\ldots,Y_k) > (s-c)^{\underline{th}}$ largest of $(Y_1,\ldots,Y_{k-t})]$

$\Rightarrow [c^{\underline{th}}$ largest of $(Y_{k-t+1},\ldots,Y_k) > (s-c+1)^{\underline{st}}$ largest of $(Y_1,\ldots,Y_{k-t})]$

$\Rightarrow [(c-1)^{\underline{st}}$ largest of $(Y_{k-t+1},\ldots,Y_k) > (s-c+1)^{\underline{st}}$ largest of $(Y_1,\ldots,Y_{k-t})]$

$\Rightarrow [(c-1)^{\underline{st}}$ largest of $(Y_{k-t+1},\ldots,Y_k) > (s-c+2)^{\underline{nd}}$ largest of $(Y_1,\ldots,Y_{k-t})]$.

From (4.4.3) and (4.4.4) we obtain (4.4.2). This completes the proof of the theorem.

When $c \leq \min(s-1,t)$, we have

(4.4.5) $$n(c,s-1) \geq n(c,s).$$

Setting $c = t \leq s - 1$ in (4.4.5) and noting the fact, $n(t,s) = n_1(s)$ when $s \geq t$ we obtain

(4.4.6) $$n_1(s - 1) \geq n_1(s).$$

Also when $c - 1 \geq \max(s+t+1-k,1)$, we have

(4.4.7) $$n(c,s) \geq n(c - 1,s - 1).$$

Setting $c = s \leq t$ in (4.4.7) and noting the fact that $n(s,s) = n_2(s)$ when $s \leq t$, we obtain

(4.4.8) $$n_2(s) \geq n(s-1).$$

The results (4.4.6) and (4.4.8) will now be stated as one result.

Corollary 4.4.1

For given k,t,P^*,d^* and the distance measure $d(x,y)$,

(4.4.9) $$n_1(s - 1) \geq n_1(s) \qquad \text{if } s - 1 \geq t,$$

and

$$n_2(s) \geq n_2(s - 1) \qquad \text{if } s \leq t.$$

As a particular case of the corollary we have

(4.4.10) $$n_1(t) \geq n_1(s) \quad \text{for } s > t,$$
and
$$n_1(t) = n_2(t) \geq n_2(s) \quad \text{for } s < t.$$

It may be interesting to relate $n_2(s_2)$ and $n_1(s_1)$ where $s_2 < t < s_1$. This is being done in the next theorem.

Theorem 4.4.2

For given k,t,P*,d* and the distance measure d(x,y)

(4.4.11) $$n_2(s_2) \leq n_1(s_1) \ ,$$

where $s_2 < t < s_1$.

Proof:

Again it is sufficient to show that for fixed (but arbitrary) values of n and θ

(4.4.12) $$Q_2(s_2) = Q(s_2,s_2) \geq Q_1(s_1) = Q(t,s_1).$$

Let $U(\cdot)$, $W(\cdot)$, $\overline{U}(\cdot)$ and $\overline{W}(\cdot)$ be the c.d.f.'s of the minimum of $(Y_{k-t+1},\ldots,Y_k)$, $(s_1-t+1)^{\underline{st}}$ largest of $(Y_1,\ldots,Y_{k-t})$, $s_2^{\underline{th}}$ largest of $(Y_{k-t+1},\ldots,Y_k)$ and the largest of $(Y_1,\ldots,Y_{k-t})$ respectively. Now we have

(4.4.13) $$Q(t,s_1) = \int_{-\infty}^{\infty} [1 - \overline{U}(x)]dW(x).$$

Further

(4.4.14) $$U(x) = P[\min (Y_{k-t+1},\ldots,Y_k) \leq x]$$

$$= P[\text{ at least one of } (Y_{k-t+1},\ldots,Y_k) \leq x]$$

$$\leq P[\text{ at least } (t-s_2+1) \text{ of } (Y_{k-t+1},\ldots,Y_k) \leq x], \text{ since } t > s_2$$

$$= \overline{U}(x) \ .$$

Also

(4.4.15) $W(x) = P[(s_1-t+1)^{\underline{st}}$ largest of $(Y_1,\ldots,Y_{k-t}) \leq x]$

$= P[$ at least $(k-s_1)$ of $(Y_1,\ldots,Y_{k-t}) \leq x]$

$\geq P[$ largest of $(Y_1,\ldots,Y_{k-t}) \leq x] = \overline{W}(x)$.

Thus from (4.4.13), (4.4.14) and (4.4.15), we obtain

$$Q(t,s_1) \geq \int_{-\infty}^{\infty} [1 - \overline{U}(x)]dW(x)$$

$$= \int_{-\infty}^{\infty} W(x)d\overline{U}(x) \geq \int_{-\infty}^{\infty} \overline{W}(x)d\overline{U}(x) = Q(s_2,s_2).$$

This completes the proof of the theorem.

Thus combining all the results proved so far, we have, for fixed k,t,P*,d* and the distance measure d(x,y)

(4.4.16) $n(c,s_1) \leq n(c,s_2) \leq n_2(s_2) \leq n_1(s_1) \leq n_1(t) = n_2(t)$,

where $c \leq s_2 < t < s_1$.

An example which deals with normal populations

An experimenter is interested in choosing 2 best out of 5 normal populations. He specifies that $d^* = \sigma$ and $P^* = .999$.

From the table I of Bechhofer (1954), one obtains that the experimenter needs 26 observations from each population to achieve his goal. If he can afford only 25 observations per population, he can relax his goal to that of selecting the best. (We used again table I of Bechhofer to arrive at the figure 25.) Using table 1 one can obtain that if he can only afford 16 observations per population, he can relax his goal to Goal 1 with s = 3. From the same table one obtains, that if he has only 15 observations per population he can relax his goal to Goal 2 with s = 1. From table 2 one obtains, that the experimenter needs 9 observations per population to relax his goal to Goal I with s = 2 and c = 1.

Table 1 - λ values† for Goal 1 and Goal 2

Goal 1 ($s > t$) P*	k = 3 t = 1, s = 2	k = 4 t = 1, s = 2	k = 4 t = 1, s = 3	k = 4 t = 2, s = 3	k = 5 t = 1, s = 2
.9995	3.639	3.965	3.185	3.878	4.156
.999	3.387	3.723	2.944	3.638	3.920
.995	2.738	3.102	2.321	3.022	3.313
.99	2.422	2.809	2.020	2.724	3.019
.98	2.076	2.472	1.689	2.412	2.698
.97	1.856	2.264	1.479	2.198	2.495
.96	1.690	2.107	1.323	2.045	2.343
.95	1.556	1.980	1.191	1.920	2.219
.90	1.092	1.542	0.747	1.494	1.793
.80	0.528	1.013	0.206	0.980	1.278
Goal 2 ($s < t$)	k = 3 t = 2, s = 1	k = 4 t = 3, s = 2	k = 4 t = 3, s = 1	k = 4 t = 2, s = 1	k = 4 t = 4, s = 3

Goal 1 ($s > t$) P*	k = 5 t = 1, s = 3	k = 5 t = 1, s = 4	k = 5 t = 2, s = 3	k = 5 t = 2, s = 4	k = 5 t = 3, s = 4
.9995	3.542	2.916	4.194	3.413	4.012
.999	3.295	2.673	3.963	3.184	3.777
.995	2.699	2.066	3.375	2.598	3.180
.99	2.410	1.781	3.096	2.314	2.893
.98	2.093	1.448	2.787	2.006	2.581
.97	1.892	1.243	2.594	1.812	2.384
.96	1.741	1.088	2.449	1.666	2.236
.95	1.619	0.962	2.332	1.547	2.117
.90	1.196	0.506	1.932	1.141	1.707
.80	0.685	*	1.453	0.652	1.215
Goal 2 ($s < t$)	k = 5 t = 4, s = 2	k = 5 t = 4, s = 1	k = 5 t = 3, s = 2	k = 5 t = 3, s = 1	k = 5 t = 2, s = 1

† These are the solutions of the equation (2.2.23).

* indicates the corresponding P* value is not of interest for these goals.

Table 2 - λ values† for Goal I

P*	k = 4, t = 2 s = 2, c = 1	k = 5, t = 2 s = 2, c = 1	k = 6, t = 2 s = 2, c = 1	k = 6, t = 2 s = 3, c = 1
.9995	2.85	2.97	3.00	2.75
.999	2.66	2.89	2.96	2.49
.995	1.85	2.30	2.61	1.78
.99	1.52	1.92	2.18	1.48
.98	1.27	1.63	1.86	1.26
.97	1.03	1.44	1.68	1.04
.96	.88	1.32	1.50	.90
.95	.76	1.21	1.42	.79
.90	.36	.79	1.04	.44
.80	*	.32	.59	*

† These are the solutions of the equation (2.2.9).

* indicates the corresponding P* value is not of interest for this goal.

References

Alam, K. and Rizvi, M. H. (1965). Selection from multivariate normal populations. Technical Report No. 65-1, Mathematics Department, The Ohio State University.

Bahadur, R. R. (1950). On a problem in the theory of k populations. Ann. Math. Statist., 21, 362-375.

Bahadur, R. R. and Goodman, L. A. (1952). Impartial decision rules and sufficient statistics. Ann. Math. Statist., 23, 553-562.

Bahadur, R. R. and Robbins, H. (1950). The problem of the greater mean. Ann. Math. Statist., 21, 469-487.

Barr, D. R. and Rizvi, M. H. (1964). Ranking and selection problems of uniform distributions. Ann. Math. Statist., 35, 1842, abstract #16.

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.

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.

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.

Birnbaum, Z. W. and Saunders, S. C. (1958). A statistical model for life lengths of materials. J. Amer. Statist. Assoc., 53, 151-160.

Cramér, H. (1946). Mathematical Methods of Statistics. Princeton University Press, Princeton.

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

Gupta, S. S. (1963). On a selection and ranking procedure for gamma populations. Ann. Inst. Statist. Math., 14, 199-216.

Gupta, S. S. and Sobel, M. (1957). On a statistic which arises in selection and ranking problems. Ann. Math. Statist., 28, 957-967.

Gupta, S. S. and Sobel, M. (1958). On selecting a subset which contains all populations better than a standard. Ann. Math. Statist., 29, 235-244.

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

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

Koopmans, B. O. (1936). On distributions admitting a sufficient statistic. Trans. Amer. Math. Soc., 39, 399-409.

Lehmann, E. L. (1959). Testing Statistical Hypotheses. John Wiley and Sons, New York.

Lehmann, E. L. (1961). Some model I problems of selection. Ann. Math. Statist., 32, 990-1012.

Mahamunulu, D. M. (1964). Note on ranking with three populations (preliminary report). Ann. Math. Statist., 35, 1851, abstract #9.

Mahamunulu, D. M. (1965). A class of ranking and selection procedures (preliminary report). Ann. Math. Statist., 36, 728, abstract #7.

Milton, R. C. (1965). Exact properties of rank order procedures under normal shift alternatives. Ph.D. Thesis, University of Minnesota.

Mosteller, F. (1948). A k-sample slippage test for an extreme population. Ann. Math. Statist., 19, 58-65.

Paulson, E. (1949). A multiple decision procedure for certain problems in analysis of variance. Ann. Math. Statist., 20, 95-98.

Paulson, E. (1952). An optimum solution to the k-sample slippage problem for the normal distribution. Ann. Math. Statist., 23, 610-616.

Rao, C. R. (1952). Advanced Statistical Methods in Biometric Research. John Wiley and Sons, New York.

Rizvi, M. H. (1963). Ranking and selection problems of normal populations using the absolute values of their means: fixed sample size case. Technical Report No. 31, Department of Statistics, Univ. of Minnesota.

Seal, K. C. (1955). On a class of decision procedures for ranking means of normal populations. Ann. Math. Statist., 26, 387-398.

Seal, K. C. (1958). On ranking parameters of scale in type III populations. J. Amer. Statist. Assoc., 53, 164-175.

Sobel, M. (1963). Single sample ranking problems with Poisson populations. Technical Report No. 19, Department of Statistics, Univ. of Minnesota.

Teichroew, D. (1955). Probabilities Associated with Order Statistics in Samples from Two Normal Populations with Equal Variances. Chemical Corps Engineering Agency, Army Chemical Center, Maryland.

Acknowledgements

The author is deeply indebted to Professor Milton Sobel for introducing him to the field of ranking and selection problems and providing invaluable suggestions and encouragement during the course of the research. Great appreciation is recorded for the help given by Miss Kathy Smith, Miss Susan Joern and Mrs. Patricia Carlson in the typing.