ON SELECTING A SUBSET OF k POPULATIONS CONTAINING THE BEST[1]

By W. J. Studden

Purdue University

1. Introduction. This paper is concerned with the problem of selecting a subset of k populations which is in some sense an optimal subset. In the usual subset selection type setup we are given k populations $\pi_1, \pi_2, \cdots, \pi_k$ with densities $f_{\theta_1}, f_{\theta_2}, \cdots, f_{\theta_k}$. In general the parameters θ_i are not known and usually range over some subset of the real line. For convenience it is assumed that the larger the parameter θ, the more preferable is the selection of the corresponding population. The population with the largest parameter is called the best and a selection of any subset containing the best population is called a correct selection. If the selection proceeds according to some rule R, then the subset selected is required to contain the best population with a specified probability γ.

For many types of densities $f_\theta(x)$, e.g. normal, gamma, binomial and more general situations, different types of rules have been proposed and numerous properties of the rules have been investigated. References to some of the literature on the subject can be found in Gupta (1965).

In this paper we first investigate the more elementary problem of defining "optimal" subset selection rules for the case where we have k *fixed* density functions and only the correct pairing of the densities and populations is unknown.

Section 2 contains a decision theoretic formulation of the selection problem and a solution is obtained under the usual symmetry conditions. In Section 3 we examine the solution for the exponential family when the parameters are in a "slippage" type configuration. In Section 4 we obtain a result for the normal case where the parameters are permitted to vary. Section 5 contains two remarks; one concerning scale invariant procedures, the other concerning the fact that certain "classical" procedures appear as limits of the procedures obtained in Section 3 for the exponential family.

The results presented below are all quite elementary, however, it is hoped that the selection procedures introduced (especially those in Theorem 3.1) will have some interest as alternatives to those procedures now in use. Theorem 4.1 indicates that the usual procedures are not the "best" for certain specific situations.

The author wishes to thank the referee and Professor R. A. Wijsman for numerous suggestions and criticisms.

Received 19 October 1966; revised 15 February 1967.

[1] This research was supported in part by the Office of Naval Research Contract NONR 1100 (26) and the Aerospace Research Laboratories Contract AF 33 (615) 67C1244 at Purdue University. Reproduction in whole or in part is permitted for any purposes of the United States Government.

2. Formulation of the problem. We formulate our problem as follows. We are given k *fixed* values $\theta_{[1]} \leqq \theta_{[2]} \leqq \cdots \leqq \theta_{[k]}$ and k populations $\pi_1, \pi_2, \cdots, \pi_k$. The correct pairing of the populations and the parameters is unknown. We are given one observation x_i from each population π_i, $i = 1, 2, \cdots, k$. The vector of observations $\mathbf{x} = (x_1, x_2, \cdots, x_k)$ is assumed to have density $f(\mathbf{x}; \boldsymbol{\theta})$ (with respect to some measure μ) where $\boldsymbol{\theta} = (\theta_1, \theta_2, \cdots, \theta_k)$ belongs to the set Ω of k-vectors consisting of permutations of the k fixed values $\theta_{[1]}, \theta_{[2]}, \cdots, \theta_{[k]}$.

Let A denote the action space consisting of the 2^k subsets of the set $\{1, 2, \cdots, k\}$. A measurable function δ defined on $X \times A$ is called a *selection procedure* provided that for each fixed k-vector of observations $x \varepsilon X$, $\delta(\mathbf{x}, a) \geqq 0$ and $\sum_A \delta(\mathbf{x}; a) = 1$. Thus if $\mathbf{x} \varepsilon X$ is an observed vector of observations, $\delta(\mathbf{x}, a)$ is the probability of selecting the subset $a \varepsilon A$. Let $\varphi_i(x) = \sum_{a \ni i} \delta(x; a)$ (summation over those subsets a containing i) denote the probability of selecting the ith population. The functions $\varphi_1, \varphi_2, \cdots, \varphi_k$ will be referred to as the *individual selection probabilities*. Note that the selection procedure δ is completely specified by the individual selection probabilities whenever the latter take on only the values zero and one.

The population associated with the largest parameter $\theta_{[k]}$ is called the "best" and a selection of a subset containing the best population is called a *correct selection* (CS). In the case where two or more of the largest parameter values are equal one of these parameters or the corresponding populations is "tagged" and called the best population.

We shall assume that a selection of the subset $a \varepsilon A$ results in a loss $L(\boldsymbol{\theta}, a) = \sum_{i \varepsilon a} L_i(\boldsymbol{\theta})$ where $L_i(\boldsymbol{\theta})$ is the loss whenever the ith population is selected. An additional loss of L will be imposed if a correct selection is not made. It is readily seen that $\sum_A L(\boldsymbol{\theta}, a) E_{\boldsymbol{\theta}}\, \delta(\mathbf{x}, a) = \sum_{i=1}^k L_i(\boldsymbol{\theta}) E_{\boldsymbol{\theta}} \varphi_i$. Our problem will be to minimize the risk

$$(2.1) \qquad R(\boldsymbol{\theta}, \boldsymbol{\varphi}) = \sum_{i=1}^k L_i(\boldsymbol{\theta}) E_{\boldsymbol{\theta}} \varphi_i + L[1 - P_{\boldsymbol{\theta}}(CS \mid \boldsymbol{\varphi})].$$

This minimization will be done under the usual symmetry conditions imposed by the group G of permutations $g(1, \cdots, k) \to (g1, \cdots, gk)$. If $h = g^{-1}$ denotes the inverse of g let $g\mathbf{x} = g \cdot (x_1, \cdots, x_k) = (x_{h1}, \cdots, x_{hk})$ and $g\boldsymbol{\theta} = (\theta_{h1}, \cdots, \theta_{hk})$. We assume that all quantities involved are invariant under G, i.e. the measure μ is invariant, $f(\mathbf{x}; \boldsymbol{\theta}) = f(g\mathbf{x}, g\boldsymbol{\theta})$, $L_i(\boldsymbol{\theta}) = L_{gi}(g\boldsymbol{\theta})$ and $\delta(\mathbf{x}, a) = \delta(g\mathbf{x}, ga)$ where $ga = \{gi \mid i \varepsilon a\}$. The invariance of the selection rule δ implies immediately that the individual selection probabilities are also invariant, i.e. that $\varphi_i(\mathbf{x}) = \varphi_{gi}(g\mathbf{x})$, $i = 1, 2, \cdots, k$. An invariant selection procedure δ or a set of invariant individual selection probabilities which minimizes (2.1) will be called *best invariant*.

Since G acts transitively on Ω it is easily seen that for invariant δ both $R_1(\boldsymbol{\theta}, \boldsymbol{\varphi}) = \sum_{i=1}^k L_i(\boldsymbol{\theta}) E_{\boldsymbol{\theta}} \varphi_i$ and $P_{\boldsymbol{\theta}}(CS \mid \boldsymbol{\varphi})$ are invariant under G, i.e. they are independent of $\boldsymbol{\theta} \varepsilon \Omega$. The class of invariant loss functions includes the following cases.

(i) $L_i(\boldsymbol{\theta}) \equiv 1$. In this situation $R_1(\boldsymbol{\theta}; \boldsymbol{\varphi}) = \sum_{i=1}^k E_{\boldsymbol{\theta}}\varphi_i$ is the expected size of the selected subset.

(ii)

$$L_i(\boldsymbol{\theta}) = 1 \text{ if } \theta_i \neq \theta_{[k]} \qquad i = 1, 2, \cdots, k$$
$$= 0 \text{ if } \theta_i = \theta_{[k]} .$$

In this case $R_1(\boldsymbol{\theta}; \boldsymbol{\varphi})$ is the expected subset size excluding the best population.

It is easily seen that the function $\boldsymbol{\varphi}$ defined in (2.2) below is the same for cases (i) and (ii). In the examples considered we restrict ourselves to case (i). It should be noted however that when the parameters θ_i are permitted to vary, it would probably be more appropriate to use case (ii) or (iii).

(iii) $L_i(\boldsymbol{\theta}) =$ rank of θ_i in the sequence $\theta_{[k]}, \theta_{[k-1]}, \cdots, \theta_{[1]}$. Here we assume that $\theta_{[k]}$ has rank one and $\theta_{[1]}$ has rank k, etc. Then $R_1(\boldsymbol{\theta}; \boldsymbol{\varphi})$ is the expected sum of ranks of the populations in the selected subset.

Let Ω_k denote those permutations of $(\theta_1, \theta_2, \cdots, \theta_k)$ such that the largest parameter value $\theta_{[k]}$ (or the tagged parameter) is in the last component. In addition let

$$p_i(\mathbf{x}; \boldsymbol{\theta}) = (1/(k-1)!) \textstyle\sum_{G_i} f(\mathbf{x}; g\boldsymbol{\theta}) \qquad i = 1, 2, \cdots, k$$

where $G_i = \{g \mid g^{-1}k = i\}$. We then have the following theorem.

THEOREM 2.1. *A selection rule δ is best invariant if and only if*

$$\varphi_k(\mathbf{x}) = 1 \text{ if } Lp_k(\mathbf{x}; \boldsymbol{\theta}) > \textstyle\sum_{i=1}^k L_i(\boldsymbol{\theta}) p_i(\mathbf{x}; \boldsymbol{\theta}) \tag{2.2}$$
$$= 0 \text{ if } Lp_k(\mathbf{x}; \boldsymbol{\theta}) < \textstyle\sum_{i=1}^k L_i(\boldsymbol{\theta}) p_i(\mathbf{x}; \boldsymbol{\theta})$$

for $\boldsymbol{\theta} \varepsilon \Omega_k$, almost everywhere μ. The functions φ_i, $i \neq k$ are defined by the invariant conditions on $\boldsymbol{\varphi}$. (Note that the quantities defining φ_k are independent of $\boldsymbol{\theta} \varepsilon \Omega_k$ so that φ_k and hence $\boldsymbol{\varphi}$ are well defined on the set where $Lp_k(\mathbf{x}; \boldsymbol{\theta}) \neq \sum_{i=1}^k L_i(\boldsymbol{\theta}) p_i(\mathbf{x}; \boldsymbol{\theta})$).

PROOF. Since our problem is invariant under G and G acts transitively on Ω there is at least one best invariant procedure and any best invariant procedure is Bayes with respect to the uniform distribution on Ω. The result now follows by noting that for $\boldsymbol{\theta} \varepsilon \Omega_k$

$$\textstyle\sum_G R(g\boldsymbol{\theta}, \varphi) - L = \sum_{j=1}^k \int \varphi_j(\mathbf{x}) [\sum_{i=1}^k L_i(\boldsymbol{\theta}) \sum_{\{g^{-1}j=i\}} f(x; g\boldsymbol{\theta})$$
$$- L \textstyle\sum_{\{g^{-1}j=k\}} f(x; g\boldsymbol{\theta})] \, d\mu.$$

Since $P_{\boldsymbol{\theta}}(CS \mid \boldsymbol{\varphi})$ is independent of $\boldsymbol{\theta} \varepsilon \Omega$ for invariant selection rules we may define $\gamma(L) = P_{\boldsymbol{\theta}}(CS \mid \boldsymbol{\varphi})$ where $\boldsymbol{\varphi}$ satisfies Equation (2.2) for a given value of L. The following corollary is immediate.

COROLLARY 2.1. *An invariant selection procedure minimizes $\sum_{i=1}^k L_i(\boldsymbol{\theta}) E_{\boldsymbol{\theta}} \varphi_i$ subject to the condition*

$$P_{\boldsymbol{\theta}}(CS \mid \boldsymbol{\varphi}) \geqq \gamma(L) \qquad \text{for all } \boldsymbol{\theta} \varepsilon \Omega \tag{2.3}$$

if and only if the individual selection probabilities are determined by (2.2).

REMARK 2.1.

(a) The function $\gamma(L)$ is nondecreasing in L so that for a fixed value of $\gamma \varepsilon (0, 1)$ the minimum of $\sum_{i=1}^{k} L_i(\boldsymbol{\theta})E_{\boldsymbol{\theta}}\varphi_i$ subject to $P_{\boldsymbol{\theta}}(CS \mid \boldsymbol{\varphi}) \geqq \gamma$ will be attained by the $\boldsymbol{\varphi}$ in (2.2) where L is the minimal value of L for which

$$P_{\boldsymbol{\theta}}(CS \mid \boldsymbol{\varphi}) \geqq \gamma.$$

(b) If two fixed k-vectors $\boldsymbol{\theta}^0$ and $\boldsymbol{\theta}$ are considered and we replace $P_{\boldsymbol{\theta}}(CS \mid \boldsymbol{\varphi})$ by $P_{\boldsymbol{\theta}^0}(CS \mid \boldsymbol{\varphi})$ in (2.1) and (2.3) then Theorem 2.1 and its Corollary remain valid when $p_k(\mathbf{x}; \boldsymbol{\theta})$ is replaced by $p_k(\mathbf{x}; \boldsymbol{\theta}^0)$ in (2.2).

3. Exponential family. The expression given in Equation (2.2) defining the selection probabilities which minimize the risk $R(\boldsymbol{\theta}; \boldsymbol{\varphi})$ is rather complicated when written down explicitly in terms of the original densities. However for the "slippage" situation when the underlying densities are from an exponential family and $L_i(\theta) \equiv 1$ the expressions simplify considerably.

We suppose that $f_\theta(x)$ is a density of the form

$$f_\theta(x) = C(\theta)e^{\theta x} \tag{3.1}$$

with respect to some σ-finite measure μ on the real line. Let

$$f_{\boldsymbol{\theta}}(\mathbf{x}) = \prod_{i=1}^{k} f_{\theta_i}(x_i) \tag{3.2}$$

and suppose further that $\theta = \theta_{[1]} = \theta_{[2]} = \cdots = \theta_{[k-1]} = \theta_{[k]} - \Delta$ for some fixed $\Delta > 0$. It then follows that for $\boldsymbol{\theta} \varepsilon \Omega_k$

$$p_k(\mathbf{x}; \boldsymbol{\theta}) = \prod_{i=1}^{k} C(\theta_i) \exp \{\theta \textstyle\sum_i^k x_j + \Delta x_k\} .$$

If $\sum'$ denotes a summation over all permutations of $x_1, x_2, \cdots, x_{k-1}$ then

$$\begin{aligned} \textstyle\sum_{i=1}^{k-1} p_i(\mathbf{x}; \boldsymbol{\theta}) &= (1/(k-1)!) \textstyle\sum_{i=1}^{k-1} \sum' \prod_{j=1}^{k} C(\theta_j) \exp \{\theta \sum x_j + \Delta x_i\} \\ &= (\textstyle\prod_1^k C(\theta_j)/(k-1)!) [\exp \theta \sum x_j] \sum' \sum_{i=1}^{k-1} e^{\Delta x_i} \\ &= \textstyle\prod_1^k C(\theta_j) [\exp \theta \sum_1^k x_j] \sum_{i=1}^{k-1} e^{\Delta x_i}. \end{aligned}$$

With the aid of Corollary 2.1 the above analysis proves the following theorem.

THEOREM 3.1. *Let* $f_{\boldsymbol{\theta}}(\mathbf{x}) = \prod_{i=1}^{k} f_{\theta_i}(x_i)$ *where* $f_\theta(x) = C(\theta)e^{\theta x}$ *and* $\theta = \theta_{[1]} = \theta_{[2]} = \cdots = \theta_{[k-1]} = \theta_{[k]} - \Delta (\Delta > 0)$. *Let S denote the size of the selected subset. An invariant selection rule δ minimizes $E_{\boldsymbol{\theta}}S$ subject to the condition* $P_{\boldsymbol{\theta}}(CS \mid \boldsymbol{\varphi}) \geqq \gamma$ *if and only if for almost all* $\mathbf{x}$

$$\begin{aligned} \varphi_k(\mathbf{x}) &= 1 \text{ if } \textstyle\sum_{i=1}^{k-1} e^{\Delta x_i} < Ce^{\Delta x_k} \\ &= 0 \text{ if } \textstyle\sum_{i=1}^{k-1} e^{\Delta x_i} > Ce^{\Delta x_k}. \end{aligned} \tag{3.3}$$

REMARK 3.1. If $f_\theta(x)$ possesses the slightly more general form $f_\theta(x) = C(\theta)e^{Q(\theta)T(x)}$ and $\theta = \theta_{[1]} = \cdots = \theta_{[k-1]} = \theta_{[k]} - \Delta$ then the expression in (3.3) defining $\varphi_k(\mathbf{x})$ is obtained by replacing x_i by $T(x_i)$ and Δ by $Q(\theta + \Delta) - Q(\theta)$.

It seems worthwhile to point out a number of properties of the function $\boldsymbol{\varphi}$

defined through (3.3) and the region

$$A_k = \{\mathbf{x} \mid \textstyle\sum_{i=1}^{k-1} e^{\Delta x_i} < Ce^{\Delta x_k}\} \tag{3.4}$$

for which we select the kth population with probability one. These properties are listed in the following lemma.

LEMMA 3.1. *The region* A_k *defined in* (3.4) is:

(i) *convex,*

(ii) *translation invariant,* i.e. *if* $(x_1, \cdots, x_k) \varepsilon A_k$ *then* $(x_1 + b, \cdots, x_k + b) \varepsilon A_k$ for all b,

(iii) *symmetric in* $x_1, \cdots, x_{k-1}$,

(iv) *if* $(x_1, \cdots, x_k) \varepsilon A_k$ *then for any* $b > 0$ $(x_1, \cdots, x_{k-1}, x_k + b) \varepsilon A_k$ *and* $(x_1, \cdots, x_i - b, \cdots, x_{k-1}, x_k) \varepsilon A_k$ *for* $i = 1, 2, \cdots, k - 1$.

(v) *if the constant* C *in* (3.3) *is* $>k - 1$ *then* $\sum_{i=1}^k \varphi_i(\mathbf{x}) \geqq 1$ *for all* $\mathbf{x}$, i.e. *we select at least one population with probability one; if* $C < k - 1$ *then* $\sum_{i=1}^k \varphi_i(\mathbf{x}) < 1$ *at least for* $\mathbf{x}$ *in a neighborhood of the equiangular line.*

PROOF. The proofs of the first four parts are immediate; part (iii) in fact is simply a restatement of the invariance imposed on $\boldsymbol{\varphi}$. To prove part (iv) we define

$$B_k = \{\mathbf{x} \mid x_k \geqq \max_i \{x_i\}\}$$

and let A_i and B_i, $i = 1, 2, \cdots, k - 1$, denote the corresponding sets obtained by interchanging x_i and x_k in A_k and B_k respectively. Since $B_i \subset A_i$, $i = 1, 2, \cdots, k$ and $\bigcup_{i=1}^k B_i = E_k =$ euclidean k-space, the result follows. The converse statement of part (v) is nearly immediate.

REMARK 3.2. With regard to part (v) in the above lemma it should be noted that situations where $\sum_{i=1}^k \varphi_i(\mathbf{x}) < 1$ for some $\mathbf{x}$ do exist. For example consider the situation of Theorem 3.1 for $k = 2$ where $f_\theta(x) = f(x - \theta)$ and $f(x)$ is the standard normal. In this case the second population is selected if $x_2 > x_1 - \Delta^{-1} \log C$. If γ is sufficiently close to $\frac{1}{2}$ and $\theta_2 > \theta_1$ then $C < 1$. In this situation $\varphi_1(\mathbf{x}) + \varphi_2(\mathbf{x}) = 0$ for $|x_1 - x_2| < \Delta^{-1} \log C^{-1}$.

4. Normal density.

In this section we consider a simple situation concerning normal populations where the parameters are permitted to vary.

The following lemma will be used. The proof will be omitted since the result follows readily from Lemma 2, p. 74 of Lehman (1959).

LEMMA 4.1. *If* $\varphi(x_1, \cdots, x_n)$ *is nonincreasing or nondecreasing in each of the variables* x_i *separately and* $f(\mathbf{x}; \boldsymbol{\theta}) = \prod_{i=1}^n f_{\theta_i}(x_i)$ *where* $f_\theta(x)$ *has a monotone likelihood ratio in* x *then* $E_{\boldsymbol{\theta}}\varphi = \psi(\theta_1, \cdots, \theta_n)$ *is nonincreasing or nondecreasing in each* θ_i *separately and in the same direction as the corresponding* x_i.

We assume that $f(\mathbf{x}; \boldsymbol{\theta}) = \prod_{i=1}^k f(x_i - \theta_i)$ where $f(x)$ is the standard normal density. For fixed Δ let $\boldsymbol{\varphi}(\mathbf{x}; \Delta)$ denote the selection probabilities defined by (3.3) where C is chosen so that $P_{\boldsymbol{\theta}}(CS \mid \boldsymbol{\varphi}(\mathbf{x}; \Delta)) = \gamma$ for all $\boldsymbol{\theta} = (\theta, \cdots, \theta, \theta + \Delta)$. Let $\Phi(\Delta)$ denote the class of invariant procedures satisfying

$$P_{\boldsymbol{\theta}}(CS \mid \boldsymbol{\varphi}) \geqq \gamma \quad \text{for all} \quad \boldsymbol{\theta} \,\varepsilon\, \Omega(\Delta)$$

where

$$\Omega(\Delta) = \{\boldsymbol{\theta} \mid \theta_{[1]} \leqq \theta_{[2]} \leqq \cdots \leqq \theta_{[k-1]} \leqq \theta_{[k]} - \Delta\}.$$

THEOREM 4.1. *For any* $\boldsymbol{\theta}$ *with* $\theta_{[1]} = \theta_{[2]} = \cdots = \theta_{[k-1]} = \theta_{[k]} - \Delta$ *the minimum value of the expected subset size* $E_{\boldsymbol{\theta}}(S \mid \varphi)$ *over the class* $\Phi(\Delta)$ *is attained by* $\varphi(\mathbf{x}; \Delta)$; i.e.

$$\min_{\Phi(\Delta)} E_{\boldsymbol{\theta}}(S \mid \varphi) = E_{\boldsymbol{\theta}}(S \mid \varphi(\mathbf{x}; \Delta)). \tag{4.1}$$

PROOF. For a fixed k-vector $\boldsymbol{\theta}$ with $\theta_{[1]} = \cdots = \theta_{[k-1]} = \theta_{[k]} - \Delta$ Corollary 2.1 asserts that $\varphi(\mathbf{x}; \Delta)$ minimizes $E_{\boldsymbol{\theta}}(S \mid \varphi)$ over those procedures satisfying $P_{\boldsymbol{\theta}}(CS \mid \varphi) \geqq \gamma$. By Lemma 4.1 the procedure with selection probabilities $\varphi(\mathbf{x}; \Delta)$ is contained in $\Phi(\Delta)$ so that $\varphi(\mathbf{x}; \Delta)$ also minimizes $E_{\boldsymbol{\theta}}(S \mid \varphi)$ over $\Phi(\Delta)$.

REMARK 4.1. For the case $k = 2$ the selection probability $\varphi_2(\mathbf{x}; \Delta)$ reduces to

$$\begin{aligned} \varphi_2(\mathbf{x}; \Delta) &= 1 \text{ if } x_2 > x_1 - c \\ &= 0 \text{ if } x_2 < x_1 - c. \end{aligned}$$

For this case it can be shown that $\varphi(\mathbf{x}; \Delta)$ possesses a stronger property, namely that Equation (4.1) is valid not only for those $\boldsymbol{\theta}$ with $\theta_{[1]} = \theta_{[2]} - \Delta$ but for all $\boldsymbol{\theta}$ for which $\theta_{[1]} \geqq \theta_{[2]} - \Delta$ or equivalently $|\theta_1 - \theta_2| \leqq \Delta$. This result can be verified with a judicious application of Remark 2.1 (b).

REMARK 4.2. The result proven in Theorem 4.1 holds in general for translation invariant exponential families since these were the only properties of the normal density that were used. However, if the density $f_\theta(x)$ is given by $f(x - \theta) = C(\theta)e^{\theta x}h(x)$ then $f_\theta(x)$ can be shown to be normal provided $h(x)$ is continuous.

5. Additional Remarks

(A) The procedures defined in Theorem 3.1 are invariant under translation of each component by a constant. This property would seem natural if the underlying densities are translation invariant but not if they are scale invariant. The results in Theorems 2.1 and 3.1 can be modified in the usual manner to produce scale invariant procedures. For example let $f(\mathbf{x}; \boldsymbol{\theta}) = \prod_{i=1}^{k} f_{\theta_i}(x_i)$ where $f_\theta(x) = (1/\theta)f(x/\theta)$, $\theta > 0$ and $f(x) = (x^{p-1}e^{-x}/\Gamma(p))$, $x > 0$, $p > 0$ and suppose $\theta = \theta_{[1]} = \theta_{[2]} = \cdots = \theta_{[k-1]}$ and $\theta_{[k]} = r\theta\,(r > 1$ and fixed). Then among all procedures which are invariant under both permutations and scale change and satisfy $P_{\boldsymbol{\theta}}(CS \mid \varphi) \geqq \gamma$ for all θ the procedure minimizing $E_{\boldsymbol{\theta}}(S \mid \varphi)$ is defined by setting $\varphi_k(x_1, \cdots, x_k) = 1$ if and only if

$$\sum_{i=1}^{k-1} [(x_k + r\sum_{j=1}^{k-1} x_j)/(x_i + r\sum_{j=1, j\neq i}^{k} x_j)]^{kp} < C.$$

(B) Consider the sequence of selection probabilities defined for $\Delta \varepsilon (0, \infty)$ by

$$\begin{aligned} \varphi_k(\mathbf{x}; \Delta) &= 1 \text{ if } \sum_{i=1}^{k-1} e^{\Delta x_i} < C(\Delta)e^{\Delta x_k} \\ &= 0 \text{ if } \sum_{i=1}^{k-1} e^{\Delta x_i} > C(\Delta)e^{\Delta x_k}. \end{aligned} \tag{5.1}$$

For $\Delta = 0$ we let

$$(5.2)\qquad \begin{aligned}\varphi_k(\mathbf{x}; 0) &= 1 \text{ if } \textstyle\sum_{j=1}^{k-1} x_j/k - 1 < x_k + C(0)\\ &= 0 \text{ if } \textstyle\sum_{j=1}^{k-1} x_j/k - 1 > x_k + C(0),\end{aligned}$$

while for $\Delta = \infty$ we define

$$(5.3)\qquad \begin{aligned}\varphi_k(\mathbf{x}; \infty) &= 1 \text{ if } \max_{1 \leqq j \leqq k-1} x_j < x_k + C(\infty)\\ &= 0 \text{ if } \max_{1 \leqq j \leqq k-1} x_j > x_k + C(\infty).\end{aligned}$$

The values $C(\Delta)$, $\Delta \varepsilon [0, \infty]$ are all chosen so that for a *fixed* set of values $\theta_{[1]} \leqq \cdots \leqq \theta_{[k]}$ the probability of a correct selection is equal to a given value γ. It should be emphasized that the present choice of $C(\Delta)$ is different from that used elsewhere in the paper.

The procedures defined in (5.1) and (5.2) are presently in use and have been considered and investigated by numerous authors; see for example Gupta (1959) and Seal (1955).

It can readily be shown as may be anticipated by the notation that the functions $\varphi_k(\mathbf{x}; \Delta)$ have limits $\varphi_k(\mathbf{x}; 0)$ and $\varphi_k(\mathbf{x}; \infty)$ almost everywhere μ as Δ approaches zero and infinity respectively.

REFERENCES

GUPTA, SHANTI S. (1965). On some multiple decision (selection and ranking) rules. Technometrics, **7** 225–245.

LEHMAN, E. L. (1959). Testing statistical hypothesis. John Wiley and Sons, New York.

LEHMAN, E. L. (1961). Some model I problems of selection. *Ann. Math. Statist.* **32** 990–1011.

SEAL, K. C. (1955). On a class of decision procedures for ranking means of normal populations. *Ann. Math. Statist.* **36** 387–397.