SUBSET SELECTION BASED ON LIKELIHOOD
FROM UNIFORM AND RELATED POPULATIONS

by

Jayanti Chotai

American Mathematical Society 1970 subject classification: Primary 62F07;
Secondary 62A10, 62F05.

Key words and phrases: Subset selection, likelihood ratio, order restrictions, uniform distribution.
Let \( \pi_1, \pi_2, \ldots, \pi_k \) be \( k (\geq 2) \) populations. Let \( \pi_i \) (\( i = 1, 2, \ldots, k \)) be characterized by the uniform distribution on \((a_i, b_i)\), where exactly one of \( a_i \) and \( b_i \) is unknown. With unequal sample sizes, suppose that we wish to select a random-size subset of the populations containing the one with the smallest value of \( \theta_i = b_i - a_i \). Rule \( R_1 \) selects \( \pi_i \) iff a likelihood-based \( k \)-dimensional confidence region for the unknown \((\theta_1, \ldots, \theta_k)\) contains at least one point having \( \theta_i \) as its smallest component. A second rule, \( R_2 \), is derived through a likelihood ratio and is equivalent to that of Barr and Rizvi (1966) when the sample sizes are equal. Numerical comparisons are made. The results apply to the larger class of densities \( g(z; \theta_i) = M(z)Q(\theta_i) \) iff \( a(\theta_i) < z < b(\theta_i) \). Extensions to the cases when both \( a_i \) and \( b_i \) are unknown and when \( \theta_{\text{max}} \) is of interest are indicated.
1. INTRODUCTION

Let $\pi_1, \ldots, \pi_k$ be $(k>2)$ given populations and assume that $\pi_i (i = 1, \ldots, k)$ is characterized by a probability distribution depending on an unknown parameter $\theta_i$. Let $\theta_1 \leq \ldots \leq \theta_k$ denote their ordered values and let $\pi_{[1]}, \ldots, \pi_{[k]}$ denote the corresponding populations. We denote the $k$-dimensional parameter space for $\theta = (\theta_1, \ldots, \theta_k)$ by $\Omega$. Based on independent samples from the populations, suppose that we are interested in selecting a random-size subset of the populations which hopefully contains the best population (which may be $\pi_{[1]}$ or $\pi_{[k]}$).

This problem has been considered extensively in the literature; see Seal (1955) and Gupta (1965, 1977). A bibliography containing over six hundred items has been compiled by Kulldorff (1977). Most of the selection procedures which have appeared (when selecting for $\pi_{[k]}$) may be expressed as

$$R(h): \text{Select } \pi_i \text{ iff } h(T_i) \geq \max_{1 \leq j \leq k} T_j,$$

where $h(x)$ is a suitable function and for each $i$, $T_i$ is a suitable estimate of $\theta_i$; see Gupta and Panchapakesan (1972). By a correct selection (CS) is meant selection of a subset that contains the best population. Under the $P^*$-approach considered in the above references, one requires

$$\inf_{\theta \in \Omega} P(\text{CS}) \geq P^*, \quad (1.1)$$

where $P^*$ is a prespecified constant. Requirement (1.1) is known as the basic probability requirement or the $P^*$-condition.

In this paper, we assume that $\pi_i (i = 1, \ldots, k)$ is characterized by $U(a_i, b_i)$, the uniform distribution on $(a_i, b_i)$. Assume that for each $i$, one of $a_i$ and $b_i$ is known and that the other is unknown. Suppose that we take a random sample $\{Z_{i1}, \ldots, Z_{in_i}\}$ of size $n_i$ from $\pi_i (i = 1, \ldots, k)$ and that the best population is the one with the smallest value of
2. \( \theta_i = b_i - a_i \). Since we may consider \( Z_{ij} - a_i \) if \( a_i \) is known and \( b_i - Z_{ij} \) if \( b_i \) is known, we shall henceforth assume that \( \pi_i \) is characterized by \( U(0, \theta_i) \) with \( \theta_i > 0 \). In the derivation of the procedure \( R_1 \) of Section 2, we consider a likelihood-based confidence region for \( \theta \) and select \( \pi_i \) iff the region contains at least one \( \theta \) having \( \theta_i \) as its smallest component.

Rule \( R \), which reduces to that given by Barr and Rizvi (1966) when \( n_1 = \ldots = n_k \), is derived through a likelihood ratio given in Section 3. For the case \( n_1 = \ldots = n_k \), comparisons are made between \( R_1 \) and \( R \) for \( k = 3 \) in Section 4 and for \( k = 10 \) in Section 5. Section 6 contains some extensions and generalizations.

Although the uniform distribution is of interest as such, there are also other reasons why the results of this paper would be of interest. Firstly, the tables and formulae given here in fact apply to a much larger class of distributions as considered in Section 6.3. Secondly, the approach used here to derive selection procedures is different from the ones usually considered in the literature where the "slippage configuration" plays an important part. For the normal means problem, a detailed study of the rules derived through the likelihood approach appears in Chotai (1978); its extension to cover an exponential class of distributions and other generalizations will be treated elsewhere. Thirdly, there has recently been a growing interest to formulate the subset selection problems in terms of realistic loss functions rather than the \( P^* \)-approach. Since Bayes procedures are often difficult to obtain explicitly, it is of interest to approximate them by simple but intuitively appealing selection procedures. From this point of view, the rules derived by the likelihood approach are natural competitors of \( R(h) \) above; see Section 4.2.

2. THE SELECTION PROCEDURE \( R_1 \)

The likelihood of the total sample

\[
\mathbf{z} = (z_{11}, \ldots, z_{1n_1}, \ldots, z_{k1}, \ldots, z_{kn_k})
\]
is given by
\[ L(\mathbf{z}; \theta) = \prod_{i=1}^{k} \theta_{i}^{-n_{i}} \text{ if } z_{ij} \in (0, \theta_{i}) \]
for \( j = 1, \ldots, n_{i}; \ i = 1, \ldots, k; \) and is zero otherwise. We denote the parameter space by
\[ \Omega = \{(\theta_{1}, \ldots, \theta_{k}): \theta_{j} > 0 \text{ for all } j\}. \]

Now for \( i = 1, \ldots, k, \) let
\[ \Omega_{i} = \{\theta \in \Omega: \theta_{i} = \hat{\theta}_{i} \}. \tag{2.1} \]

In words, \( \Omega_{i} \) is the subspace of \( \Omega \) where the \( i: \text{th} \) component is the smallest. Now the maximum likelihood estimator of \( \theta \) is given by \( \hat{\theta} = (Y_{1}, \ldots, Y_{k}) \), where \( Y_{i} \) denotes the maximum of the observations from \( \pi_{i} \) \((i = 1, \ldots, k)\). Let \( c_{1}, \ 0 \leq c_{1} \leq 1, \) be a given constant and let
\[ \Omega(c_{1}) = \{\theta \in \Omega: L(\mathbf{z}; \theta) \geq c_{1}L(\mathbf{z}; \hat{\theta})\}. \]

Now consider the following selection procedure
\[ R_{i}: \text{Select } \pi_{i} \text{ iff } \Omega(c_{1}) \cap \Omega_{i} \text{ is nonempty.} \]

We thus include \( \pi_{i} \) in the selected subset iff a likelihood-based confidence region for the unknown \( \theta \) contains at least one point having its \( i: \text{th} \) component as the smallest. This is equivalent to requiring that
\[ \sup_{\theta \in \Omega_{i}} L(\mathbf{z}; \theta) \geq c_{1}L(\mathbf{z}; \hat{\theta}). \tag{2.2} \]

Let \( \theta^{*} = (\theta_{1}^{*}, \ldots, \theta_{k}^{*}) \) denote the value of \( \theta \) that gives supremum in (2.2). Since \( \theta_{j} \geq Y_{j} \) for all \( j, \) it is easy to see that
\[ \theta_{j}^{*} = \begin{cases} Y_{j} & \text{if } Y_{j} < Y_{i} \\ Y_{i} & \text{if } Y_{j} \geq Y_{i} \end{cases}. \tag{2.3} \]
Therefore, our rule may be expressed as

$$R_1: \text{Select } \pi_i \text{ iff } \prod_{j \in J_i} (Y_j/Y_i)_{n_j} \geq c_1,$$

where $J_i = \{j: Y_j \leq Y_i\}$.

It may be noted that the distribution of $Y_j^{n_j}$ is $U(0, \theta_j)$ for each $j$.

Given $X = \chi$, let $\psi_i(\chi)$ take on the value one if $\pi_i$ is selected and zero otherwise. Obviously $R_1$ is just; that is, for $i = 1, \ldots, k$, the function $\psi_i(\chi)$ is decreasing in $Y_i$ and increasing in each $Y_j$, $j \neq i$. Now for $j = 1, \ldots, k$, let $p_j = P(\pi_j \text{ is included in the selected subset})$. It follows easily from Seal (1958, Theorem 4.1) that if $n_1 = \ldots = n_k$, then $R_1$ is monotone; that is, $\theta_i \leq \theta_j$ implies $p_i \geq p_j$. We therefore have the following theorem.

**Theorem 2.1** Procedure $R_1$ is just and scale invariant. Furthermore, $R_1$ is monotone if $n_1 = \ldots = n_k$.

It may be noted that, as shown in Gupta and Nagel (1971), it follows from the above theorem that $P(\text{CS})$, as a function of $\Theta$, attains its infimum at a point where $\theta_k = \ldots = \theta_k$. Also, this infimum is independent of the common parameter value, which may therefore be set equal to unity. The following lemma simplifies our task of determining $c_1$ required to satisfy the $P^*$-condition.

**Lemma 2.2** Assume that $\theta_1 = \ldots = \theta_k = \theta$ and that $n_1 \leq \ldots \leq n_k$. We have $p_k \leq \ldots \leq p_1$.

**Proof** We may set $\theta = 1$. Choose $i$ and $k$ with $i < k$ arbitrarily and keep them fixed for rest of the proof. Let $r = n_i/n_k$ and consider the random variables $Y'_1, \ldots, Y'_k$ defined by $Y'_i = Y_i^r$, $Y'_k = Y_k^{1/r}$ and $Y'_j = Y_j$ for the remaining $j$. Then the dis-
tribution of \((Y_1', ..., Y_k')\) is the same as that obtained by interchanging \(Y_i\) and \(Y_j\) in \((Y_1, ..., Y_k)\). The lemma follows if we show that
\[
P[\prod_{j \in J_\lambda} (Y_j/Y_j')^{n_j} \geq c_1] \leq P[\prod_{j \in J_\lambda'} (Y_j/Y_j')^{n_j'} \geq c_1] \tag{2.4}
\]
where \(J_\lambda = \{j: Y_j < Y_j'\}\), \(J_\lambda' = \{j: Y_j' < Y_j\}\) and where \((n_1', ..., n_k')\) is obtained by interchanging \(n_i\) and \(n_j\) in \((n_1, ..., n_k)\).

Now it is easy to see that
\[
(Y_i/Y_i')^{n_i'} = (Y_i/Y_i')^{n_i} \geq (Y_i/Y_i')^{n_i}
\]
since we have assumed that \(n_i \leq n_i'\). Since \(J_\lambda' \subseteq J_\lambda'\),
\[
\prod_{j \in J_\lambda'} (Y_j/Y_j')^{n_j'} \geq \prod_{j \in J_\lambda} (Y_j/Y_j')^{n_j}
\]
and so (2.4), and consequently the lemma, follow.

The following theorem enables us to determine the required \(c_1\)-value.

**Theorem 2.3** Let \(d_1 = - \ln c_1\) and \(n_1 \leq ... \leq n_k\). For given \(P^*\), the value of \(d_1\) required to satisfy the \(P^*\)-condition is obtained by solving for \(d_1\) the equation
\[
P^* = A + \sum_{m=1}^{k-1} G_m(d_1) B_m \tag{2.5}
\]
where
\[
A = \sum_{m=0}^{k-1} (-1)^m \sum_{a \in S_m} (N_a + 1)^{-1}
\]
and
\[
B_m = \sum_{p=m}^{k-1} \left(\frac{p}{p-m}\right) (-1)^{p-m} \sum_{a \in S_p} (N_a + 1)^{-1}
\]
and where \(S_m\) denotes the set of all subsets of \(\{1, 2, ..., k-1\}\) having exactly \(m\) elements. Also, \(N_a = \sum_{j \in a} n_j/n_k\); and \(G_m(*)\)
denotes the cumulative distribution function for the standard gamma distribution with parameter \( m \).

**Proof** By Theorem 2.1 and Lemma 2.2 it suffices to assume that 
\( \theta_1 = \ldots = \theta_k = 1, \ n_1 \leq \ldots \leq n_k \) and then calculate \( d, 1 \) such that \( P^* = p_k \). But

\[
P_k = P(Y_1 > Y_k, \ldots, Y_{k-1} > Y_k) +
\]

\[
\sum_{m=1}^{k-1} \sum_{a \in S_m} \mathbb{P}(Y_j < Y_k \text{ for } j \in a, Y_j > Y_k \text{ for } j \notin a, \prod_{j=1}^{n_k} \frac{Y_j}{Y_k} > c_1).
\]

Now the random variables \( X_1, \ldots, X_k \), defined by \( X_j = -n_j \ln Y_j \) are independent, each with the standard exponential distribution. With \( \alpha_j = n_j/n_k \), we obtain

\[
p_k = E_1 + \sum_{m=1}^{k-1} \sum_{a \in S_m} E(a),
\]

where

\[
E_1 = P(X_1 \leq \alpha_1 X_k, \ldots, X_{k-1} \leq \alpha_{k-1} X_k)
\]

and

\[
E(a) = P(X_j > \alpha_j X_k \text{ for } j \in a, X_j \leq \alpha_j X_k \text{ for } j \notin a, \prod_{j=1}^{n_k} (X_j - \alpha_j X_k) < d_k).
\]

Now

\[
E_1 = \int_{0}^{\infty} \prod_{j=1}^{k-1} \left(1 - e^{-\alpha_j x} \right) e^{-x} \, dx,
\]

which is equal to \( A \). Also,

\[
E(a) = \int_{0}^{\infty} \prod_{j=1}^{k-1} \left(1 - e^{-\alpha_j x} \right) e^{-x} \, dx.
\]

...
Since the exponential distribution lacks memory, the above integral is equal to

\[ G_m(d_1) \int_0^\infty \prod_{j \in a} e^{-\lambda_j x} \prod_{j \notin a} (1 - e^{-\lambda_j x}) e^{-x} \, dx. \]

Now expanding the second product in the integrand above, we obtain

\[ \sum_{a \in S_m} E(a) = G_m(d_1) \sum_{p=m}^{k-1} \left( \frac{p}{p-m} \right) (-1)^{p-m} \sum_{a \in S_p} (N_a + 1)^{-1}. \]

We have thus proved (2.5), which completes the proof.

When the sample sizes are all equal, (2.5) simplifies to

\[ p^* = \frac{1}{k} + \sum_{m=1}^{k-1} \frac{1}{G_m(d_1)} \sum_{\nu=0}^{k-1-m} \binom{k-1}{\nu, m} (-1)^\nu (\nu + m + 1)^{-1} \]

(2.6)

where

\[ \binom{k-1}{\nu, m} = \frac{(k-1)!}{\nu! m! (k-1-\nu-m)!}. \]

For selected values of \( k \) and \( p^* \), Table I gives the value of \( d_1 \) satisfying (2.6).

3. THE SELECTION PROCEDURE \( R \)

In this section we are concerned with the following selection procedure

\[ R: \text{Select } \pi_i \iff \sup_{\theta \in \Omega_i} L(z; \theta) > c \sup_{\theta \in \Omega_i'} L(z; \theta) \]

where \( \Omega_i' = \{ \theta \in \Omega_i: \theta_i = \theta_j [1] \text{ or } \theta_i = \theta_j [2] \} \) and where \( \Omega_i \) is given by (2.1).
TABLE I
Values of \( d_1 \) to Implement \( R_1 \) With Equal Sample Sizes

<table>
<thead>
<tr>
<th>( P^* )</th>
<th>0.75</th>
<th>0.90</th>
<th>0.95</th>
<th>0.975</th>
<th>0.99</th>
</tr>
</thead>
<tbody>
<tr>
<td>( k )</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>0.693</td>
<td>1.609</td>
<td>2.302</td>
<td>2.983</td>
<td>3.901</td>
</tr>
<tr>
<td>3</td>
<td>1.556</td>
<td>2.765</td>
<td>3.622</td>
<td>4.450</td>
<td>5.512</td>
</tr>
<tr>
<td>4</td>
<td>2.344</td>
<td>3.795</td>
<td>4.789</td>
<td>5.732</td>
<td>6.926</td>
</tr>
<tr>
<td>5</td>
<td>3.108</td>
<td>4.775</td>
<td>5.891</td>
<td>6.935</td>
<td>8.187</td>
</tr>
<tr>
<td>6</td>
<td>3.860</td>
<td>5.725</td>
<td>6.945</td>
<td>8.071</td>
<td>9.457</td>
</tr>
<tr>
<td>8</td>
<td>5.347</td>
<td>7.595</td>
<td>9.023</td>
<td>10.320</td>
<td>11.902</td>
</tr>
<tr>
<td>11</td>
<td>7.564</td>
<td>10.345</td>
<td>12.053</td>
<td>13.576</td>
<td>15.408</td>
</tr>
<tr>
<td>13</td>
<td>9.040</td>
<td>12.159</td>
<td>14.040</td>
<td>15.701</td>
<td>17.683</td>
</tr>
<tr>
<td>15</td>
<td>10.517</td>
<td>13.965</td>
<td>16.010</td>
<td>17.801</td>
<td>19.925</td>
</tr>
<tr>
<td>17</td>
<td>11.996</td>
<td>15.764</td>
<td>17.967</td>
<td>19.883</td>
<td>22.140</td>
</tr>
<tr>
<td>18</td>
<td>12.735</td>
<td>16.663</td>
<td>18.942</td>
<td>20.918</td>
<td>23.240</td>
</tr>
<tr>
<td>20</td>
<td>14.216</td>
<td>18.456</td>
<td>20.887</td>
<td>22.980</td>
<td>25.426</td>
</tr>
</tbody>
</table>

Intuitive justification for this approach is clear. Now the likelihood in \( \Omega_i \) is maximized by \( \theta^* \) given by (2.3). It is also easy to see that the likelihood in \( \Omega_i^1 \) is maximized by \( \theta' \) given by \( \theta'_j = Y_j \) if \( \left( \frac{Y_j}{Y_i} \right)^{\frac{1}{n}} = \min \left( \frac{Y_j}{Y_r} \right)^{\frac{1}{r}} \) or if \( Y_j \geq Y_i \), and by \( \theta'_j = Y_i \) otherwise. This leads us to express \( R \) as follows.

\[
R: \text{Select } \pi_i \text{ iff } \min_{1 \leq j \leq k} \left( \frac{Y_j}{Y_i} \right)^{n_j} \geq c.
\]

When all the sample sizes are equal, the procedure turns out to be the same as that proposed by Barr and Rizvi (1966), and is of the type \( R(h) \) given in Section 1.

By using arguments similar to those of Section 2, it can be shown that the results of Theorem 2.1 and Lemma 2.2 are also valid.
The following theorem enables us to determine the c-value for R.

**Theorem 3.1** Let \( d = -\ln c \) and \( n_1 \leq \ldots \leq n_k \). For given \( P^* \), the value of \( d \) required to satisfy the \( P^* \)-condition is obtained by solving for \( d \) the equation

\[
P^* = \sum_{m=0}^{k-1} (-1)^m \sum_{a \in S_m} (N_a + md + 1)^{-1}
\]

where \( S_m \) denotes the set of all subsets of \( \{1, 2, \ldots, k-1\} \) having exactly \( m \) elements and \( N_a = \sum_{j \in a} n_j/n_k \).

**Proof** We may assume \( \theta_1 = \ldots = \theta_k = 1 \) and then set \( P^* = p_k \).
By the transformation \( X_j = n_j \ln Y_j \) and with \( \alpha_j = n_j/n_k \), we have

\[
p_k = P(X_j - \alpha_j X_k \leq d \text{ for } j = 1, \ldots, k)
\]

\[
= \int_0^{\infty} \prod_{j=1}^{k-1} (1 - e^{-\alpha_j x - d}) e^{-x} dx
\]

which equals the right hand side of (3.1), thus proving the theorem.

For the case of equal sample sizes, (3.1) simplifies to

\[
P^* = [1 - (1 - c)^k]/ck.
\]

Table II below gives the value of \( d = -\ln c \) satisfying (3.2) for selected values of \( k \) and \( P^* \).

**4. THE CASE OF THREE POPULATIONS AND COMMON SAMPLE SIZE**

In this section, we investigate in detail the performances of \( R_1 \) and \( R \) for the case \( k = 3 \) and when a random sample of size \( n \) is taken from each population. For simplicity in no-
**TABLE II**

Values of $d$ to Implement $\mathcal{R}$ With Equal Sample Sizes

<table>
<thead>
<tr>
<th>$P^*$</th>
<th>0.75</th>
<th>0.90</th>
<th>0.95</th>
<th>0.975</th>
<th>0.99</th>
</tr>
</thead>
<tbody>
<tr>
<td>$k$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>0.6931</td>
<td>1.6094</td>
<td>2.3026</td>
<td>2.9957</td>
<td>3.9120</td>
</tr>
<tr>
<td>3</td>
<td>1.2901</td>
<td>2.2674</td>
<td>2.9785</td>
<td>3.6803</td>
<td>4.6017</td>
</tr>
<tr>
<td>4</td>
<td>1.6636</td>
<td>2.6612</td>
<td>3.3784</td>
<td>4.0831</td>
<td>5.0062</td>
</tr>
<tr>
<td>5</td>
<td>1.9353</td>
<td>2.9430</td>
<td>3.6632</td>
<td>4.3694</td>
<td>5.2933</td>
</tr>
<tr>
<td>6</td>
<td>2.1488</td>
<td>3.1627</td>
<td>3.8847</td>
<td>4.5917</td>
<td>5.5159</td>
</tr>
<tr>
<td>7</td>
<td>2.3247</td>
<td>3.3426</td>
<td>4.0658</td>
<td>4.7733</td>
<td>5.6981</td>
</tr>
<tr>
<td>8</td>
<td>2.4744</td>
<td>3.4951</td>
<td>4.2191</td>
<td>4.9271</td>
<td>5.8521</td>
</tr>
<tr>
<td>9</td>
<td>2.6045</td>
<td>3.6274</td>
<td>4.3521</td>
<td>5.0604</td>
<td>5.9856</td>
</tr>
<tr>
<td>10</td>
<td>2.7196</td>
<td>3.7442</td>
<td>4.4694</td>
<td>5.1779</td>
<td>6.1033</td>
</tr>
<tr>
<td>11</td>
<td>2.8228</td>
<td>3.8488</td>
<td>4.5744</td>
<td>5.2831</td>
<td>6.2086</td>
</tr>
<tr>
<td>12</td>
<td>2.9164</td>
<td>3.9435</td>
<td>4.6694</td>
<td>5.3783</td>
<td>6.3038</td>
</tr>
<tr>
<td>13</td>
<td>3.0020</td>
<td>4.0299</td>
<td>4.7561</td>
<td>5.4652</td>
<td>6.3908</td>
</tr>
<tr>
<td>14</td>
<td>3.0808</td>
<td>4.1095</td>
<td>4.8360</td>
<td>5.5451</td>
<td>6.4708</td>
</tr>
<tr>
<td>15</td>
<td>3.1538</td>
<td>4.1833</td>
<td>4.9099</td>
<td>5.6191</td>
<td>6.5449</td>
</tr>
<tr>
<td>16</td>
<td>3.2219</td>
<td>4.2519</td>
<td>4.9787</td>
<td>5.6880</td>
<td>6.6138</td>
</tr>
<tr>
<td>17</td>
<td>3.2857</td>
<td>4.3162</td>
<td>5.0431</td>
<td>5.7525</td>
<td>6.6783</td>
</tr>
<tr>
<td>18</td>
<td>3.3456</td>
<td>4.3765</td>
<td>5.1036</td>
<td>5.8131</td>
<td>6.7389</td>
</tr>
<tr>
<td>19</td>
<td>3.4021</td>
<td>4.4335</td>
<td>5.1607</td>
<td>5.8702</td>
<td>6.7961</td>
</tr>
<tr>
<td>20</td>
<td>3.4556</td>
<td>4.4873</td>
<td>5.2146</td>
<td>5.9242</td>
<td>6.8501</td>
</tr>
<tr>
<td>25</td>
<td>3.6871</td>
<td>4.7202</td>
<td>5.4479</td>
<td>6.1576</td>
<td>7.0834</td>
</tr>
<tr>
<td>30</td>
<td>3.8750</td>
<td>4.9089</td>
<td>5.6368</td>
<td>6.3466</td>
<td>7.2727</td>
</tr>
<tr>
<td>35</td>
<td>4.0331</td>
<td>5.0676</td>
<td>5.7957</td>
<td>6.5057</td>
<td>7.4318</td>
</tr>
<tr>
<td>40</td>
<td>4.1696</td>
<td>5.2045</td>
<td>5.9328</td>
<td>6.6428</td>
<td>7.5690</td>
</tr>
<tr>
<td>45</td>
<td>4.2896</td>
<td>5.3250</td>
<td>6.0533</td>
<td>6.7634</td>
<td>7.6896</td>
</tr>
<tr>
<td>50</td>
<td>4.3968</td>
<td>5.4324</td>
<td>6.1609</td>
<td>6.8710</td>
<td>7.7973</td>
</tr>
</tbody>
</table>

In Section 4.1, we compare $R_1$ and $R$ under the $P^*$-approach. In Section 4.2, we assume a loss function. We then compare the procedures in terms of minimum expected loss for a given model when the $c_1$-value optimal for $R_1$ and the $c$-value optimal for $R$ are used.
4.1 The P*-approach

The selected subset $S$ would be one of the seven possible nonempty subsets of the three populations. We use the notation $s_j$, $s_{ij}$ and $s_{123}$ to denote the probability that $S = \{\pi_j\}$, $S = \{\pi_i, \pi_j\}$ and $S = \{\pi_1, \pi_2, \pi_3\}$, respectively. The expressions for these probabilities are derived in the Appendix (Section 7).

We begin the comparisons with the following theorem.

**Theorem 4.1** For $k = 3$, we have $P(CS|R_1) \geq P(CS|R)$ for any parameter configuration and for any $P^*$.

**Proof** Using the expressions given in the Appendix and the relation $P(CS) = 1 - s_2 - s_3 - s_{23}$, we obtain

$$3\delta_2(P(CS|R_1) - 1)/\delta_1 = c_1(\ln c_1 + 1) - 3c_1(\delta_2 + 1)/2\delta_1. \quad (4.1)$$

$$3\delta_2(P(CS|R) - 1)/\delta_1 = c^2 - 3c(\delta_2 + 1)/2\delta_1. \quad (4.2)$$

Now the constants $c_1$ and $c$ are obtained through

$$\begin{align*}
\begin{cases}
P^* = 1 + c_1(\ln c_1) / 3 - 2c_1 / 3 \\
P^* = 1 - c + c^2 / 3
\end{cases}
\end{align*} \quad (4.3)$$

which in turn imply

$$c^2 = c_1 \ln c_1 - 2c_1 + 3c. \quad (4.4)$$

Substituting (4.4) into (4.2), we see by comparing (4.1) with (4.2) that the inequality $P(CS|R_1) \geq P(CS|R)$ is equivalent to

$$c_1(2\delta_1 - \delta_2 - 1) \geq c(2\delta_1 - \delta_2 - 1),$$

which is equivalent to $c_1 \leq c$. It is a straightforward matter to verify $c_1 \leq c$ by examining (4.3), which proves the theorem.
With \( p_j \) denoting the probability of including \( \pi_j \) in the selected subset, we let \( \mathbb{E}(|a'|) = p_2 + p_3 \) denote the expected number of nonbest populations selected, and \( P(CS) = p_1 \). Let

\[
\Psi = \sum_{j \in a} j/|a|
\]

denote the average rank of the selected set \( a \) of indices and \( \mathbb{E}(\Psi) \) its expected value. For a good selection rule satisfying the \( P^* \)-condition, we desire the value of \( P(CS) \) to be high and the values of \( \mathbb{E}(|a'|) \) and \( \mathbb{E}(\Psi) \) to be low for any given parameter configuration.

To make comparisons between \( R_1 \) and \( R \) under various configurations of the underlying parameters, the following three types of configurations of \((\theta_1, \theta_2, \theta_3) = (\delta_1^{1/n}, \delta_2^{1/n}, 1)\) will be considered (with \( \delta \leq 1 \)):

- **(A)** \( \delta_1 = \delta, \ \delta_2 = 1 \)
- **(B)** \( \delta_1 = \delta^2, \ \delta_2 = \delta \)
- **(C)** \( \delta_1 = \delta_2 = \delta \).

Table III gives performance characteristics of the rules for each of (A), (B) and (C) with \( \delta = j/40 \); for \( P^* = 0.75 \) with \( j = 1, 4(5)39 \) and for \( P^* = 0.95 \) with \( j = 1(2)9(5)19(10)39 \). It can be seen from the table that in terms of \( \mathbb{E}(|a'|) \) or \( \mathbb{E}(\Psi) \), \( R_1 \) performs better than \( R \) for larger values of \( \delta \) or \( P^* \) whereas the opposite holds for smaller values of \( \delta \). It can be observed that for (B) and (C), \( R_1 \) usually gives smaller \( p_3 \). This is also true for (A) when \( \delta \) is large.

Tables IV gives lower bound to the value of \( \delta \) for which \( R_1 \) performs better than \( R \) with regards to different criteria, configuration types and values of \( P^* \).

It may be remarked that if the experimenter employing the \( P^* \)-approach is willing to rely on a probability model for \((\delta_1, \delta_2)\), we may compare the rules by taking expectations (over the parameter space) of the criteria of importance.
### TABLE III
Performance Characteristics of $R_1$ (Upper Entry) and $R$ (Lower Entry)

For $P^* = 0.75$

<table>
<thead>
<tr>
<th>$\delta$</th>
<th>configuration (A)</th>
<th></th>
<th>configuration (B)</th>
<th></th>
<th>configuration (C)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$P(CS)$</td>
<td>$P_2 = P_3$</td>
<td>$E(</td>
<td>a'</td>
<td>)$</td>
<td>$E(\Psi)$</td>
</tr>
<tr>
<td>$0.75$</td>
<td>$0.757$</td>
<td>$0.746$</td>
<td>$1.493$</td>
<td>$1.993$</td>
<td>$0.761$</td>
</tr>
<tr>
<td>$0.80$</td>
<td>$0.756$</td>
<td>$0.747$</td>
<td>$1.494$</td>
<td>$1.994$</td>
<td>$0.758$</td>
</tr>
<tr>
<td>$0.85$</td>
<td>$0.792$</td>
<td>$0.725$</td>
<td>$1.449$</td>
<td>$1.957$</td>
<td>$0.810$</td>
</tr>
<tr>
<td>$0.90$</td>
<td>$0.784$</td>
<td>$0.730$</td>
<td>$1.460$</td>
<td>$1.961$</td>
<td>$0.799$</td>
</tr>
<tr>
<td>$0.95$</td>
<td>$0.827$</td>
<td>$0.696$</td>
<td>$1.393$</td>
<td>$1.915$</td>
<td>$0.853$</td>
</tr>
<tr>
<td>$1.00$</td>
<td>$0.814$</td>
<td>$0.707$</td>
<td>$1.415$</td>
<td>$1.921$</td>
<td>$0.838$</td>
</tr>
<tr>
<td>$1.05$</td>
<td>$0.859$</td>
<td>$0.658$</td>
<td>$1.317$</td>
<td>$1.866$</td>
<td>$0.890$</td>
</tr>
<tr>
<td>$1.10$</td>
<td>$0.844$</td>
<td>$0.675$</td>
<td>$1.350$</td>
<td>$1.872$</td>
<td>$0.873$</td>
</tr>
<tr>
<td>$1.15$</td>
<td>$0.891$</td>
<td>$0.606$</td>
<td>$1.212$</td>
<td>$1.805$</td>
<td>$0.922$</td>
</tr>
<tr>
<td>$1.20$</td>
<td>$0.875$</td>
<td>$0.626$</td>
<td>$1.252$</td>
<td>$1.806$</td>
<td>$0.906$</td>
</tr>
<tr>
<td>$1.25$</td>
<td>$0.921$</td>
<td>$0.529$</td>
<td>$1.057$</td>
<td>$1.719$</td>
<td>$0.948$</td>
</tr>
<tr>
<td>$1.30$</td>
<td>$0.907$</td>
<td>$0.541$</td>
<td>$1.083$</td>
<td>$1.707$</td>
<td>$0.936$</td>
</tr>
<tr>
<td>$1.35$</td>
<td>$0.951$</td>
<td>$0.406$</td>
<td>$0.813$</td>
<td>$1.583$</td>
<td>$0.970$</td>
</tr>
<tr>
<td>$1.40$</td>
<td>$0.939$</td>
<td>$0.378$</td>
<td>$0.756$</td>
<td>$1.531$</td>
<td>$0.962$</td>
</tr>
<tr>
<td>$1.45$</td>
<td>$0.979$</td>
<td>$0.212$</td>
<td>$0.424$</td>
<td>$1.321$</td>
<td>$0.988$</td>
</tr>
<tr>
<td>$1.50$</td>
<td>$0.973$</td>
<td>$0.176$</td>
<td>$0.351$</td>
<td>$1.268$</td>
<td>$0.985$</td>
</tr>
<tr>
<td>$1.55$</td>
<td>$0.995$</td>
<td>$0.058$</td>
<td>$0.115$</td>
<td>$1.090$</td>
<td>$0.997$</td>
</tr>
<tr>
<td>$1.60$</td>
<td>$0.993$</td>
<td>$0.045$</td>
<td>$0.090$</td>
<td>$1.072$</td>
<td>$0.996$</td>
</tr>
</tbody>
</table>
TABLE IV
Lower Bound to the Value of $\delta$ for which $R_1$ Performs Better than $R$

<table>
<thead>
<tr>
<th>configuration</th>
<th>criterion</th>
<th>$P^*$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>0.75</td>
</tr>
<tr>
<td>(A)</td>
<td>$E(</td>
<td>a'</td>
</tr>
<tr>
<td></td>
<td>$E(\Psi)$</td>
<td>.45</td>
</tr>
<tr>
<td></td>
<td>$P_3$</td>
<td>.30</td>
</tr>
<tr>
<td>(B)</td>
<td>$E(</td>
<td>a'</td>
</tr>
<tr>
<td></td>
<td>$E(\Psi)$</td>
<td>.46</td>
</tr>
<tr>
<td></td>
<td>$P_3$</td>
<td>.30</td>
</tr>
<tr>
<td>(C)</td>
<td>$E(</td>
<td>a'</td>
</tr>
<tr>
<td></td>
<td>$E(\Psi)$</td>
<td>.00</td>
</tr>
<tr>
<td></td>
<td>$P_3$</td>
<td>.00</td>
</tr>
</tbody>
</table>

4.2 Some Other Comparisons

Selection of a subset of $k$ populations may be carried out by using an approach other than the $P^*$-approach. One such example is the Bayesian approach. However, except for certain types of loss functions and priors, Bayesian procedures are complicated as regards derivation and application. For that reason, if simple procedures like $R_1$ and $R$ are available that do almost as well as the Bayes procedure, they may be preferable. The problem then reduces to that of determining the constant for each of these simple rules that best approximates the Bayes procedure. However, it may be pointed out that determination of this optimal value would usually be a cumbersome task. For the normal means problem, comparisons between different rules in this respect appear in Chernoff and Yahav (1977), Chotai (1978) and Gupta and Hsu (1978).

For the present problem, it may be of interest to ask which one of $R_1$ and $R$ performs better in the above sense. Generally, it is reasonable to expect the answer to depend on the sort of loss function and the prior distribution assumed.
To make a limited comparison between $R_1$ and $R$, we consider the following loss function:

$$L = |a'| + \alpha \cdot ICS(\varnothing, a),$$

where $|a'|$ denotes the number of nonbest populations selected, $ICS(\varnothing, a)$ equals zero if the best population is included in the selected subset and unity otherwise, and where $\alpha \geq 0$ is a given constant. Note that for the loss function, $L_1$, appearing in Gupta and Hsu (1978), we have $L_1 = L + 1 - ICS(\varnothing, a)$. Assume that $(\delta_1, \delta_2) = (\theta_1^n, \theta_2^n)$ have the same joint distribution as $(U_{(1)}, U_{(2)})$, where $U_{(1)} \leq U_{(2)}$ are the order statistics based on two independent random variables with the uniform distribution on the interval $(0,1)$. It may be pointed out that since the choice of this model is based on mathematical simplicity rather than on application considerations, the limitations of the present comparisons should be borne in mind.

The conditional expectation $E(L)$ of $L$ for given $\delta_1$ and $\delta_2$ is

$$E(L) = 2 - 2s_1 + (\alpha - 1)s_2 + (\alpha - 1)s_3 - s_{12} - s_{13} + \alpha s_{23}.$$

For $R_1$, an explicit expression for $E(L)$ for each of the cases $c_1 \leq \delta_1 \delta_2$, $\delta_1 \delta_2 \leq c_1 \leq \delta_1/\delta_2$ and $\delta_1/\delta_2 \leq c_1$ may be obtained through the expressions given in the Appendix. Similarly, for $R$ we have the cases $c \leq \delta_1$, $\delta_1 \leq c \leq \delta_1/\delta_2$ and $\delta_1/\delta_2 \leq c$. The following expectations of $E(L)$ with respect to $(\delta_1, \delta_2)$ for $R_1$ and $R$ respectively may be obtained by straightforward but lengthy computations:

$$\ell(R_1) = 2 + [(37\alpha - 179)/6 + (71 - 4\alpha)(\ln c_1)/3$$

$$-7(\ln c_1)^2 + (\ln c_1)^3]c_1/18,$$

$$\ell(R) = 2 + [(5\alpha - 31)/4 + (25 - 2\alpha)c/9$$

$$+ 4.5 \ln c - 10 c(\ln c)/3 + c(\ln c)^2/3]c/3.$$
TABLE V

Performances of $R_1$ (Upper Entry) and $R$ (Lower Entry) for the Given Model

\[
\begin{array}{cccc}
\alpha & d\text{-value} & l^* & P^* \\
1.28 & .0 & .781 & .333 \\
 & .0 & .781 & .333 \\
2.0 & .400 & .991 & .464 \\
 & .325 & .984 & .451 \\
3.0 & .851 & 1.199 & .594 \\
 & .670 & 1.196 & .575 \\
4.0 & 1.226 & 1.348 & .684 \\
 & .966 & 1.361 & .668 \\
6.0 & 1.840 & 1.545 & .797 \\
 & 1.493 & 1.599 & .792 \\
8.0 & 2.344 & 1.668 & .861 \\
 & 1.983 & 1.750 & .869 \\
10.0 & 2.779 & 1.750 & .901 \\
 & 2.461 & 1.846 & .917 \\
15.0 & 3.686 & 1.866 & .952 \\
 & 3.669 & 1.957 & .975 \\
20.0 & 4.435 & 1.922 & .975 \\
 & 4.937 & 1.989 & .993 \\
\infty & \infty & 2.0 & 1.0 \\
\infty & 2.0 & 1.0 \\
\end{array}
\]

It turns out that the optimal value of $c_1$ or $c$ is unity if $\alpha \leq 37/29 = 1.28$, in which case only one population is selected. For several values of $\alpha$, Table V gives (the d-values)
\[d_1 = -\ln c_1 \text{ and } d = -\ln c,\]
where $c_1$ and $c$ are the optimal values. The table also gives the expected loss $l^*$ and the value of $P^* = \inf_{\omega \in \Omega} P(C_S)$ attained, when these optimal values are used.

Our study reveals that for the given model, $R_1$ performs better than $R$ if $\alpha$ exceeds approximately 3.5.

5. THE CASE OF MANY POPULATIONS AND COMMON SAMPLE SIZE

When $k$ is large, determination of probabilities of selec-
ting the various possible subsets becomes lengthy for arbitrary parameter configurations. Assuming common sample size $n$, we therefore restrict our comparison between $R_1$ and $R$ to the slippage configuration:

$$\theta_1 = \delta^{1/n} \leq 1, \theta_2 = \ldots = \theta_k = 1.$$ 

Numerical comparisons will be made for $k = 10$ under the $P^*$-approach.

For this, let $Y_1, \ldots, Y_k$ be independent random variables such that $Y_1^n$ has the uniform distribution on $(0, \delta)$, while each $Y_j^n, j \neq 1$, is uniform on $(0, 1)$. Now for $R_1$,

$$P(CS|R_1) = p_1(R_1) = A + \sum_{m=1}^{k-1} \binom{k-1}{m} B_m$$

where $A = P(Y_2 \geq Y_1, \ldots, Y_k \geq Y_1)$ and

$$B_m = P(\prod_{j=2}^{m+1} (Y_j/Y_1)^n \geq c_1, Y_2 \leq Y_1, \ldots, Y_{m+1} \leq Y_1, Y_{m+2} \geq Y_1, \ldots, Y_k \geq Y_1).$$

Setting $X_j = -\ln Y_j^n$, we obtain

$$A = P(X_2 \leq X_1, \ldots, X_k \leq X_1)$$

$$= \int_{-\ln \delta}^{\infty} (1 - e^{-x})^{k-1} e^{-x - 1} \, dx = [1 - (1 - \delta)^k]/k\delta.$$

Also, with $d_1 = -\ln c_1$

$$B_m = \int_{-\ln \delta}^{\infty} P(\sum_{j=2}^{m+1} (X_j - x) \leq d_1, X_2 > x, \ldots, X_{m+1} > x)$$

$$\cdot P(X_{m+2} \leq x, \ldots, X_k \leq x)e^{-x\delta - 1} \, dx.$$

Since the exponential distribution lacks memory, we get
\[ B_m = \delta^{-1} G_m(d_1) \int_{-\infty}^{\infty} (1 - e^{-x})^{k-l-m}e^{-(m+1)x} \, dx \]
\[ = \delta^m G_m(d_1) \sum_{\nu=0}^{k-l-m} \frac{(k-l-m)}{\nu} (-\delta)^\nu (\nu + m + 1)^{-1}, \]

where \( G_m(\cdot) \) is the cumulative distribution function for standard gamma distribution with parameter \( m \).

As regards the probability \( p_k(R_1) \) of selecting each of the nonbest populations, we have

\[ p_k(R_1) = A_1 + A_2 + \sum_{m=1}^{k-2} \binom{k-2}{m} C_m + \sum_{m=1}^{k-2} \binom{k-2}{m} D_m, \]

where

\[ A_1 = P(X_1 \leq X_k, \ldots, X_{k-1} \leq X_k) \]
\[ A_2 = P(X_1 - X_k \leq d_1, X_1 > X_k, X_j \leq X_k \text{ for } 2 \leq j \leq k-1) \]
\[ C_m = P(\sum_{j=1}^{m+1} (X_j - X_k) \leq d_1, X_j > X_k \text{ for } 1 \leq j \leq m+1; \]
\[ D_m = P(\sum_{j=2}^{m+1} (X_j - X_k) \leq d_1, X_j > X_k \text{ for } 2 \leq j \leq m+1; \]
\[ X_1 \leq X_k; X_j \leq X_k \text{ for } m+2 \leq j \leq k-1). \]

Using the property that the exponential distribution lacks memory, computing each of the above terms and collecting them, the expression for \( p_k(R_1) \) splits into the two cases \( \delta \leq c_1 \) and \( c_1 \leq \delta \) as follows.

\[ p_k(R_1|\delta \leq c_1) = (k-1)^{-1} + (\delta k(k-1))^{-1} c_1 ((1 - \delta/c_1)^{k-1}) \]
\[ + \sum_{m=1}^{k-2} \binom{k-2}{m} \sum_{\nu=0}^{m+1} \frac{(k-2)}{\nu} (-\delta/c_1)^\nu \]
\[ \cdot G_m(d_1(\nu + m + 2)) (\nu + m + 1)^{-1} (\nu + m + 2)^{-m-1} \]
where
\[
\binom{k-2}{\nu, m} = \frac{(k-2)!}{(\nu! \cdot m! \cdot (k-2-\nu-m)!)}.
\]

Also,
\[
p_k(R_1 | c_1 < \delta) = \frac{(k-1)^{-1} - (\delta k(k-1))^{-1} c_1}{k-2 \frac{k-2-m}{c_1}}
+ \sum_{m=1}^{k-2} \sum_{\nu=0}^{m-1} \binom{k-2}{\nu, m} (-1)^{\nu} \left(\frac{\delta}{c_1}\right)^{\nu+m+1}
\cdot \left( G_m(d_1(\nu+m+2) - G_m((d_1 + \delta n \delta)(\nu+m+2)) \right)
\cdot (\nu+m+1)^{-1}(\nu+m+2)^{-m-1} + G_m(d_1 + \delta n \delta)(\nu+m+1)^{-1}
\cdot c_1(d_1 + \delta n \delta)^m[m! \cdot \delta(\nu+m+2)]^{-1}\}
\]

As regards $R$,
\[P(CS|R) = \left[1 - (1 - \delta)^k\right] / k\delta d\]
\[p_k(R | \delta \leq c) = \left[1 - \left[1 - (1 - \delta)^k\right] / \delta\right] c(k-1)^{-1}\]
\[p_k(R | c \leq \delta) = \{k\delta - 1 + (1-c)^{-1} [(k-1)c - k\delta]\} [\delta c(k-1)]^{-1}\]

Using the above expressions with $k = 10$, we obtain Table VI, which gives these probabilities for the slippage configuration and for selected $P^*$. The table indicates that unless $\delta$ is small, $R_1$ is preferable to $R$ with respect to $P_{10}$. Also, $P(CS|R_1) \geq P(CS|R)$ seems to hold for all $\delta$.

6. EXTENSIONS AND GENERALIZATIONS

6.1 The Case when Both $a_i$ and $b_i$ are Unknown

If both the endpoints of the intervals are unknown, then the reasoning of previous sections would yield the same rules with $Y_i$ ($i = 1, \ldots, k$) replaced by $W_i$, the sample range from $\pi_i$. Theorem 2.1, Lemma 2.2 and the corresponding results for $R$ would also hold for these rules. It may be noted that when the
### TABLE VI

Performances of $R_1$ (Upper Entry) and $R$ (Lower Entry) for the Slippage Configuration with $\delta = 0.75^j$

<table>
<thead>
<tr>
<th>j</th>
<th>$\delta$</th>
<th>$P^{(CS)}$</th>
<th>$P_{10}$</th>
<th>$P^{(CS)}$</th>
<th>$P_{10}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.7500</td>
<td>0.87</td>
<td>0.73</td>
<td>0.98</td>
<td>0.94</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.80</td>
<td>0.74</td>
<td>0.96</td>
<td>0.95</td>
</tr>
<tr>
<td>2</td>
<td>0.5625</td>
<td>0.93</td>
<td>0.71</td>
<td>0.99</td>
<td>0.94</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.85</td>
<td>0.73</td>
<td>0.97</td>
<td>0.94</td>
</tr>
<tr>
<td>3</td>
<td>0.4219</td>
<td>0.97</td>
<td>0.69</td>
<td>1.00</td>
<td>0.93</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.88</td>
<td>0.72</td>
<td>0.98</td>
<td>0.94</td>
</tr>
<tr>
<td>5</td>
<td>0.2373</td>
<td>0.99</td>
<td>0.64</td>
<td>1.00</td>
<td>0.91</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.93</td>
<td>0.68</td>
<td>0.99</td>
<td>0.93</td>
</tr>
<tr>
<td>7</td>
<td>0.1335</td>
<td>1.00</td>
<td>0.59</td>
<td>1.00</td>
<td>0.89</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.96</td>
<td>0.60</td>
<td>0.99</td>
<td>0.92</td>
</tr>
<tr>
<td>9</td>
<td>0.0751</td>
<td>1.00</td>
<td>0.54</td>
<td>1.00</td>
<td>0.86</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.96</td>
<td>0.47</td>
<td>1.00</td>
<td>0.88</td>
</tr>
<tr>
<td>12</td>
<td>0.0317</td>
<td>1.00</td>
<td>0.46</td>
<td>1.00</td>
<td>0.81</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.99</td>
<td>0.22</td>
<td>1.00</td>
<td>0.79</td>
</tr>
<tr>
<td>15</td>
<td>0.0134</td>
<td>1.00</td>
<td>0.39</td>
<td>1.00</td>
<td>0.75</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1.00</td>
<td>0.10</td>
<td>1.00</td>
<td>0.55</td>
</tr>
<tr>
<td>19</td>
<td>0.0042</td>
<td>1.00</td>
<td>0.29</td>
<td>1.00</td>
<td>0.66</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1.00</td>
<td>0.03</td>
<td>1.00</td>
<td>0.18</td>
</tr>
<tr>
<td>23</td>
<td>0.0013</td>
<td>1.00</td>
<td>0.20</td>
<td>1.00</td>
<td>0.56</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1.00</td>
<td>0.01</td>
<td>1.00</td>
<td>0.06</td>
</tr>
</tbody>
</table>

Sample sizes are equal, the rule $R$ for the present case reduces to that given by McDonald (1976). For the rule $R_1$ in the present case, determination of the constants required to satisfy the $P^*$-condition would be difficult. In conclusion, it may be remarked that McDonald (1978) considers subset selection rules of type $R$ based on quasi-ranges for the present problem.
6.2 Subset Selection for the Population with the Largest Parameter

If selection of $\pi_k$ is of interest, then the approach of Section 2 used to derive $R_1$ leads to the following rule $R'$:

Select $\pi_i$ iff $\left(\frac{Y_i}{Y_{(k)}}\right)^n_i \geq c$,

where $Y_{(1)} \leq Y_{(2)} \leq \ldots \leq Y_{(k)}$.

However, the approach of Section 3 leads to the unreasonable rule that (in the case of common sample size) selects only the population corresponding to $Y_{(k)}$ if $\left(\frac{Y_{(k-1)}}{Y_{(k)}}\right)^n < c$, and which selects all the populations otherwise. It may be noted that the result of Theorem 2.1 can be shown to hold for rule $R'$. Also, it can be shown using the technique of the proof of Lemma 2.2 that if $\theta_1 = \ldots = \theta_k$ and $n_1 \leq \ldots \leq n_k$, we have $p_1 \leq \ldots \leq p_k$.

Let it also be noted that rule $R'$ reduces to the rule given by Barr and Rizvi (1966) if the sample sizes are equal.

6.3 Extensions to a Larger Class of Distributions

Following Barr and Rizvi (1966), we may extend the results of the previous sections to the following class of distributions given in Hogg and Craig (1956). Let the variables $Z_{ij}$ ($j = 1, \ldots, n_i; i = 1, \ldots, k$) be independent; the distribution of $Z_{ij}$ having density

$$g(z; \theta_i) = \begin{cases} M(z)Q(\theta_i), & a(\theta_i) < z < b(\theta_i) \\
0, & \text{elsewhere} \end{cases}$$

where $a, b, M$ and $Q$ satisfy the following restrictions:

(i) $M(z)$ is positive and continuous,

(ii) $a'(\theta)$ and $b'(\theta)$ are continuous, and

either

(iii) $a(\theta)$ is constant, $b(\theta)$ strictly monotone (or vice versa) and $\sup a(\theta) = \inf b(\theta)$,
or

(iv) \( a(\theta) \) is strictly monotone decreasing, \( b(\theta) \) strictly
monotone increasing (or vice versa) and \( \sup a(\theta) = \inf b(\theta) \).

The relation

\[
\frac{b(\theta)}{1/Q(\theta)} = \int_a^b M(z)dz
\]

shows that \( 1/Q(\theta) \) is strictly monotone, and also reveals whether it is decreasing or increasing.

As noted in Barr and Rizvi (1966), the distribution of

\( 1/Q(V_i) \), where \( V_i \) is the maximum likelihood estimator of \( \theta_i \)
and also complete and sufficient for \( \theta_i \), is given by the distribution of the largest item of a random sample of size \( n_i \) from
the uniform distribution on \( (0, 1/Q(\theta_i)) \).

Therefore, we may replace the variables \( Y_i \) by \( 1/Q(V_i) \) in
the rules above and proceed exactly as given there, using the
same tables. However, which one of \( R_1, R \) or \( R' \) is derived
depends on whether each of the functions \( a(\theta) \) and \( b(\theta) \) is
strictly increasing (+), strictly decreasing (−) or constant (−).
For each of the cases compatible with the given restrictions,
Table VII gives the rules derived.

7. APPENDIX

We now derive expressions for the probabilities

\[ s_i = P(S = \pi_i) \]

and

\[ s_{ij} = P(S = \{\pi_i, \pi_j\}) \]

referred to in Section 4. In what follows, assume that \( Z_1, Z_2 \) and \( Z_3 \) are independent, each with uniform distribution on the interval \((0, 1)\).
TABLE VII

The Types of Rules Derived

<table>
<thead>
<tr>
<th>Selection for the smallest $\theta_i$</th>
<th>Selection for the largest $\theta_i$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$b(\theta)$</td>
<td>$\uparrow$</td>
</tr>
<tr>
<td>$a(\theta)$</td>
<td>$R'$</td>
</tr>
<tr>
<td>$R_1$</td>
<td>$R_1$</td>
</tr>
</tbody>
</table>

Procedure $R_1$

$s_1 = \mathbb{P}(c_1 \delta_2 Z_2 > \delta_1 Z_1, c_1 Z_3 > \delta_1 Z_1)$

$$s_1 = \int_0^1 \mathbb{P}(c_1 \delta_1 Z_2 > \delta_1 x, c_1 Z_3 > \delta_1 x) dx$$

$$s_1 = \begin{cases} 
(1/2 - \delta_1/6) \delta_2 c_1 / \delta_1 & \text{if } c_1 \leq \delta_1 / \delta_2 \\
1 - \delta_1 / 2 \delta_2 c_1 - \delta_1 / 2 c_1 + \delta_1^2 / 3 \delta_2 c_1^2 & \text{if } \delta_1 / \delta_2 \leq c_1
\end{cases}$$

Similarly,

$s_2 = (1/2 - \delta_1/6) \delta_1 c_1 / \delta_2,$

$s_3 = \delta_1 c_1 / 2 - \delta_1^2 c_1 / 6 \delta_2.$
Now,

\[ s_{12} = s'_{12} + s''_{12}, \text{ where} \]

\[ s'_{12} = P(c_1 \delta_2 z_2 < \delta_1 z_1 < \delta_2 z_2, \delta_1 \delta_2 z_1 z_2 < c_1 z_3^2) \]

\[ = \int P(c_1 \delta_2 x < \delta_1 z_1 < \delta_2 x, \delta_1 \delta_2 z_1 x < c_1 z_3^2) dx \]

\[ = \int \int P(c_1 \delta_2 x/\delta_1 < z_1 < \delta_2 x/\delta_1, z_1 < c_1 y^2/\delta_1 \delta_2 x) dy dx \]

\[ x=0 \ y=0 \]

and where \( s''_{12} \) is obtained by interchanging \( \delta_1 z_1 \) and \( \delta_2 z_2 \) in \( s'_{12} \). Remembering the restriction \( \delta_1 \leq \delta_2 \), the above integral may be obtained by straightforward (but lengthy) computations.

Doing this and collecting the terms, we obtain

\[ s_{12} = \begin{cases} 
D_1 & \text{if } c_1 \leq \delta_1 \delta_2 \\
D_2 & \text{if } \delta_1 \delta_2 < c_1 \leq \delta_1/\delta_2 \\
D_3 & \text{if } \delta_1/\delta_2 < c_1 
\end{cases} \]

where

\[ D_1 = c_1 \{5/96_1 \delta_2 + [\ln(\delta_1 \delta_2/c_1)]/3\delta_1 \delta_2 - \delta_2/2\delta_1 - \delta_1/2\delta_2 + 2\delta_2^2/96_1 + 2\delta_1^2/\delta_2 \}, \]

\[ D_2 = 1 - 4(\delta_1 \delta_2/c_1)^{1/2} + c_1 \{- \delta_2/2\delta_1 - \delta_1/2\delta_2 + 2\delta_2^2/96_1 + 2\delta_1^2/\delta_2 \}, \]

\[ D_3 = \delta_1/2\delta_2 c_1 - 2\delta_2^2/96_1 c_1^2 + c_1 \{- \delta_1/2\delta_2 + 2\delta_2^2/96_2 \}. \]
In the same way as above, we also obtain the following expressions.

\[ s_{13} = \begin{cases} T_1 & \text{if } c_1 < \delta_1/\delta_2 \\ T_2 & \text{if } \delta_1/\delta_2 < c_1 \end{cases} \]

where

\[ T_1 = c_1 \left\{ (\delta_2^2/3\delta_1) \ln(\delta_1/\delta_2 c_1) + 5\delta_2^2/18\delta_1 - \delta_1/2 + 2\delta_1^2/9\delta_2 \right\}, \]
\[ T_2 = \delta_1/2c_1 - 2\delta_1^2/9\delta_2 c_1 + c_1 \left\{ -\delta_1/2 + 2\delta_1^2/9\delta_2 \right\}. \]

Also,

\[ s_{23} = - (\delta_1^2 c_1/3\delta_2) (\ln c_1). \]

The probability \( s_{123} \) is obtained by subtraction of all the above probabilities from unity.

**Procedure R**

The expressions for \( s_1 \), \( s_2 \) and \( s_3 \) are exactly the same as those for \( R_1 \), except for the fact that \( c_1 \) is replaced by \( c \). Now,

\[ s_{12} = s_{12}' + s_{12}'' \],

where

\[ s_{12}' = P(\delta_1 Z_1 < \delta_2 Z_2 < \delta_1 Z_1/c, \ z_3 > \delta_1 Z_1/c) \]
\[ \quad = \int P(\delta_1 x/\delta_2 < Z_2 < \delta_1 x/\delta_2 c) P(\delta_1 x/c) \, dx \]

and where \( s_{12}'' \) is obtained by interchanging \( \delta_1 Z_1 \) and \( \delta_2 Z_2 \).

We thus obtain

\[ s_{12} = \begin{cases} F_1 & \text{if } c \leq \delta_1 \\ F_2 & \text{if } \delta_1 \leq c \leq \delta_1/\delta_2 \\ F_3 & \text{if } \delta_1/\delta_2 \leq c \end{cases} \]
where

\[ F_1 = \left(\frac{c}{2\delta_1}\right)\left(1 - \delta_2 + \delta_2^2/3 - c/3\delta_2\right) \]
\[ + \left(\frac{c}{2\delta_2}\right)\left(1 - \delta_1 + \delta_1^2/3 - c/3\delta_1\right), \]
\[ F_2 = 1 - \left(\frac{c\delta_1}{2\delta_2}\right)(1 - \delta_1/3) - \left(\frac{c\delta_2}{2\delta_1}\right)(1 - \delta_2/3) \]
\[ - \delta_1/2c + \delta_1^2/6\delta_2c, \]
\[ F_3 = \left(\frac{\delta_1}{2c\delta_2}\right)(1 + \delta_1/3 - 2\delta_1/3c) - \left(\frac{c\delta_1}{2\delta_2}\right)(1 - \delta_1/3). \]

Similarly,

\[ s_{13} = \begin{cases} H_1 & \text{if } c < \delta_1/\delta_2 \\ H_2 & \text{if } \delta_1/\delta_2 < c \end{cases} \]

where

\[ H_1 = \left(\frac{c\delta_2^2}{3\delta_1}\right)(1/2 - c) - \left(\frac{c\delta_1}{2}\right)(1 - \delta_1/3\delta_2) + c\delta_2/2 \]
\[ H_2 = \left(\frac{\delta_1^2}{3c\delta_2}\right)(1/2 - 1/c) - \left(\frac{c\delta_1}{2}\right)(1 - \delta_1/3\delta_2) + \delta_1/2c. \]

Also,

\[ s_{23} = (1 - c)c\delta_1^2/3\delta_2. \]

The probability \( s_{123} \) is obtained by subtraction from unity.

**ACKNOWLEDGEMENT**

The author is grateful to professor Gunnar Kulldorff for considerable encouragement and valuable comments during the course of this work. The author is also grateful to docent Georg Lindgren for some criticism.

**BIBLIOGRAPHY**


