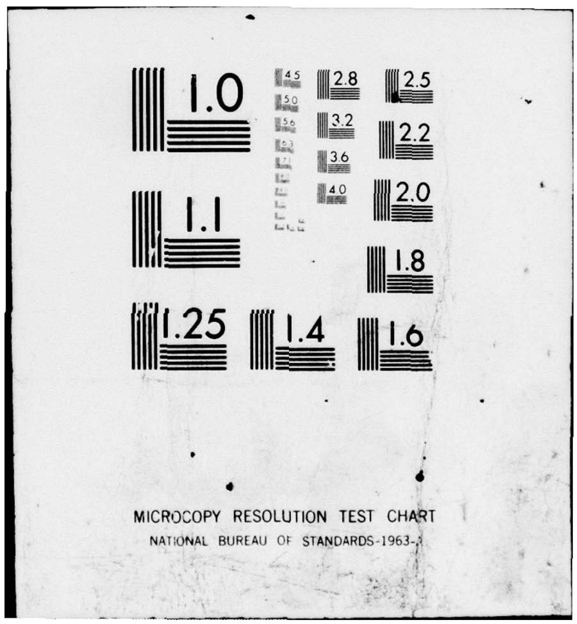
AD-A080 582 FLORIDA STATE UNIV TALLAHASSEE DEPT OF STATISTICS F/G 12/1
A MINIMAX AND ADMISSIBLE SUBSET SELECTION RULE FOR THE LEAST PR--ETC(U)
SEP 79 R L BERGER DAAG29-79-C-0158
UNCLASSIFIED FSU-STATISTICS-M522 ARO-16713.4-M NL

| OF |

AD
A080582

END
DATE
FILMED
2-80
DDC


1.0
1.1
1.25
1.4
1.6
1.8
2.0
2.2
2.5
2.8
3.2
3.6
4.0
4.5
5.0
5.6
MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

(12)

(6)

A MINIMAX AND ADMISSIBLE SUBSET SELECTION RULE
FOR THE LEAST PROBABLE MULTINOMIAL CELL[1].

by

(10) Roger L. Berger
The Florida State University

(14) FSU-Statistics Report-M522, TR-
USARO Technical Report D39

Abbreviated title: Minimax and Admissible Subset Selection

(9) Technical rept.,

The Florida State University
Department of Statistics
Tallahassee, Florida 32306

(18) ARO

(11) September, 1979

(12) 16

(19) 16713.4-M

[1]This research was supported by the U.S. Army Research Office Grant No.
(15) DAAG29-79-C-0158

AMS 1970 Subject Classifications: Primary 62F07; Secondary 62C05.

Key words and phrases: Minimax, admissible, subset selection, multinomial, Schur function, stochastic majorization.

400 277

YB

ABSTRACT

In this paper a new subset selection rule for selecting a subset containing the least probable multinomial cell is defined. The rule is shown to be minimax and admissible in the class of rules which have a preassigned probability of at least P^* of selecting the least probable cell provided that P^* is sufficiently large. The loss used is the number of non-best cells selected.

A Minimax and Admissible Subset Selection Rule for the Least Probable Multinomial Cell

1. INTRODUCTION. In this paper, subset selection problems for the multinomial distribution are considered. In these problems, the aim is to select a non-empty subset of the cells which contains the cell with the lowest cell probability. Having restricted attention to rules which have a high probability of including the least probable cell, the goal is to find a rule which effectively excludes the cells associated with the larger cell probabilities. This leads to the use of the number of non-best cells selected as a measure of the loss to the experimenter. In this paper, a subset selection rule is presented which is minimax and admissible for this problem. The rule is simple and easy to implement and in some cases is similar to a rule proposed and studied by Nagel (1970).

Alam and Thompson (1972) considered the problem of selecting the single least probable cell. The subset selection problem for the multinomial distribution has been previously considered by Gupta and Nagel (1967), Nagel (1970), Panchapakesan (1971) and Gupta and Huang (1975). Berger (1979b) described a class of minimax multinomial selection rules. Minimax selection rules for multinomial and other distributions have been considered by Berger (1979a) and Bjornstad (1978). Berger and Gupta (1979) found minimax and admissible subset selection rules for location parameters but the class of selection rules considered was restricted. To this author's knowledge, this is the first time minimax and admissible subset selection rules have been derived for the multinomial or any other problem.

Section 2 contains the necessary notation for a formulation of the problem. The selection rule is defined in Section 3. The minimaxity and admissibility of the rule is proven in Section 4.

2. NOTATION AND FORMULATION. Let $\underset{\sim}{X} = (X_1, \ldots, X_k)$ be a multinomial random vector with $\sum_{i=1}^k X_i = n$. $\underset{\sim}{x}$ and $\underset{\sim}{y}$ will denote vectors in the sample space of $\underset{\sim}{X}$. Let $\underset{\sim}{p} = (p_1, \ldots, p_k)$ be the unknown cell probabilities with $\sum_{i=1}^k p_i = 1$. The ordered cell probabilities will be denoted by $p_{[1]} \leq \cdots \leq p_{[k]}$. The goal of the experimenter is to select a subset of the cells including the best cell, the cell associated with $p_{[1]}$. A correct selection, CS, is the selection of any subset which contains the best cell.

The action space A for a subset selection problem is the $2^k - 1$ non-empty subsets of $\{1, 2, \ldots, k\}$. In general a selection rule is, for each $\underset{\sim}{x}$, a probability distribution on A. But as described in Berger (1979a), for our purposes a selection rule can be defined by the individual selection probabilities, $\psi(\underset{\sim}{x}) = (\psi_1(\underset{\sim}{x}), \ldots, \psi_k(\underset{\sim}{x}))$, where $\psi_i(\underset{\sim}{x})$ is the probability of including the i^{th} cell having observed $\underset{\sim}{X} = \underset{\sim}{x}$. A necessary and sufficient condition on ψ to insure the existence of selection rule which always selects a non-empty subset is $\sum_{i=1}^k \psi_i(\underset{\sim}{x}) \geq 1$ for all $\underset{\sim}{x}$.

Let P^* be a preassigned fixed number such that $1/k < P^* < 1$. As is traditional, the only selection rules to be considered are those which satisfy the P^*-condition, viz., $\inf_{\underset{\sim}{p}} P_{\underset{\sim}{p}}(CS|\psi) \geq P^*$. The set of all selection rules which satisfy the P^*-condition will be denoted by D_{P^*}.

The loss to be used herein is the number of non-best cells selected, S'. A non-best cell is any cell for which $p_i > p_{[1]}$. Thus the risk for a selection rule ψ at a parameter point $\underset{\sim}{p}$, i.e., the expected number of non-best cells selected, can be calculated from the individual selection probabilities by $E_{\underset{\sim}{p}}(S'|\psi) = \sum_{i \in a(\underset{\sim}{p})} E_{\underset{\sim}{p}} \psi_i(\underset{\sim}{X})$ where $a(\underset{\sim}{p}) = \{i \in \{1, 2, \ldots, k\}: p_i > p_{[1]}\}$. This definition of the loss and risk differs from the definition used elsewhere (see e.g. Berger (1979a)) if $p_{[1]} = p_{[2]}$ but it agrees with the usual definition if $p_{[1]} < p_{[2]}$ and has the advantage of being permutationally invariant. It is easily checked that the minimax and admissible selection rule to be derived herein is also minimax and admissible for the definition of S' used in Berger (1979a).

The subset selection problem as defined above is invariant under the group of permutations on the sample space. See Ferguson (1967) for the general definitions of invariance. If a selection rule is invariant under the group of permutations then these two relationships are true about the individual selection probabilities: (1) For every $i \in \{2, \ldots, k\}$ and every $\underset{\sim}{x}$, $\psi_i(\underset{\sim}{x}) = \psi_1(\underset{\sim}{y})$ where $y_1 = x_i$, $y_i = x_1$ and $y_j = x_j$ for j not equal to 1 or i; and (2) For every $i \in \{1, \ldots, k\}$, $\psi_i(\underset{\sim}{x}) = \psi_i(\underset{\sim}{y})$ where $x_i = y_i$ and $(y_1, \ldots, y_{i-1}, y_{i+1}, \ldots, y_k)$ is a permutation of $(x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_k)$. Permutationally invariant selection rules will be of interest since, by Theorem 2, page 156 of Ferguson (1967), a selection rule is admissible in the class of all selection rules if it is admissible in the class of invariant selection rules.

Finally, some results involving Schur functions and stochastic majorization will be used in the subsequent sections. All the notations, definitions and conventions will be as presented in Proschan and Sethuraman (1977) and Nevius, Proschan and Sethuraman (1977) and will not be repeated herein.

3. A CLASS OF SELECTION RULES.

In this section a class of selection rules is defined. The form of these rules is examined and the fact that these rules satisfy the P*-condition for certain values of P* is proven.

Selection rules of the following form will be examined. Let $0 < c \leq 1$ be a fixed constant. Define the individual selection probabilities by

$$(3.1) \qquad \psi_i^*(\underset{\sim}{x}) = \begin{cases} 1 & \sum_{\substack{j=1 \\ j \neq i}}^{k} c^{x_j} < M \\ \alpha & \qquad = \\ 0 & \qquad > \end{cases}$$

where the numbers $0 \leq \alpha < 1$ and $0 < M \leq k - 1$ are chosen so that $E_{\underset{\sim}{p}_0} \psi_i^*(\underset{\sim}{X}) = P^*$ where $\underset{\sim}{p}_0 = (1/k, \ldots, 1/k)$. The ψ_i^* defined by (3.1) satisfy the invariance property (1) of Section 2 so in the future all the discussion will be in terms of ψ_1^*. Now further constraints will be placed on c which will further limit the form of ψ_1^*. For each $\underset{\sim}{y}$, define $A(\underset{\sim}{y}) = \{\underset{\sim}{x}\colon\ x_1 = y_1$ and $(x_2, \ldots, x_k)$ is a permutation of $(y_2, \ldots, y_k)\}$.

<u>Lemma 3.1</u>. There exists $\varepsilon > 0$ such that, if $1 - \varepsilon < c \leq 1$, then

$$\psi_1^*(\underset{\sim}{x}) = \begin{cases} 1 & x_1 < t \text{ or } x_1 = t \text{ and } \sum_{j=2}^k c^{x_j} < M \\ \alpha & \underset{\sim}{x} \in A(\underset{\sim}{y}) \\ 0 & \text{otherwise} \end{cases} \tag{3.2}$$

for some $t \in \{0, \ldots, n\}$ and some $\underset{\sim}{y}$.

<u>Proof</u>. Let $f(c, \underset{\sim}{x}) = \sum_{j=2}^k c^{x_j}$. Clearly $f(c, \underset{\sim}{x}) = f(c, \underset{\sim}{y})$ for every c if $\underset{\sim}{x} \in A(\underset{\sim}{y})$. The lemma will be true if the following two facts are true for every $1 - \varepsilon < c < 1$: (i) $f(c, \underset{\sim}{x}) \neq f(c, \underset{\sim}{y})$ if $\underset{\sim}{x} \notin A(\underset{\sim}{y})$ and (ii) $f(c, \underset{\sim}{x}) > f(c, \underset{\sim}{y})$ if $x_1 > y_1$. To see (i), fix $\underset{\sim}{x}$ and $\underset{\sim}{y}$ with $\underset{\sim}{x} \notin A(\underset{\sim}{y})$. Then $f(c, \underset{\sim}{x})$ and $f(c, \underset{\sim}{y})$ are two different polynomials in c. Hence $f(c, \underset{\sim}{x}) = f(c, \underset{\sim}{y})$ for only a finite number of values for c. But $f(1, \underset{\sim}{x}) = k - 1 = f(1, \underset{\sim}{y})$. So there exists $\varepsilon > 0$ such that if $1 - \varepsilon < c < 1$ then $f(c, \underset{\sim}{x}) \neq f(c, \underset{\sim}{y})$. By considering all such pairs $\underset{\sim}{x}$ and $\underset{\sim}{y}$ (there are only a finite number of such pairs in the sample space) and taking the minimum ε obtained, an $\varepsilon > 0$ which works for any pair $\underset{\sim}{x}$ and $\underset{\sim}{y}$ is obtained. To see (ii), note that

$$\frac{d}{dc} f(c, \underset{\sim}{x})\Big|_{c=1} = \sum_{j=2}^k x_j = n - x_1 < n - y_1 = \sum_{j=2}^k y_j = \frac{d}{dc} f(c, \underset{\sim}{y})\Big|_{c=1} \,.$$

Since $f(c, \underset{\sim}{x})$ and $f(c, \underset{\sim}{y})$ are continuous functions of c, inequality (i) implies either $f(c, \underset{\sim}{x}) > f(c, \underset{\sim}{y})$ for every $1 - \varepsilon < c < 1$ or $f(c, \underset{\sim}{x}) < f(c, \underset{\sim}{y})$ for every $1 - \varepsilon < c < 1$. The inequality of the derivatives implies $f(c, \underset{\sim}{x}) > f(c, \underset{\sim}{y})$. ||

Note that every value of c satisfying $1 - \varepsilon < c < 1$ gives rise to the same ordering of the $\underset{\sim}{x}$'s. That is if the $\underset{\sim}{x}$'s in the sample space were to be ordered according to the value of the function $\sum_{j=2}^{k} c^{x_j}$, the same ordering would result from every c satisfying $1 - \varepsilon < c < 1$. Thus (3.2) defines only one selection rule, not different selection rules for different values of c.

Henceforth it will be assumed that c has been chosen so that ψ_1^* has the form given in (3.2).

To insure that the selection rule ψ^* will always select at least one cell, the individual selection probabilities must satisfy $\sum_{i=1}^{k} \psi_i^*(\underset{\sim}{x}) \geq 1$ for all $\underset{\sim}{x}$. This will be true if $P^* \geq P_{\underset{\sim}{p}_0}(X_1 < n/k) + P_{\underset{\sim}{p}_0}(\underset{\sim}{X} = (n/k, \ldots, n/k))/k$. This lower bound converges to 1/2 as $n \to \infty$. Thus the rule ψ^* cannot be used for very small P^*. In practice P^* is usually chosen to be near one so this is not a serious restriction. In the following theorem the range of possible P^* values is restricted even further in order to ensure that ψ^* satisfies the P^*-condition.

<u>Theorem 3.1</u>. Let $P^* \geq P_{\underset{\sim}{p}_0}(X_1 < n(k - 1)/k)$. Then

$$\inf_{\underset{\sim}{p}} P_{\underset{\sim}{p}}(CS|\psi^*) = \inf_{\underset{\sim}{p} \in \mathcal{P}_1} E_{\underset{\sim}{p}} \psi_1^*(\underset{\sim}{X}) = E_{\underset{\sim}{p}_0} \psi_1^*(\underset{\sim}{X}) = P^*$$

where $\mathcal{P}_1 = \{\underset{\sim}{p}: p_1 = p_{[1]}\}$.

<u>Proof</u>. The first equality is true by the invariance of ψ^*. The last equality is true by the definition of ψ^*. The last equality is true by the definition of ψ^*. Only the middle equality remains to be proven.

Let $\underset{\sim}{p} \in P_1$. Define $\underset{\sim}{p}'$ by $p_1' = \dots = p_{k-1}' = p_1$ and $p_k' = 1 - (k-1)p_1$. First it will be shown that $E_{\underset{\sim}{p}}\psi_1^*(\underset{\sim}{X}) \geq E_{\underset{\sim}{p}'}\psi_1^*(\underset{\sim}{X})$. Since $c > 0$, c^x is a convex function of x. Thus $\sum_{i=2}^k c^{x_i}$ is a Schur convex function of $(x_2, \dots, x_k)$. Thus $\psi_1^*(t, x_2, \dots, x_k)$ is a Schur concave function of $(x_2, \dots, x_k)$ on the set $\{(x_2, \dots, x_k)\colon \sum_{i=2}^k x_k = n - t\}$. The conditional distribution of $(X_2, \dots, X_k)$ given $X_1 = t$ is a multinomial distribution. So, by Application 4.2a of Nevius, Proschan and Sethuraman (1977), $E_{\underset{\sim}{p}}(\psi_1^*(X)\,|\,X_1 = t)$ is a Schur concave function of $(p_2, \dots, p_k)$ for fixed p_1. Thus $E_{\underset{\sim}{p}}(\psi_1^*(\underset{\sim}{X})\,|\,X_1 = t) \geq E_{\underset{\sim}{p}'}(\psi_1^*(\underset{\sim}{X})\,|\,X_1 = t)$ since $(p_2', \dots, p_k')$ majorizes $(p_2, \dots, p_k)$. On the other hand, $P_{\underset{\sim}{p}}(X_1 < t) = P_{\underset{\sim}{p}'}(X_1 < t)$ and $P_{\underset{\sim}{p}}(X_1 = t) = P_{\underset{\sim}{p}'}(X_1 = t)$ since these probabilities depend only on p_1 and $p_1 = p_1'$. Hence

$$\begin{aligned} E_{\underset{\sim}{p}}\psi_1^*(\underset{\sim}{X}) &= P_{\underset{\sim}{p}}(X_1 < t) + E_{\underset{\sim}{p}}(\psi_1^*(\underset{\sim}{X})\,|\,X_1 = t)P_{\underset{\sim}{p}}(X_1 = t) \\ &\geq P_{\underset{\sim}{p}'}(X_1 < t) + E_{\underset{\sim}{p}'}(\psi_1^*(X)\,|\,X_1 = t)P_{\underset{\sim}{p}'}(X_1 = t) \\ &= E_{\underset{\sim}{p}'}\psi_1^*(\underset{\sim}{X}). \end{aligned}$$

It remains to show that, for any $\underset{\sim}{p}$ of the form $(p, \dots, p, q)$ where $q = 1 - (k-1)p$ and $p \leq 1/k$; $E_{\underset{\sim}{p}}\psi_1^*(\underset{\sim}{X}) \geq E_{\underset{\sim}{p}_0}\psi_1^*(\underset{\sim}{X})$. By examining the derivative of $p^{n-x_k}(1 - (k-1)p)^{x_k}$ with respect to p, it is easily verified that this expression is a non-decreasing function of p on $0 \leq p \leq 1/k$ if $x_k \leq n/k$. If $P^* \geq P_{\underset{\sim}{p}_0}(X_1 < n(k-1)/k)$ then $t \geq n - n/k$. Thus $\psi_1^*(\underset{\sim}{x}) < 1$ implies $x_k \leq n/k$. This further implies that if $\psi_1^*(\underset{\sim}{x}) < 1$ then $P_{\underset{\sim}{p}}(\underset{\sim}{X} = \underset{\sim}{x}) = n!\, p^{n-x_k}(1 - (k-1)p)^{x_k}/(x_1!\dots x_k!) \leq n!\, (1/k)^{n-x_k}(1 - (k-1)(1/k))^{x_k}/(x_1!\dots x_k!) = P_{\underset{\sim}{p}_0}(\underset{\sim}{X} = \underset{\sim}{x})$. Let $T = \{\underset{\sim}{x}\colon\ \psi_1^*(\underset{\sim}{x}) = 0\}$. Then

$$E_{\underset{\sim}{p}}\psi_1^*(\underset{\sim}{X}) = 1 - \sum_{\underset{\sim}{x}\in T} P_{\underset{\sim}{p}}(\underset{\sim}{X} = \underset{\sim}{x}) - (1-\alpha)\sum_{\underset{\sim}{x}\in A(\chi)} P_{\underset{\sim}{p}}(\underset{\sim}{X} = \underset{\sim}{x})$$

$$\geq 1 - \sum_{\underset{\sim}{x}\in T} P_{\underset{\sim}{p}_0}(\underset{\sim}{X} = \underset{\sim}{x}) - (1-\alpha)\sum_{\underset{\sim}{x}\in A(\chi)} P_{\underset{\sim}{p}_0}(\underset{\sim}{X} = \underset{\sim}{x})$$

$$= E_{\underset{\sim}{p}_0}\psi_1^*(\underset{\sim}{X}).$$

This verifies the middle equality. ||

The result of Nevius, Proschan and Sethuraman used in the above proof was also proved by Rinott (1973).

Further values of P^* for which ψ^* will satisfy the P^* condition are given by Theorem 3.2.

Theorem 3.2. Let $P^* = P_{\underset{\sim}{p}_0}(X_1 < t)$ for some t. Then $\inf_{\underset{\sim}{p}} P(CS|\psi^*) = P^*$.

Proof. If $P^* = P_{\underset{\sim}{p}_0}(X_1 < t)$, then $\psi_1^*(\underset{\sim}{x}) = 1$ if $x_1 < t$ and $\psi_1^*(\underset{\sim}{x}) = 0$ if $x_1 \geq t$. The equality follows from the MLR property of the binomial distribution. ||

The values of P^* specified by Theorem 3.2 correspond to certain simple rules, investigated by Nagel (1970), for selecting the most probable cell.

Henceforth it will be assumed that P^* was chosen so that the condition of either Theorem 3.1 or 3.2 is satisfied. The restriction used in the proof of Theorem 3.1 that $P_{\underset{\sim}{p}}(\underset{\sim}{X} = \underset{\sim}{x}) \leq P_{\underset{\sim}{p}_0}(\underset{\sim}{X} = \underset{\sim}{x})$ for all $\underset{\sim}{x}$ with $\psi_1^*(x) < 1$ is rather strong. The fact that ψ^* satisfies the P^*-condition for some smaller values of P^*, as given by Theorem 3.2, leads the author to believe that ψ^* satisfies the P^*-condition for a wider range of values. But this has not been proven.

4. MINIMAXITY AND ADMISSIBILITY OF THE SELECTION RULE ψ^*.

In this section the minimaxity and admissibility of the selection rule ψ^*, defined by (3.1) and (3.2), in the class of rules D_{P^*} with respect to the loss S' is proven. First the minimaxity of ψ^* will be investigated.

Theorem 4.1. If $P^* \geq P_{\underset{\sim}{p}_0}(X_1 \leq n/2) - (k - 1)P_{\underset{\sim}{p}_0}(\underset{\sim}{X} = (n/2, n/2, 0, \ldots 0))/2$ then ψ^* is minimax with respect to S'.

Remark 4.1. If n is an odd number, the second term in this lower bound for P^* is zero. The only case in which this lower bound is larger (more restrictive) than the bound given in Theorem 3.1 is if $k = 2$ and n is even. In this case, this bound is the same as that given in Section 3 to ensure $\sum_{i=1}^{k} \psi_i^*(\underset{\sim}{x}) \geq 1$.

The following two lemmas will be used in the proof of Theorems 4.1 and 4.2.

Lemma 4.1. (a) If $\psi \in \mathcal{D}_{P^*}$ then $E_{\underset{\sim}{p}_0}\psi_i(\underset{\sim}{X}) \geq P^*$ for $1 \leq i \leq k$.

(b) The minimax value for S' is $(k - 1)P^*$.

(c) If ψ is minimax then $\sum_{i=2}^{k} E_{\underset{\sim}{p}_0}\psi_i(\underset{\sim}{X}) = (k - 1)P^*$.

Proof. These facts follow from the observation that $E_{\underset{\sim}{p}}\psi_i(\underset{\sim}{X})$ is a continuous function of $\underset{\sim}{p}$ for any selection rule and $\underset{\sim}{p}_0$ can be considered as the limit of a sequence of parameter points for which $p_i = p_{[1]} < p_{[2]}$. See Theorem 3.1 of Berger (1979a) for a similar proof with more details. ||

Lemma 4.2. If $P^* \geq P_{\underset{\sim}{p}_0}(X_1 \leq n/2) - (k - 1)P_{\underset{\sim}{p}_0}(\underset{\sim}{X} = (n/2, n/2, 0, \ldots, 0))/2$ then $S(\underset{\sim}{x}) = \sum_{i=1}^{k} \psi_i^*(\underset{\sim}{x})$ is a Schur concave function of $\underset{\sim}{x}$ on the sample space.

Proof. The inequality assumed for P^* and the definition (3.2) of ψ^* implies either $t > n/2$ or $t = n/2$, $y = (n/2, n/2, 0, \ldots, 0)$ and $\alpha \geq \frac{1}{2}$. (Recall t is defined to be an integer). Suppose $\underset{\sim}{x}$ majorizes $\underset{\sim}{y}$. Without loss of generality it will be assumed that $x_1 \geq x_2 \cdots \geq x_k$ and $y_1 \geq y_2 \geq \cdots \geq y_k$. Let $f(c, \underset{\sim}{x}) = \sum_{i=2}^{k} c^{x_i}$.

Case 1: $y_1 < t$ or $y_1 = t$ and $f(c, \underset{\sim}{y}) < M$. Then $S(\underset{\sim}{y}) = k \geq S(\underset{\sim}{x})$.

Case 2: $y_1 = t = n/2, f(c, \underset{\sim}{y}) = M$. Then $\underset{\sim}{y} \in A(n/2, n/2, 0, \ldots, 0)$. Since $\underset{\sim}{x}$ majorizes $\underset{\sim}{y}$, either $\underset{\sim}{x} \in A(n/2, n/2, 0, \ldots, 0)$ or $x_1 > t$. In the first case $S(\underset{\sim}{y}) = S(\underset{\sim}{x})$ and in the second case $S(\underset{\sim}{y}) = (k - 2) + 2\alpha \geq (k - 1) = S(\underset{\sim}{x})$ since $\alpha \geq \frac{1}{2}$.

Case 3: $y_1 = t > n/2$. Since $\sum_{i=1}^k x_i = n = \sum_{i=1}^k y_i$, $x_1 \geq y_1 > n/2$ implies $x_i < n/2 < t$ and $y_i < n/2 < t$ for $2 \leq i \leq k$. So $\sum_{i=2}^k \psi_i^*(\underset{\sim}{x}) = k - 1 = \sum_{i=2}^k \psi_i^*(\underset{\sim}{x})$. If $x_1 > y_1$, $\psi_1^*(\underset{\sim}{y}) \geq 0 = \psi_1^*(\underset{\sim}{x})$. If $x_1 = y_1$, then $(x_2, \ldots, x_k)$ majorizes $(y_2, \ldots, y_k)$. ψ_1^* is a Schur concave function of $(x_2, \ldots, x_k)$ for fixed x_1 (as in the proof of Theorem 3.1) so $\psi_1^*(\underset{\sim}{y}) \geq \psi_1^*(\underset{\sim}{x})$. In either case, $S(\underset{\sim}{y}) \geq S(\underset{\sim}{x})$.

Case 4: $y_1 > t$. As in Case 3, $x_i < n/2 \leq t$ and $y_i < n/2 \leq t$ for $2 \leq i \leq k$ so $S(\underset{\sim}{y}) = k - 1 = S(\underset{\sim}{x})$. ||

Proof of Theorem 4.1. $S(\underset{\sim}{x})$ is a Schur concave function of $\underset{\sim}{x}$ by Lemma 4.2. By Application 4.2a of Nevius, Proschan and Sethuraman (1977), $E_{\underset{\sim}{p}} S(\underset{\sim}{X}) = \sum_{i=1}^k E_{\underset{\sim}{p}} \psi_i(\underset{\sim}{X})$ is a Schur concave function of $\underset{\sim}{p}$ and thus is maximized at $\underset{\sim}{p}_0$. By the definition of ψ^*, $E_{\underset{\sim}{p}_0} S(\underset{\sim}{X}) = kP^*$. Fix $\underset{\sim}{p}$. Assume $p_j = p_{[1]}$.

$$
\begin{aligned}
(k - 1)P^* &= kP^* - P^* \\
&\geq E_{\underset{\sim}{p}} S(\underset{\sim}{X}) - P^* \\
&\geq E_{\underset{\sim}{p}} S(\underset{\sim}{X}) - P_{\underset{\sim}{p}}(CS|\psi^*) \\
&= E_{\underset{\sim}{p}} S(\underset{\sim}{X}) - E_{\underset{\sim}{p}} \psi_j^*(\underset{\sim}{X}) \\
&= \sum_{\substack{i=1 \\ i \neq j}}^k E_{\underset{\sim}{p}} \psi_i^*(\underset{\sim}{X}) \geq \sum_{i \in a(\underset{\sim}{p})} E_{\underset{\sim}{p}} \psi_i^*(X) = E_{\underset{\sim}{p}}(S'|\psi^*).
\end{aligned}
$$

By Lemma 4.1b, ψ^* is minimax with respect to S'. ||

Remark 4.2: The proof of Theorem 4.1 also shows that ψ^* is minimax with respect to the loss S, the number of populations selected, which was investigated by Berger (1979a), Bjørnstad (1978) and many others. See these two papers for further references. But no admissibility claims can be made for ψ^* for the loss S.

Now the admissibility of ψ^* will be proved.

Theorem 4.2: ψ^* is admissible with respect to S'.

Proof. By Theorem 2, page 156 of Ferguson (1967), it suffices to prove that ψ^* is admissible in the class of permutation invariant rules in $\mathcal{D}_{P^*}$. Let $\underset{\sim}{p} = (p, q, \ldots, q)$ where $(k - 1)q + p = 1$ and $1 - \varepsilon < p/q < 1$ for the ε specified by Lemma 3.1. For any invariant rule ψ

$$E_{\underset{\sim}{p}}(S'|\psi) = \sum_{i=2}^{k}\sum_{\underset{\sim}{x}}\psi_i(\underset{\sim}{x})n!\prod_{j=1}^{k}(p_j^{x_j}/x_j!)$$

$$= \sum_{\underset{\sim}{x}}\psi_1(\underset{\sim}{x})\sum_{i=2}^{k}[n!(p_1^{x_i}/x_i!)(p_i^{x_1}/x_1!)\prod_{\substack{j=2\\ j\neq i}}^{k}(p_j^{x_j}/x_j!)]$$

$$= \sum_{\underset{\sim}{x}}\psi_1(\underset{\sim}{x})\sum_{i=2}^{k}n!q^n(p/q)^{x_i}/(x_1!\ldots x_k!)$$

By Lemma 4.1a, every permutation invariant rule in $\mathcal{D}_{P^*}$ satisfies $E_{\underset{\sim}{p}_0}\psi_1(\underset{\sim}{x}) \geq P^*$. By the Neymann-Pearson Lemma (See Lehmann (1959)) any such individual selection probability, ψ_1, which minimizes $E_{\underset{\sim}{p}}(S'|\psi)$ must satisfy

$$\psi_1(\underset{\sim}{x}) = \begin{cases} 1 & q^n\sum_{i=2}^{k}(p/q)^{x_i} < C/k^n \\ 0 & \qquad\qquad > \end{cases}$$

where C is a constant and the factorials have cancelled from both sides of the inequalities. This is equivalent to

$$\text{(4.1)} \qquad \psi_1(\underset{\sim}{x}) \begin{cases} 1 & \sum_{i=2}^{k} c^{x_i} < M \\ 0 & \quad > \end{cases}$$

where $c = p/q$ and $M = C/(qk)^n$. (4.1) is the form given by (3.1) and, since $1 - \varepsilon < p/q < 1$, (4.1) is of the form (3.2) be Lemma 3.1. Furthermore, by Lemma 3.1 the set of $\underset{\sim}{x}$ in the sample space for which $\sum_{i=2}^{k} c^{x_i} = M$ is $A(\underset{\sim}{y})$ for some $\underset{\sim}{y}$. Since $\psi_1(\underset{\sim}{x})$ is invariant, $\psi_1(\underset{\sim}{x})$ is constant on $A(\underset{\sim}{y})$. So $\psi_1(\underset{\sim}{x})$ must be exactly of the form (3.2). (The Neyman-Pearson Lemma would have allowed different values of α for different $\underset{\sim}{x}$'s in $A(\underset{\sim}{y})$). Thus ψ_1^* corresponds to the unique permutation invariant rule in $\mathcal{D}_{P^*}$ which minimizes $E_{\underset{\sim}{p}}(S'|\psi)$. Thus ψ^* is admissible among the permutation invariant rules in $\mathcal{D}_{P^*}$. ||

The results of Sections 3 and 4 show that for fixed values of k and n, if P* is sufficiently large, ψ^* is minimax and admissible with respect to S'. This result could be extended if ψ^* could be shown to satisfy the P*-condition for smaller values of P* since the bound on P* in Theorem 3.1 is usually the largest. More work is needed to find minimax and admissible selection rules for smaller values of P*.

The problem of selecting a subset containing the most probable cell also leads to rules of the form (3.1) where now $c \geq 1$. But the author has been unable to verify the P*-condition except in certain special cases corresponding to Theorem 3.2. This problem too requires further investigation.

REFERENCES

Alam, K. and Thompson, J. R. (1972). On selecting the least probable multinomial event. Ann. Math. Statist., 43, 1981-1990.

Berger, R. L. (1979a). Minimax subset selection for loss measured by subset size. Ann. Statist., 7.

Berger, R. L. (1979b). Minimax subset selection for the multinomial distribution. Florida State University Statistics Department Report M497.

Berger, R. L. and Gupta, S. S. (1979). Minimax subset selection rules with applications to unequal variance (unequal sample size) problems. Scand. J. Statist.

Bjørnstad, J. F. (1978). The subset selection problem. II. On the optimality of some subset selection procedures. Purdue University Statistics Department Mimeo Series 78-27.

Ferguson, T. S. (1967). Mathematical Statistics: A Decision Theoretic Approach. Academic Press.

Gupta, S. S. and Huang, D. Y. (1975). On subset selection procedures for Poisson populations and some applications to the multinomial selection problems. Applied Statistics (R. P. Gupta, ed.). North Holland Pub. Co.

Gupta, S. S. and Nagel, K. (1967). On selection and ranking procedures and order statistics from the multinomial distribution. Sankhya Ser. B, 29, 1-34.

Lehmann, E. L. (1959). Testing Statistical Hypotheses. John Wiley and Sons.

Nagel, K. (1970). On subset selection rules with certain optimality properties. Ph.D. Thesis, Purdue University.

Nevius, S. E., Proschan, F. and Sethuraman, J. (1977). Schur functions in statistics II. Stochastic majorization. Ann. Statist., 5, 263-274.

Panchapakesan, S. (1971). On subset selection procedures for the most probable event in a multinomial distribution. Statistical Decision Theory and Related Topics. (S. S. Gupta and J. Yackel, eds.). Academic Press.

Proschan, F. and Sethuraman, J. (1977). Schur functions in statistics I. The preservation theorem. Ann. Statist., 5, 256-262.

Rinott, Y. (1973). Multivariate majorization and rearrangement inequalities with some applications to probability and statistics. Israel J. Math., 15, 60-77.

UNCLASSIFIED
SECURITY CLASSIFICATION OF THIS PAGE

REPORT DOCUMENTATION PAGE

1. REPORT NUMBER

FSU No. M522
USARO No. D39

2. GOVT ACCESSION NO.

3. RECIPIENT'S CATALOG NUMBER

4. TITLE

A MINIMAX AND ADMISSIBLE SUBSET SELECTION RULE FOR THE LEAST PROBABLE MULTINOMIAL CELL

5. TYPE OF REPORT & PERIOD COVERED

Technical Report

6. PERFORMING ORG. REPORT NUMBER

FSU Statistics Report M522

7. AUTHOR(s)

Roger L. Berger

8. CONTRACT OR GRANT NUMBER(s)

USARO-DAAG29-79-C-0158

9. PERFORMING ORGANIZATION NAME & ADDRESS

The Florida State University
Department of Statistics
Tallahassee, Florida 32306

10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS

11. CONTROLLING OFFICE NAME & ADDRESS

U.S. Army Research Office-Durham
P.O. Box 12211
Research Traingle Park, NC 27709

12. REPORT DATE

September, 1979

13. NUMBER OF PAGES

12

14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office)

15. SECURITY CLASS (of this report)

Unclassified

15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this report)

Approved for public release; distribution unlimited.

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

18. SUPPLEMENTARY NOTES

19. KEY WORDS

Minimax, admissible, subset selection, multinomial, Schur function, stochastic majorization.

20. ABSTRACT

In this paper a new subset selection rule for selecting a subset containing the least probable multinomial cell is defined. The rule is shown to be minimax and admissible in the class of rules which have a preassigned probability of at least P* of selecting the least probable cell provided that P* is sufficiently large. The loss used is the number of non-best cells selected.