DESIGNING EXPERIMENTS FOR SELECTING A NORMAL \(^1\) POPULATION WITH A LARGE MEAN AND A SMALL VARIANCE

by

Thomas J. Santner
Cornell University
Ithaca, New York

Ajit C. Tamhane
Northwestern University
Evanston, Illinois

\(^1\) This research was performed while the second author was visiting Cornell University in 1982-83 on a sabbatical leave with partial support from the U.S. Army Research Office-Durham Contract DAAG-29-81-K0168.
Key Words: Indifference-zone approach, Two-stage procedures, Heterogeneous variances.
DESIGNING EXPERIMENTS FOR
SELECTING A NORMAL POPULATION WITH A LARGE
MEAN AND A SMALL VARIANCE\(^1\)

Thomas J. Santner
Cornell University,
Ithaca, New York

Ajit C. Tamhane
Northwestern University
Evanston, Illinois

ABSTRACT

Consider \( p > 2 \) normal populations with the \( i \)th one having mean \( \mu_i \) and variance \( \sigma_i^2 (1 \leq i \leq p) \). All the parameters are unknown. Let \( \mu[1] \leq \cdots \leq \mu[p] \) and \( \sigma[1]^2 \leq \cdots \leq \sigma[p]^2 \) be the ordered means and variances, respectively. For specified integers \( a \) and \( b \) \((1 \leq a, b \leq p)\), consider sets \( A = \{i: \mu_i \geq \mu_{[p-a+1]}\} \), \( B = \{i: \sigma_i^2 \leq \sigma_{[b]}^2\} \), and \( C = A \cap B \). The experimenter's goal is to select the population associated with \( \max\{\mu_i: i \in C\} \) if \( C \) is nonempty and select no population if \( C \) is empty.

It is desired to design the experiment to guarantee a prespecified level of the probability of a correct selection \( (P(C)) \) for \( \mu = (\mu_1, \mu_2, \ldots, \mu_p) \) and \( \sigma^2 = (\sigma_1^2, \sigma_2^2, \ldots, \sigma_p^2) \) lying in a suitably defined "preference-zone."

We propose different two-stage procedures for the cases: (I) \( a + b > p \) when \( C \) cannot be empty, and (II): \( a + b \leq p \). The infimums of the \( P(C) \) of the two procedures are studied. Computer programs are described for determining the design constants for the procedures based on these infimums. Simulation results are given to assess the conservatism of certain bounds employed in deriving the infimums.

\(^1\)This research was performed while the second author was visiting Cornell University in 1982-83 on a sabbatical leave with partial support from the U.S. Army Research Office-Durham Contract DAAG-29-81-K0168.

Key Words: Indifference-zone approach, Two-stage procedures, Heterogeneous variances.
1. INTRODUCTION AND SUMMARY

Let $\Pi_1, \Pi_2, \ldots, \Pi_p$ be $p \geq 2$ normal populations where observations $X_{ij}$ from $\Pi_i$ are independently distributed as $N(\mu_i, \sigma_i^2)$ ($i = 1, 2, \ldots, p$; $j = 1, 2, \ldots$). All the means $\mu_i$ and variances $\sigma_i^2$ are unknown. For this setup Dudewicz and Dalal (1975) and Rinott (1978) studied the problem of selecting the population having the largest mean. However, in many practical situations an experimenter will not be satisfied in selecting the population with the largest mean if it has (nearly) the largest variance. For example, suppose that a manufacturing process is to be selected from $p \geq 2$ possible alternatives. It may be desired to select a process that is highly productive (large mean daily output) as well as highly reliable (small day-to-day variation in the output). If no such process exists then it may be desired to select none.

Motivated by the above consideration, the present paper proposes a new selection goal that is appropriate when both means and variances are important. In §2 we give a mathematical formulation of this new selection goal and an associated probability requirement based on the indifference-zone approach of Bechhofer (1954). This formulation requires the experimenter to specify two integers $a$ and $b$ ($1 \leq a, b \leq p$) such that any $\Pi_i$ with mean $\mu_i$ among the $a$ largest ones and variance $\sigma_i^2$ among the $b$ smallest ones is regarded as a "good" population. The goal then is to select the population with the largest mean from the set of "good" populations if that set is nonempty and to select no population if it is empty. In §3 we propose different two-stage selection procedures for the two cases: (I) $a + b > p$, in which case the set of "good" populations cannot be empty, and (II) $a + b \leq p$. The probabilities of a correct selection ($P(CS)$) associated with these two-stage procedures are derived in §4 and their inifimums over the appropriate region ("preference-zone") of the parameter space are studied. Based on the expressions for these inifimums §5 gives computer programs for calculating the constants required to implement the procedures. The conservatism of certain bounds employed in deriving the inifimums on the $P(CS)$ is also studied in this section by Monte Carlo simulation. Some concluding remarks are made in §6.
2. FORMULATION OF THE PROBLEM

Denote \( \mu = (\mu_1, \mu_2, \ldots, \mu_p) \), \( \sigma^2 = (\sigma_1^2, \sigma_2^2, \ldots, \sigma_p^2) \), \( \omega = (\mu, \sigma^2) \) and the parameter space of \( \omega \) by \( \Omega \). Let \( \mu[1] \leq \mu[2] \leq \cdots \leq \mu[p] \) and \( \sigma^2[1] \leq \sigma^2[2] \leq \cdots \leq \sigma^2[p] \) be the ordered means and variances, respectively. For specified integers \( a \) and \( b \) (\( 1 \leq a, b \leq p \)) and \( \omega \in \Omega \) define the index sets

\[
A = A(\mu; a) = \{i : \mu_i \geq \mu[p-a+1]\}
\]

and

\[
B = B(\sigma^2; b) = \{i : \sigma_i^2 \leq \sigma^2[b]\}.
\]

Let the index set of "good" populations be denoted by

\[
C = C(\omega; a, b) = A(\mu, a) \cap B(\sigma^2, b)
\]

and let \( c = \operatorname{card} (C) \); clearly \( \max(a+b-p, 0) \leq c \leq \min(a, b) \).

The experimenter's goal is to select a population associated with \( \max(\mu_i : i \in C) \) if \( C \) is nonempty and select no population if \( C \) is empty. If a selection procedure achieves this goal then a correct selection (CS) is said to have been made.

Remark 2.1. If \( a + b > p \) then \( C \) cannot be empty and \( \max(\mu_i : i \in C) = \max(\mu_i : i \in \beta) \). In this case the above goal reduces to selecting the population with the largest mean among the ones having the \( b \) smallest variances. Note that this goal is independent of the value of \( a \) as long as \( a + b > p \). This special goal was considered by Handa and Jain (1982).

In the special subcase \( b = p \) of \( a + b > p \), the goal reduces to selecting the population with the largest mean \( \mu[p] \) which was considered by Dudewicz and Dalal (1975) and Rinott (1978). In the subcase \( (a, b) = (p, 1) \) of \( a + b > p \), the goal reduces to selecting the population with the smallest variance \( \sigma^2[1] \) which was considered by Bechhofer and Sobel (1954).

To formulate the probability requirement for the above selection goal we adopt the indifference-zone approach of Bechhofer (1954). In this formulation we assume that the experimenter can specify three constants \( \{\delta, \theta, \gamma\} \) where \( \delta > 0 \), \( \theta > 1 \) and \( 1/p < \gamma < 1 \). The first two constants (together with specified integers \( a \) and \( b \)) define the following preference-zone:
\[ \Omega^* = \bigcap_{i=0}^{p} \Omega^* \]

where
\[ \Omega^*_0 = \{ \omega \in \Omega : C = \phi, \sigma^2_{[b+1]}/\sigma^2_{[b]} \geq \theta, \mu_{[p-a+1]} - \mu_{[p-a]} \geq \delta \} \]

and for \( 1 \leq i \leq p, \)
\[ \Omega^*_i = \{ \omega \in \Omega : i \in C, \sigma^2_{[b+1]}/\sigma^2_{[b]} \geq \theta, \mu_i \geq \mu_j + \delta \forall j \in C, j \neq i \]
\[ \& \forall j \notin A \} . \]

The indifference-zone is \( \Omega = \Omega^* . \)

Remark 2.2. If \( a = p \) then the last condition in the definition of \( \Omega^*_0 \) is vacuous. If \( b = p \) then the second condition in the definition of \( \Omega^*_i (0 \leq i \leq p) \) is vacuous. Also note that if \( c > 1 \) (resp., \( c = 1 \)) then for \( i = 1, \ldots, p \) the last condition in the definition of \( \Omega^*_i \) is equivalent to \( \mu_i \geq \mu_j + \delta \forall j \in C, j \neq i \) (resp., \( \mu_i \geq \mu_j + \delta \forall j \notin A \)).

Our probability requirement is
\[ \inf_{\Omega^*_i} \frac{P(C \cap S)}{P(S)} \geq \gamma . \]  \hfill (2.1)

An equivalent way of stating (2.1) which will be used in the sequel is as follows: Let \( d_i \) denote the decision to select the population \( \Pi_i (1 \leq i \leq p) \) and \( d_0 \) denote the decision to select no population. Then the probability requirement is
\[ \inf_{\Omega^*_i} \frac{P(d_i)}{P(S)} \geq \gamma \quad (0 \leq i \leq p). \]  \hfill (2.2)

3. TWO-STAGE SELECTION PROCEDURES \( P_I \) AND \( P_{II} \)

3.1 Description of the Procedures

Since \( \sigma^2 \) is assumed to be unknown and the parameter of primary interest for selection purposes is the mean of a population, it follows from the arguments of Dantzig (1940) and Dudewicz (1971) that except for the case \( a = p, b = 1 \), no single-stage procedure can guarantee (2.2).

(For \( a = p, b = 1 \), when the goal is to select the population associated with \( \sigma^2_{[1]} \), a single-stage procedure can be used, e.g., the Bechhofer-Sobel (1954) procedure.)

We now propose two-stage procedures \( P_I \) and \( P_{II} \) for Case I: \( a + b > p \) and Case II: \( a + b \leq p \), respectively.
Case I \((a + b > p)\): Procedure \(P_I\)

**Stage 1:** For \(i = 1, 2, \ldots, p\), take a random sample of size \(n_i \geq 2\) (fixed in advance) from \(\Pi_i\) and calculate the sample variances

\[
S^2_i = \left[ \sum_{j=1}^{n_i} X^2_{ij} - \left( \sum_{j=1}^{n_i} X_{ij} \right)^2 / n_i \right] / (n_i - 1).
\]

Let \(S^2_{[1]} \leq S^2_{[2]} \leq \ldots \leq S^2_{[k]}\) be the ordered sample variances,

\[
\hat{B} = \{i: S^2_i \leq S^2_{[b]}\},
\]

and

\[
N_i = \max(n_i, \langle(S_i h/\delta)^2 \rangle), \quad i \in \hat{B}
\]

where \(\langle x \rangle\) denotes the smallest integer \(\geq x\) and \(h > 0\) is a predetermined constant. The choice of \(h\) is discussed in Remark 4.3.

**Stage 2:** For \(i \in \hat{B}\) take an additional random sample of size \((N_i - n_i)\) from \(\Pi_i\) and calculate the overall sample mean \(\overline{X}_i = \sum_{j=1}^{N_i} X_{ij} / N_i\). Select the population yielding \(\max(\overline{X}_i: i \in \hat{B})\).

Case II \((a + b \leq p)\): Procedure \(P_{II}\)

**Stage 1:** The first stage is the same as that of \(P_I\) except that \(N_i\) must be calculated by (3.2) for all \(\Pi_i\) \((1 \leq i \leq p)\).

**Stage 2:** For \(i = 1, 2, \ldots, p\) take an additional random sample of size \((N_i - n_i)\) from \(\Pi_i\) and calculate the overall sample mean \(\overline{X}_i\). Let \(\overline{X}_{[1]} \leq \overline{X}_{[2]} \leq \ldots \leq \overline{X}_{[p]}\) be the ordered sample means,

\[
\hat{A} = \{i: \overline{X}_i \geq \overline{X}_{[p-a+1]}\},
\]

\(\hat{B}\) be as in (3.1) and \(\hat{C} = \hat{A} \cap \hat{B}\). Select no population if \(\hat{C} = \emptyset\); otherwise, select the population yielding \(\max(\overline{X}_i: i \in \hat{C})\).

For convenience, we shall hereafter assume that \(n_i = n\) for \(1 \leq i \leq p\). Thus each \(S^2_i\) is based on the same degrees of freedom \(\nu = n - 1\).

### 3.2 Some Remarks on the Procedures

(i) There are two design constants, \(n\) and \(h\), in \(P_I\) and \(P_{II}\) which must be chosen to guarantee (2.1) for given \(p\) and specified \((a, b, \delta, \theta, \gamma)\). (For \(P_I\) the design constants do not depend on \(a\).) For fixed \(n\),
h must be chosen so that the infimum of $P_{\omega}(CS)$ over $\omega^*$ is $\geq \gamma$. It will be shown that (see Remarks 4.2 and 4.3) that such a finite $h$ exists iff $n > n_0$ where $n_0$ is a positive integer which depends on $p, b, \theta$ and $\gamma$.

(ii) It follows from Remark 2.1 that in $P_{I}$ we do not need to estimate sets $A$ and $C = A \cap B$ while in $P_{II}$ we do. Thus in Case I it suffices to select the population yielding the largest $\overline{X}_i$ among the ones in the estimated set $\hat{B}$. However, in Case II we must estimate set $C$ to determine whether to make decision $d_0$. Thus it is necessary to sample from all populations in the second stage.

(iii) Ideally one would like to use the data from both the stages to estimate the $\sigma^2_i$ and the set $B$. However, this poses the same technical difficulties that were present in Stein's original (1945) problem.

4. PROBABILITIES OF A CORRECT SELECTION OF $P_{I}$ AND $P_{II}$

4.1 Case I (a + b > p): Procedure $P_{I}$

Since $C \neq \phi$ in this case, it suffices to consider the infimum of $P_{\omega}(d_p|P_{I})$ over $\omega^*$ (say) in order to guarantee (2.2). Our result is given in the following theorem.

**Theorem 4.1.** Let $\Phi(*)$ and $F_{\chi}(*)$ denote the cdf's of a $N(0,1)$ and a $\chi^2$ random variable (rv), respectively, where $v = n - 1$. Let

$$G(v,w) = G(v,w|h,v,\theta) = \int_0^\infty \Phi\left(\frac{h}{\sqrt{v+v/y}}\right) dF_{\chi}(y).$$

Then for given $p$ and specified $\{a, b, \delta, \theta\}$ where $b < p$ and $a + b < p$,

$$\inf_{\omega^*} P_{\omega}(d_p|P_{I}) \geq (p-b) \int_0^{\omega_h} \int_0^{b-1} d\Phi_{\chi}(v)$$

$$x(1 - F_{\chi}(w))^{p-b-1} dF_{\chi}(w).$$

**Proof.** Fix $\omega \in \omega^*$ and thus $A = A(\mu; a), B = B(\sigma^2; b)$ and $C = A \cap B$ with $p \in C$. Consider

$$P_{\omega}(d_p|P_{I}) = \sum_{\mathcal{Q}} P_{\omega}(\overline{X}_p > \overline{X}_i \forall i \in Q, i \neq p; S_i^2 < S_j^2 \forall i \in Q, j \neq Q)$$

where the sum is over all subsets $\mathcal{Q}$ of $\{1, 2, \ldots, p\}, p \in \mathcal{Q}$ with card($\mathcal{Q}$) = $b$. A lower bound on (4.3) is given by the term for $\mathcal{Q} = B$, viz.,
\[ P \{ \bar{X}_p > \bar{X}_i \; \forall i \in B, \; i \neq p; \; S_i^2 < S_j^2 \; \forall i \in B, \; j \neq B \}. \] (4.4)

This bound is sharp for \( \omega \in \Omega_p^* \); if \( \overline{A} \cap \overline{B} = \phi \) (where \( \overline{A} \) and \( \overline{B} \) denote the complements of \( A \) and \( B \), respectively) then we can choose a sequence \( \{\omega\} \) in \( \Omega_p^* \) such that \( \mu_i > +\infty \; \forall i \in A \cap B \) which makes all the terms in (4.3) for \( Q \neq B \) approach zero yielding (4.4). If \( \overline{A} \cap \overline{B} \neq \phi \) then we can choose the sequence such that, in addition, \( \sigma_j^2 > +\infty \; \forall j \in \overline{A} \cap \overline{B} \) thus yielding (4.4).

We shall find the infimum of the lower bound (4.4) which we first write as

\[ P \{ Y_{pi} < \frac{\mu_p - \mu_i}{\sqrt{\sigma_p^2 + \sigma_i^2}} \; \forall i \in B, \; i \neq p; \; Z_i < \frac{\sigma_j^2}{\sigma_i^2} \; \forall i \in B, \; j \neq B \}. \] (4.5)

Here

\[ Y_{ij} = \frac{X_j - X_i + \mu_i - \mu_j}{\sqrt{\sigma_i^2 + \sigma_j^2}} \; (i \neq j, \; 1 \leq i, \; j \leq p), \; Z_i = \frac{\nu S_i^2}{\sigma_i^2} \; (1 \leq i \leq p). \] (4.6)

For fixed \( \sigma_i^2 \), (4.5) is increasing in each \( \mu_p - \mu_i \) for \( i \in B, \; i \neq p \). So a lower bound on the infimum of (4.5) with respect to (wrt) \( \mu \) is obtained by noting that for \( i \in B, \; i \neq p \),

\[ \frac{\mu_p - \mu_i}{\sqrt{\sigma_p^2 + \sigma_i^2}} \geq \frac{\delta}{\sqrt{\sigma_p^2 + \sigma_i^2}} \] (by the definition of \( \Omega_p^* \))

\[ \geq \frac{h}{\sqrt{\sigma_p^2 + \sigma_i^2} + \nu S_i^2} \] (by (3.2))

\[ = \frac{h}{\sqrt{\nu Z_p + \nu Z_i}} \] (by (4.6)).

Now note that conditional on the \( \{Z_i: 1 \leq i \leq p\} \), the \( \{Y_{pi}: i \in B, \; i \neq p\} \) are multivariate normal with zero means, unit variances and nonnegative correlations which depend on the unknown variances. Therefore we apply the Slepian (1962) inequality to obtain a further lower bound on (4.5) by regarding the \( Y_{pi} \) as conditionally uncorrelated.
N(0,1) rv's. Finally, the infimum of this last bound wrt $\sigma_i^2$ is attained when $\sigma_j^2/\sigma_i^2 = 0 \ \forall \ i \in B, j \not\in B$. An integral expression for this infimum is obtained by conditioning on $\min_{j \neq B} Z_j = w$, $Z_p = v$, and $Z_i = u_i$ for $i \in B, i \neq p$ and noting that the $Z_i$ are iid $\chi_v^2$ rv's. This expression is

$$\left(\frac{p-b}{b}\right) \int_0^\infty \left[ \prod_{i \in B} \int_0^{\infty} \Phi \left( \frac{h}{\sqrt{v/u_i} + v/v_i} \right) dF_v(u_i) \right] dF_v(v)$$

$$\times \left[ (1 - F_v(w))^{b-1} dF_v(w) \right]$$

which upon substituting (4.1) gives (4.2). □

**Remark 4.1.** The expression (4.2) is not applicable for $b = p$ because in that case the event $\{S_i^2 \leq S_j^2 \ \forall \ i \in B, j \not\in B\}$ is vacuous. Dropping the event $\{B = B\}$ in (4.4) yields Rinott's (1978) expression (see his equation (12)) for the infimum of the P(CS). For $(a,b) = (p,1)$, (4.2) reduces to Bechhofer and Sobel's (1954) expression (see their equation (19)) for the infimum of the P(CS).

**4.2 Case II $(a + b \leq p)$: Procedure $P_{II}$**

Since $C$ can be empty in this case we must consider both (i) the infimum of $P_\omega(d_0|P_{II})$ over $\Omega_0^\ast$, and (ii) the infimum of $P_\omega(d_p|P_{II})$ over $\Omega_p^\ast$ (say). The result for (i) is given in Theorem 4.2 while that for (ii) is given in Theorem 4.3.

**Theorem 4.2.** Let

$$H(v,w) = H(v,w|v,\theta,h) = \int_{w/\theta}^\infty \Phi \left( \frac{h}{\sqrt{v/y} + v/y} \right) dF_v(y). \quad (4.8)$$

Then for given $p$ and specified $(a,b,\delta,\theta)$ where $a + b \leq p$,

$$\inf_{\Omega_0^\ast} P_\omega(d_0|P_{II}) \geq \int_0^\infty \left[ \int_0^W (H(w,w))^a \left[ \int (H(u,0))^a dF_v(u) \right] dF_v(v) \right] dF_v(w) \quad (4.9)$$

where $d = p - (a + b)$.

**Proof.** Fix $\omega \in \Omega_0^\ast$ and thus $A = A_{\omega}(a)$, $B = B_{\omega}(b)$ and $C = A \cap B = \phi$.

Consider
\[ P_\omega(d_0 | P_{II}) = \sum_{Q, R \sim} P_\omega(\hat{A} = \hat{Q}, \hat{B} = \hat{R}) \]  

(4.10)

where the sum is over all sets \( Q, R \) which are mutually exclusive subsets of \( \{1, 2, \ldots, p\} \) with \( \text{card (Q)} = a \) and \( \text{card (R)} = b \). As in the proof of Theorem 4.1, a sharp lower bound on (4.10) is

\[ P_\omega(\hat{A} = A_0, \hat{B} = B) \]  

(4.11)

This lower bound is sharp because for fixed sets \( A, B \) we can choose a sequence \( \{\omega_\lambda\} \) in \( Q_0^* \) such that \( \mu_i - \sigma^2_i \gamma \forall i \notin A \) and \( \sigma^2_j + \sigma^2_k \gamma \forall j \notin B \) thus yielding (4.11). We now find the infimum of (4.11) which we first write as

\[ P_\omega(Y_{ij} < \frac{\mu_i - \mu_j}{\sqrt{\sigma^2_i N_j + \sigma^2_j N_i}}, \forall i \in A, j \notin A; Z_k < (\sigma^2_k / \sigma^2_k)Z_k \forall k \in B, \forall \lambda \notin B). \]  

(4.12)

Here the \( Y \)'s and \( Z \)'s are as defined in (4.6). Arguing as in Theorem 4.1, a lower bound on (4.12) is

\[ P(Y_{ij} < \frac{h}{\sqrt{v/Z_i + v/Z_j}}, Y_{ik} < \frac{h}{\sqrt{v/Z_i + v/Z_k}}; Z_k < \theta Z_i, Z_k < \theta Z_j) \]  

(4.13)

\[ \forall i \in A, j \notin A \cup B, k \in B \]

where, by using the Slepian inequality, the \( Y \)'s are uncorrelated \( N(0,1) \) rv's conditional on the \( Z \)'s. To write an integral expression for (4.10) we relabel the \( Z \)'s as follows: the \( Z_k = \max(Z_k; k \in B) \) as \( W \), the \( Z_j \) for \( j \notin A \cup B \) as \( (V_1, \ldots, V_d) = V \), the \( Z_k \) for \( k \in B, k \neq k^* \) as \( (U_1, \ldots, U_{b-1}) = U \) and the \( Z_i \) for \( i \in A \) as \( (T_1, \ldots, T_a) = T \). Also let \( T \) denote a generic \( \chi^2_v \) rv. Then by the usual conditioning arguments we obtain

\[ b \int \int \cdots \int \left[ \sum_{w/\theta} W \right] \frac{a}{\Pi} \int \int \phi \left( \frac{h}{\sqrt{v/t_i + v/w}} \right) \int \int \phi \left( \frac{h}{\sqrt{v/t_i + v/w}} \right) \]

\[ x \int \int \phi \left( \frac{h}{\sqrt{v/t_i + v/w}} \right) \]

\[ d \int \int \phi \left( \frac{h}{\sqrt{v/t_i + v/w}} \right) \]

\[ = b E_Y \left\{ E_U \left\{ E_T \left\{ I(T > W/\theta) \phi \left( \frac{h}{\sqrt{v/T + v/W}} \right) \right\} \right\} \right\} \]
\[ x \left[ \prod_{k=1}^{b-1} I(U_k < W) \phi \left( \frac{h}{\sqrt{v/T + v/U_k}} \right) \right] \quad (4.14) \]

where \( I(E) \) denotes the indicator function of event \( E \) and \( E_X \) denotes the expectation with respect to \( X \). Now (4.14) is a \((p-a+1)\)-fold integral which is difficult to evaluate even on a computer. So we seek a computationally feasible lower bound on it. To this end we note in (4.14) that for fixed \( U, V \) and \( W \), the expectation with respect to \( T \) is \( E_T \{ \prod_{i=1}^{3} f_i(T) \} \) where

\[
\begin{align*}
 f_1(T) &= I(T > W/\theta) \phi \left( \frac{h}{\sqrt{v/T + v/W}} \right), \\
 f_2(T) &= \prod_{k=1}^{b} \left\{ I(U_k < W) \phi \left( \frac{h}{\sqrt{v/T + v/U_k}} \right) \right\}, \\
 f_3(T) &= \prod_{j=1}^{d} \left\{ I(V_j > W/\theta) \phi \left( \frac{h}{\sqrt{v/T + v/V_j}} \right) \right\}.
\end{align*}
\]

Each \( f_i(T) \) is increasing in \( T \) and therefore the \( f_i(T) \) are associated rv's (conditioned on \( U, V \) and \( W \)) as a consequence of properties \( P_3 \) and \( P_4 \) stated in Esary, Proschan and Walkup (1967). Therefore

\[ E_T \{ \prod_{i=1}^{3} f_i(T) \} \geq \prod_{i=1}^{3} E_T \{ f_i(T) \}. \quad (4.15) \]

Using this fact and the mutual independence of the \( U_k \) and \( V_j \) we obtain the following lower bound on (4.14):

\[
\begin{align*}
 b \ E_W \left[ E_T \left[ I(T > W/\theta) \phi \left( \frac{h}{\sqrt{v/T + v/W}} \right) \right] \right] \\
 \times \prod_{k=1}^{b-1} \left[ E_U \left\{ E_T \left[ I(U_k < W) \phi \left( \frac{h}{\sqrt{v/T + v/U_k}} \right) \right] \right\}\right] \\
 \times \prod_{j=1}^{d} \left[ E_V \left\{ E_T \left[ I(V_j > W/\theta) \phi \left( \frac{h}{\sqrt{v/T + v/V_j}} \right) \right] \right\}\right].
\end{align*}
\]
\[
\begin{align*}
&= b \left\{ \int_0^\infty \phi \left( \frac{h}{\sqrt{v/t + v/w}} \right) dF_v(t) \right\}^a \left\{ \int_0^\infty \phi \left( \frac{h}{\sqrt{v/t + v/u}} \right) dF_v(t) \right\}^a b^{-1} \\
&\quad \times \left\{ \int_0^\infty \phi \left( \frac{h}{\sqrt{v/t + v/w}} \right) dF_v(t) \right\}^a dF_v(w) \left( \int_0^\infty \phi \left( \frac{h}{\sqrt{v/t + v/w}} \right) dF_v(t) \right)^a dF_v(v) \right\} \right. \\
&\text{which upon substituting (4.8) gives (4.9).} \quad \Box \\
\end{align*}
\]

We next turn to the study of the infimum of \( P_\omega (d_p | P_{II}) \) over \( \omega \in \Omega_p^* \).

**Theorem 4.3.** For given \( p \) and specified \( \{a, b, \delta, \theta\} \) with \( a + b \leq p \),

\[
\inf_{\omega \in \Omega_p^*} P_\omega (d_p | P_{II}) \geq \int_0^\infty \int_0^\infty \left\{ H(v, \theta w) \right\}^{p-M-1} (G(v, w))^{b-1} \\
\quad \times \left\{ (p-M) \phi \left( \frac{h}{\sqrt{v}/V + v/W} \right) \right\} \left\{ 1 - F_v(w) \right\} + (a-m) H(v, \theta w) \right\} dF_v(v) \\
\quad \times (1 - F_v(w))^{a-m-1} dF_v(w) \\
\]  

(4.16)

where \( m = \min(a, b) \), \( M = \max(a, b) \) and \( G(\cdot, \cdot) \) and \( H(\cdot, \cdot) \) are defined in (4.1) and (4.8), respectively.

**Proof.** Fix \( \omega \in \Omega_p^* \) and thus \( A = A(\omega; a), B = B(\omega; b) \) and \( C = A \cap B \) with \( p \in C \). We obtain

\[
P_\omega (d_p | P_{II}) = P_\omega \left\{ p \in \hat{A} \cap \hat{B}; X_p > X_j \forall j \in \hat{C}, j \neq p \right\} \\
\geq P_\omega \left\{ X_p > X_i \forall i \notin A; S_k^2 < S_{\lambda}^2 \forall k \in B, \lambda \notin B \right\} \times \left\{ X_p > X_j \forall j \in \hat{A} \cap B, j \neq p \right\} \\
\geq P_\omega \left\{ X_p > X_i \forall i \notin A; X_p > X_j \forall j \in C, j \neq p \right\} \times S_k^2 < S_{\lambda}^2 \forall k \in B, \lambda \notin B \}
\]

\[
\geq P_\omega \left\{ Y_i < \frac{h}{\sqrt{v/Z_p + v/Z_i}} \forall i \notin A; Y_j < \frac{h}{\sqrt{v/Z_p + v/Z_j}} \forall j \in C, j \neq p; Z_k < \theta Z_{\lambda} \forall k \in B, \lambda \notin B \right\} \\
\]  

(4.17)

where, by using the Slepian inequality, the \( Y \)'s are uncorrelated \( N(0,1) \) rv's conditional on the \( Z \)'s. Now (4.17) is clearly minimized when \( c \)
is maximized, i.e., \( c = m = \min(a, b) \) and \( a + b - c = M = \max(a, b) \). To obtain an integral expression for this probability, condition on 
\[ Z_{\lambda^*} = \min_{\lambda \neq B} Z_{\lambda} = w, \ Z_p = v \quad \text{and} \quad Z_i = u_i \quad \forall i \notin A, \ i \neq \lambda^* \quad \text{and} \quad Z_j = u_j \quad \forall j \in C, \ j \neq p. \]
Then considering the two cases: \( \lambda^* \in A \cap B \)
and \( \lambda^* \in A \cap B \) separately, and combining the two resulting integrals we obtain (4.16). □

Remark 4.2. For fixed \( v \) as \( h \to \infty \) we see that \( G(v, w) \to F_v(\theta w) \),
\( H(v, w) \to 1 - F_v(w/\theta) \) \( \forall \ v, \ w, \) and in particular \( H(v, 0) \to 1. \) By applying
the dominated convergence theorem we then see that as \( h \to \infty \) (4.2), (4.9) and
(4.16) all approach the expression

\[
\left( \frac{p-b}{b} \right) \int_0^\infty \{F_v(\theta w)\}^b \{1 - F_v(w)\}^{p-b-1} \ dF_v(w)
\]

\[ = b \int_0^\infty \{F_v(w)\}^{b-1} \{1 - F_v(w/\theta)\}^{p-b} \ dF_v(w) \]

\[ = P(\hat{B} = B) \] (4.18)

where this probability is calculated under the slippage configuration
\( \sigma_1^2 = \cdots = \sigma^2 = \sigma_{b+1}^2/\theta = \cdots = \sigma_p^2/\theta. \) Thus the procedure is
limited by the accuracy with which it can estimate \( B \) in the first stage.

As \( v \to \infty \), (4.18) converges to unity so that any preassigned \( \gamma \) can be
attained by choosing the first stage sample size sufficiently large.

Furthermore it is easy to see that (4.2), (4.9) and (4.16) all
approach unity as \( h, \nu \to \infty \) so that \( h/\nu \to \infty \). Thus there exist many
\( (h, \nu) \) values which solve the equation obtained by setting each one of the
above expressions equal to the specified value \( \gamma \in (1/p, 1) \). If \( \nu_0 \)
is the smallest integer which makes (4.18) \( \geq \gamma \) then clearly one must have
\( \nu > \nu_0 \) (i.e., \( n > \nu_0 + 1 = n_0 \), say) to assure a finite solution in \( h. \)

Remark 4.3. For given \( p \) and specified \( \{a, b, \delta, \theta, \gamma\} \) the design con-
stant \( h \) to be used in \( P_I \) is obtained by first choosing \( \nu > \nu_0 \) and
then solving the equation obtained by setting (4.2) equal to \( \gamma \).

Similarly the design constant \( h \) to be used in \( P_{II} \) is obtained
by first choosing \( \nu > \nu_0 \), then determining \( h_1 \) (resp., \( h_2 \)) as the
solution to the equation obtained by setting (4.9) (resp., (4.16)) equal
to \( \gamma \), and finally letting \( h = \max(h_1, h_2) \).

Remark 4.4. The question of determining an "optimal" choice of \( (h, \nu) \) is
not addressed in the present article. One could choose \( (h, \nu) \) to
minimize the expected total sample size $E(N) = \sum_{i=1}^{p} \mathcal{E}(N_i)$. We note that it is possible to derive exact expressions for $E(N)$ for $P_I$ and $P_{II}$, although we have not given them here.

5. COMPUTATION OF CONSTANTS AND SIMULATION

RESULTS FOR $P_I$ AND $P_{II}$

5.1 Computation of Constants for $P_I$ and $P_{II}$

The various multiple integrals required in the computation of the $h$ values for $P_I$ and $P_{II}$ were evaluated using Gauss-Laguerre quadrature for the outer integral and Gauss-Legendre quadrature for the inner integrals (except, for the $H(\nu,0)$ function appearing in (4.9) for which the former method was used). In applying the Gauss-Legendre quadrature method if the upper limit of any inner integral was $\infty$, it was truncated at a value $U$ chosen so that the truncation error is less than $10^{-8}$. Note that this error is on the conservative side, i.e., it makes the value of the integral smaller and hence the value of $h$ larger. Thirty-two point quadrature formulae were used for both the methods, the weights and zeroes of the corresponding polynomials being obtained from Stroud and Secrest (1966).

A bisection method was used to solve the equations resulting from setting (4.2), (4.9) and (4.16) equal to $\gamma$, respectively. Programs for evaluating the integrals and solving the equations were written in APL and were run on an IBM PC-1 machine in single precision (8 byte reals). The programs are included in the Appendix.

We studied the performance of $P_I$ and $P_{II}$ using a Monte Carlo simulation. For this purpose we first computed the $h$ values (correct to second decimal place) for two cases which are given in Table 1.

<table>
<thead>
<tr>
<th>Case</th>
<th>Table(s)</th>
<th>p</th>
<th>a</th>
<th>b</th>
<th>$\theta$</th>
<th>$n_0$</th>
<th>$\gamma$</th>
<th>h</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>2</td>
<td>5</td>
<td>3</td>
<td>3</td>
<td>2.0</td>
<td>45</td>
<td>0.90</td>
<td>2.89</td>
</tr>
<tr>
<td>II</td>
<td>3A,3B</td>
<td>5</td>
<td>2</td>
<td>2</td>
<td>2.0</td>
<td>45</td>
<td>0.90</td>
<td>3.59</td>
</tr>
</tbody>
</table>
We note that in Case II, \( h = h_1 = \max(h_1, h_2) \) where \( h_1 \) (resp., \( h_2 \)) is from Theorem 4.2 (resp., Theorem 4.3). It is not generally true, however, that \( h_1 \geq h_2 \).

5.2 Simulation Results

For simulating \( P_I \) (resp., \( P_{II} \)) the constants listed under Case I (resp., Case II) in Table 1 were used; \( \delta \) was set equal to 0.2 in both cases. The performance of \( P_I \) was studied for six different \( \omega \) while that of \( P_{II} \) was studied for twelve different \( \omega \). These parameter configurations are listed in Table 2 for \( P_I \) and in Table 3A (the \( C = \phi \) case) and Table 3B (the \( C \neq \phi \) case) for \( P_{II} \).

Both simulation programs were written in FORTRAN and are included in Santner and Tamhane (1983). The IMSL subroutines GGNML and GGCMS were used to generate a set of \( p \) independent standard normal and \( p \) independent chi-square random variables, respectively, in each run of the simulation experiment. The \( X_i \) and \( S_i^2 \) values for different parameter configurations were obtained from the same basic set of these random variables by making the usual transformations. A total of 5000 runs was used in each simulation experiment. All simulations were performed on Cornell University's IBM 3081 computer.

The results for \( P_I \) are displayed in Table 2. We first note that the smallest \( P(\text{CS}) \) is attained at the configuration \( (\mu_1, c_1) \) which has \( A = \{1,2,5\} \), \( B = \{3,4,5\} \) and \( C = \{5\} \). From the proof of Theorem 4.1 we see that this configuration is close to the least favorable configuration which theoretically minimizes the \( P(\text{CS}) \) of \( P_I \) for \( \omega \) in the preference zone when \( \overline{A} \cap \overline{B} = \phi \), viz., \( \mu_i^{\sim} = + \infty \) \( \forall i \in A \cap B \), \( \mu_{p_i} - \mu_i = \delta \) \( \forall i \in B \) (\( i \neq p \)) and \( \sigma_j^2/\sigma_i^2 = 0 \) \( \forall i \in B \), \( j \neq B \). We further see that this smallest \( P(\text{CS}) \) nearly equals the nominal value of \( \gamma = 0.90 \) indicating that very little probability is lost by the use of the Slepian inequality and the inequality used at the second step of (4.7). The latter inequality is known to be quite sharp. The sharpness of the Slepian inequality here is due to the fact that conditional on the \( S_i^2 \) for \( i \in B = \{3,4,5\} \) there is only a single correlation, that between \( Y_{5,3} \) and \( Y_{5,4} \), which is set equal to zero; it can be shown that the value of this correlation coefficient is approximately 1/2. In general, there are
### TABLE 2

<table>
<thead>
<tr>
<th>( P_1 )</th>
<th>( \mu_1 )</th>
<th>( \mu_2 )</th>
<th>( \sigma_1^2 )</th>
<th>( \sigma_2^2 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.8992</td>
<td>((5,5,0,0,0.2))</td>
<td>((0,0,0,0.2))</td>
<td>((2,2,1,1,1))</td>
<td>((3,3,1,1,1))</td>
</tr>
<tr>
<td>0.9360</td>
<td>((0.0043))</td>
<td>((0.0035))</td>
<td>((0.0020))</td>
<td>((0.0029))</td>
</tr>
<tr>
<td>0.9906</td>
<td>((0.9806))</td>
<td>((0.9544))</td>
<td>((0.9954))</td>
<td>((0.9996))</td>
</tr>
<tr>
<td>0.2078</td>
<td>((20.15, 51.78))</td>
<td>((38.80, 47.87))</td>
<td>((45.35, 51.90))</td>
<td>((45.19, 44.75))</td>
</tr>
<tr>
<td>0.2078</td>
<td>((20.15, 51.78))</td>
<td>((38.80, 47.87))</td>
<td>((45.35, 51.90))</td>
<td>((45.19, 44.75))</td>
</tr>
<tr>
<td>0.2078</td>
<td>((20.15, 51.78))</td>
<td>((38.80, 47.87))</td>
<td>((45.35, 51.90))</td>
<td>((45.19, 44.75))</td>
</tr>
</tbody>
</table>

The estimated standard error of each estimate is given in parentheses below it. All estimates are based on 5000 runs.
### TABLE 3A
Monte Carlo Estimates for $P_{II}$ when $C$ is Empty

<table>
<thead>
<tr>
<th>$\sigma^2$</th>
<th>$\mu$</th>
<th>$P_\omega$(CS)</th>
<th>$E_\omega(N_1)$</th>
<th>$E_\omega(N_2)$</th>
<th>$E_\omega(N_3)$</th>
<th>$E_\omega(N_4)$</th>
<th>$E_\omega(N_5)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\tilde{\sigma}_1^2 = (1,1,2,2,2)$</td>
<td>$\mu_1 = (0,0,0.2,0.2)$</td>
<td>0.9388 (0.0034)</td>
<td>322.68 (68.05)</td>
<td>320.02 (68.45)</td>
<td>647.83 (139.10)</td>
<td>643.48 (137.98)</td>
<td>640.73 (136.30)</td>
</tr>
<tr>
<td></td>
<td>$\mu_2 = (-1,-1,-1,0.2,0.2)$</td>
<td>0.9594 (0.0028)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\tilde{\sigma}_2^2 = (1,1,3,3,3)$</td>
<td>$\mu_1 = (0,0,0.2,0.2)$</td>
<td>0.9760 (0.0022)</td>
<td>322.68 (68.05)</td>
<td>320.02 (68.45)</td>
<td>971.49 (208.65)</td>
<td>964.98 (206.97)</td>
<td>960.83 (204.45)</td>
</tr>
<tr>
<td></td>
<td>$\mu_2 = (-1,-1,-1,0.2,0.2)$</td>
<td>0.9988 (0.0005)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

1/ The estimated standard error of each estimate is given in parentheses below it. All estimates are based on 5000 runs.
<table>
<thead>
<tr>
<th>$\sigma_1^2$</th>
<th>$\mu$</th>
<th>$P_\omega^{(CS)}$</th>
<th>$E_\omega^{(N_1)}$</th>
<th>$E_\omega^{(N_2)}$</th>
<th>$E_\omega^{(N_3)}$</th>
<th>$E_\omega^{(N_4)}$</th>
<th>$E_\omega^{(N_5)}$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$\mu_1 = (5,0,0,0,0.2)$</td>
<td>0.9396 (0.0034)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>$\mu_2 = (1,0,0,0,0.2)$</td>
<td>0.9396 (0.0034)</td>
<td>644.86</td>
<td>639.54</td>
<td>647.83</td>
<td>321.99</td>
<td>320.61</td>
</tr>
<tr>
<td></td>
<td>$\mu_3 = (0,0,0,0,0.2)$</td>
<td>0.9580 (0.0028)</td>
<td>(136.10)</td>
<td>(136.91)</td>
<td>(139.10)</td>
<td>(68.99)</td>
<td>(68.15)</td>
</tr>
<tr>
<td></td>
<td>$\mu_4 = (-1,-1,-1,-1,0.2)$</td>
<td>0.9690 (0.0025)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

$\sigma_2^2 = (2,2,2,1,1)$

| $\mu_1 = (5,0,0,0,0.2)$ | 0.9782 (0.0021)   |                     |                   |                   |                   |                   |                   |
| $\mu_2 = (1,0,0,0,0.2)$ | 0.9782 (0.0021)   | 967.05              | 959.07             | 971.49            | 321.99            | 320.61            |
| $\mu_3 = (0,0,0,0,0.2)$ | 0.9892 (0.0015)   | (204.15)            | (205.35)           | (208.65)          | (68.99)           | (68.15)           |
| $\mu_4 = (-1,-1,-1,-1,0.2)$ | 0.9996 (0.0003)   |                     |                     |                   |                   |                   |                   |

$\sigma_2^2 = (3,3,3,1,1)$

$1/$ The estimated standard error of each estimate is given in parentheses below it. All estimates are based on 5000 runs.
correlations between the $Y_{pi}$ for $i \in B$, $i \neq p$, each approximately 1/2, which are set equal to zero by the Slepian inequality. Thus for larger values of $b$, use of the Slepian inequality will result in a greater loss in probability.

For $P_1$ the estimates in Table 2 show that $\mu_1 < \mu_2 < \mu_3$ in order of increasing favorability of a CS for fixed $\sigma_i^2$, $i = 1, 2$ whereas $\sim_1 < \sim_2$ in order of increasing favorability of a CS for fixed $\mu_j$, $j = 1, 2, 3$. These results can be explained intuitively as follows: For fixed $\sigma_i^2$, the distribution of $\hat{B}$ is identical for all $\mu_i$. However, if $\mu_i$ is viewed as the limit of the vectors $(0, 0, k, 0, 0)$ as $k \rightarrow 0$ then $A = \{1, 2, 5\}$ for both $\mu_1$ and $\mu_2$ but if either $P_1$ or $P_2$ is in $\hat{B}$ then a CS is much more likely under $\mu_2$ than $\mu_1$. Similar heuristics apply to the comparison of $\mu_3$ with $\mu_2$. For fixed $\mu_j$, $j = 1, 2, 3$, the procedure has a higher probability of correctly identifying the true $B$ set under $\sigma_i^2$ then under $\sigma_1^2$ and hence the value of the P(CS) is larger under $\sigma_i^2$.

Finally we note that for $i \in B$ (resp., $j \notin B$), $E(N_i)$ (resp., $E(N_j)$) roughly equals $\sigma_i^2(h/d)^2$ (resp., $n_0$) if $\sigma_j^2/\sigma_i^2$ is sufficiently large $\forall i \in B$, $j \notin B$.

We now turn to the simulation results for $P_{II}$. The configurations considered in Table 3A correspond to $A = \{4, 5\}$, $B = \{1, 2\}$ and $C = \phi$. The estimate of the P(CS) given for each $\omega$ is the proportion of times no population was selected. Here the configuration $(\mu_1, \sigma_1^2)$ is nearly least favorable and the achieved P(CS) in this configuration is significantly higher than the nominal value of $\gamma = 0.90$ indicating that (4.9) gives a quite conservative lower bound on the actual P(CS). The explanation for this again lies in the fact that the lower bound (4.11) is sharp iff $\mu_i - \mu_j \rightarrow +\infty \forall i \in A$, $j \notin A$ and $\sigma_j^2/\sigma_i^2 \rightarrow +\infty \forall i \in B$, $j \notin B$; for any other configuration such as $(\mu_1, \sigma_1^2)$, the actual P(CS) is greater by an amount equal to the terms ignored in arriving at the lower bound (4.11). Some extra loss in P(CS) results by the use of the Slepian inequality and the inequality (4.15). Also note that $\mu_1 < \mu_2$ (resp., $\sigma_1^2 < \sigma_2^2$) in the order of favorability for a CS because for fixed $\sigma_i^2$, $i = 1, 2$ (resp., $\mu_j$, $j = 1, 2$) the probability of correctly
identifying the true A (resp., B) set is higher under $\mu_2$ (resp., $\sigma_2^2$) than under $\mu_1$ (resp., $\sigma_1^2$).

The configurations considered in Table 3B correspond to $A = \{1, 5\}$, $B = \{4, 5\}$ and $C = \{5\}$. For configurations $\overset{\sim}{\mu}_3$ and $\overset{\sim}{\mu}_4$ one can also take $A = \{4, 5\}$ and $C = \{4, 5\}$. Once again, the configuration $(\mu_1, \sigma_1^2)$ is nearly least favorable and the explanation for the excess P(CS) in this configuration over the nominal value $\gamma = 0.90$ is similar to that given in the preceding case. For fixed $\sigma_1^2$ or $\sigma_2^2$, $\overset{\sim}{\mu}_1$ through $\overset{\sim}{\mu}_4$ are successively more favorable configurations. Similarly for fixed $\mu_j (1 \leq j \leq 4)$, $\sigma_2^2$ is more favorable than $\sigma_1^2$.

Finally we turn to the discussion of the estimates of the $E(N_i)$ in Tables 3A and 3B. Since $P_{II}$ samples from all populations in both the stages, each $E(N_i)$ depends on $\sigma_i^2$ but not on $\sigma_j^2$ for $j \neq i$. This is in contrast to the case of $P_I$ where each $E(N_i)$ depends on all the $\sigma_j^2$ because whether a population is sampled in the second stage or not in $P_I$ depends on whether it enters set $\hat{B}$. Since in Tables 3A and 3B, $n_0$ is much smaller than $\sigma_i^2(h/\delta)^2 \forall i$, all the $E(N_i)$ roughly equal $\sigma_i^2(h/\delta)^2$.

6. CONCLUDING REMARKS

In this article we have provided one of many possible formulations of the general problem of selecting the normal population with a large mean and a small variance. Here a large mean (resp., a small variance) is defined relative to the set of means (resp., variances) of all populations at hand. Alternatively, one may wish to define these quantities relative to some known standards. Thus one may specify the values $\mu_0$ and $\sigma_0^2$ and regard any $\mu_i \geq \mu_0$ (resp., $\sigma_i^2 \leq \sigma_0^2$) as a large mean (resp., a small variance). In some situations one may have a control population which is normal with mean $\mu_0$ and variance $\sigma_0^2$ both of which are unknown. In such cases, in addition to sampling from the $p$ contending populations, one can also sample from the control population in two stages to obtain estimates of $\mu_0$ and $\sigma_0^2$ with which the estimates of the $\mu_i$ and $\sigma_i^2 (1 \leq i \leq p)$ can then be compared. The authors are studying such a procedure.
BIBLIOGRAPHY


APPENDIX

A.1 The APL Programs for Computation of $h$

The APL functions PCSTHM41, PCSTHM42 and PCSTHM43 calculate (4.2), (4.9) and (4.16), respectively for given values of the global variables:

- $P = \text{value of } p$
- $A = \text{value of } a$
- $B = \text{value of } b$
- $C = \text{value of } c$
- $\text{THETA} = \text{value of } \theta$
- $N = \text{value of } n_0$
- $H = \text{value of } h$
- $\text{ZLA} = \text{vector of zeroes of Laguerre polynomial of some order}$
- $\text{WLA} = \text{vector of weights corresponding to ZLA}$
- $\text{ZLE} = \text{vector of zeroes of Legendre polynomial of some order (not necessarily the same order as the Laguerre polynomial)}$
- $\text{WLE} = \text{vector of weights corresponding to ZLE}$

Warning: These functions assume $N$ is odd.

Once values of $H$ are known which give probabilities bracketing the desired $\gamma$ value then the function BISECTION can be used to determine the $H$ value which make (4.2), (4.9) or (4.16) equal to $\gamma$. First define the global variables.

- $\text{HI} = \text{a value of } H \text{ giving a } P(\text{CS}) > \gamma$.
- $\text{LO} = \text{a value of } H \text{ giving a } P(\text{CS}) < \gamma$.
- $\text{PSTAR} = \text{desired } \gamma \text{ value}$.

Then invoke the function BISECTION 'PCSNAME' where PCSNAME is PCSTHM41, PCSTHM42 or PCSTHM43. The function currently stops when the calculated $P(\text{CS})$ differs from $\gamma$ by less than $10^{-4}$. The stopping criteria can be changed in line [6].

There are several other notes that are relevant to modifying the functions. The function PHI $X$ calculates the standard normal cdf of a vector argument $X$. PHI currently sets any argument greater than 4.5 equal to 1; this convention can be changed in line [3]. The functions PCSTHM42 and PCSTHM43 calculate an upper bound $U$ so that the upper tail of the chi-square distribution with $(N-1)$ degrees of freedom is less than $10^{-8}$. This tolerance can be altered in line [4] of either PCSTHM42 or PCSTHM43.
RES=PCSTHM4;I;NU;U;NUO2;NUO2M1;NUO2B;CONST;ZLEP1;WRK1;WRK2;TMP;ALF

;BETA;MLWLA;MLULE;FCAC1;FCAC2;FCAC3;SM

[1] NUO2*(NU*NU-1)+2
[2] NUO2M1+NUO2-1
[3] CHIBAR*FBAR(2*ZLA)
[5] SM+ALB
[6] CM+ALB
[7] CONST*THETA*NUO2B
[8] CONST*CONST*:((2*(NUO2B+NU*(1+P-CM)))/(1+NUO2M1)+(P+B-CM))
[9] ZLEP1+ZLE+1
[10] U+MAKEU
[11] START:
[12] ALF=((-2*ZLA)+U)
[13] BETA=(-ALF)+2*U
[14] MLWLA=WLAK(*ZLA*(1+NUO2+NUO2B))*(ALF*(-1+P-CM))*(CHIBAR*(-1+A-1))
[15] TMP=(pZLA)^(11+1)+1
[16] LOOP:=(ALF(1)10)/NEXTI
[17] WRK1=-(ALF1I*ZLE)+BETA(I)
[18] WRK2=ZLEP1+THETA*ZLA[I]
[19] FAC1=-(UW*WRK1+UO2M1)(-WRK1/4)+,MAKE5PHI
[20] FAC3+FAC1*(-1+P-CM)
[21] FAC1+(A-SM)*ALFI*FAC1)*/(2*NU)/(1+NUO2M1)
[22] FAC1+FAC1((P-CM)*(CHBARI)*/MAKE6PHI)
[23] FAC2=(MLWLA+WLE*(ZLEP1+UO2M1)(-WRK2/2)+,MAKE7PHI/(B-1)
[24] TMP=1I+MLX1,MAKE2X*(FAC1*FCAC2*FCAC3)
[25] NEXTI:=(pZLA)^(11+1)+1)/LOOP
[26] RES=CONSTX(/MLWLA*TMP)

MAKE1PHI(I)

RES=MAKE1PHI;TMP;NZ;NZLA;DIM;IND

[1] TMP=H=(((NU=(2*POW)++)+(NU=(2*POW)))X0.5)
[2] NZ=NZLA+NZL
[3] DIM=NZLA,NZL
[4] TMP=PHI(IND*,(NZ*,LNZ))/,TMP
[5] RES=DIPO(IND/TMP)
[6] RES=RES+DIPO/IND*(IND*,(NZ*,(NZ))/,RES
[7] RES=((RES1E-150)*RES)+(RES1E-150)*1E-150

MAKE3PHI(O)

RES=MAKE3PHI;TMP;DIM

[1] DIM=TMP=H=(((NUO2+ZLA)++,NU=(ZLA1I*XLEP1)))X0.5)
[2] RES=DIPOPHI(,TMP)

MAKE2PHI(O)

RES=MAKE2PHI

[1] RES=PHI H=(((NUO2+ZLA1I)*(2*POW))X0.5)
```

```
A.2 The FORTRAN Programs for Simulation

SIM1 and SIM2 are programs to simulate $P_I$ and $P_{II}$, respectively. Both programs require the following as inputs:

- $P =$ value of $p$ (upper limit of 10 but can be raised by increasing the dimensions of the various arrays)
- $A =$ value of $a$
- $B =$ value of $b$
- $\Delta =$ value of $\delta$
- $\Theta =$ value of $\theta$
- $N =$ value of $n$
- $H =$ value of $h$
- $J_{TOT} =$ number of $\omega$ configurations (upper limit of 6 for SIM1 and 12 for SIM2; these limits can be raised by increasing the dimensions of the various arrays)
- $N_{TOT} =$ number of runs
- $\text{MEAN}(I,J) =$ value of $\mu_i$ for the $J$th configuration
- $\text{VAR}(I,J) =$ value of $\sigma_i^2$ for the $J$th configuration ($1 \leq I \leq P$, $1 \leq J \leq J_{TOT}$).

DSEED is a double precision seed required by IMSL subroutines GGYNML and GGCHS. Its initial value is specified in line 22 in SIM1 and in line 23 in SIM2; this value can be changed if necessary. Both the programs assume that $\pi_p$ is the "best" population (when $C$ is nonempty); when inputting the MEAN and VAR values the populations should be labelled accordingly.

As output, SIM1 gives the estimates of the $P(CS)$ (i.e., the probability that $\pi_p$ is selected), the probability that $\pi_p$ enters the set $\hat{B}$, and $E(N_i)$ for $i = 1,\ldots,p$, along with their estimated standard errors. SIM2 gives the estimates of the probability that no population is selected (which is the $P(CS)$ when $C = \phi$), the probability that $\pi_p$ is selected (which is the $P(CS)$ when $C \neq \phi$), and $E(N_i)$ for $i = 1,\ldots,p$, along with their estimated standard errors.
Program SIM1

DIMENSION MEAN(10,6),SIGMA(10,6),VAR(10,6),SVAR(10,6),TVAR(10,6),
1N(10,6),NCS(6),INB(10,6),NSUM(10,6),NSUMSQ(10,6),ASN(10),ASNSD(10)
2,NENT(6),R(10),W(50)
DOUBLE PRECISION DSeed
INTEGER P,A,B
REAL MEAN
READ (5,*) P,A,B,Delta,Theta,,NO,H
READ (5,*) JTOT,NTOT
DO 5 J=1,JTOT
READ (5,*) (MEAN(I,J),I=1,P)
READ (5,*) (VAR(I,J),I=1,P)
NCS(J)=0
NENT(J)=0
DO 10 I=1,P
SIGMA(I,J)=SQRT(VAR(I,J))
NSUM(I,J)=0
NSUMSQ(I,J)=0
10 CONTINUE
5 CONTINUE
G=(H/DELTA)**2
NU=NO-1
DSeed=121146.DO
DO 100 I=1,NTOT
CALL GGNML(DSeed,P,R)
DO 15 I=1,P
CALL GGCHO(DSeed,NU,W,CHISQ)
DO 20 J=1,JTOT
INB(I,J)=I
SVAR(I,J)=(VAR(I,J)*CHISQ)/NU
TVAR(I,J)=SVAR(I,J)
20 CONTINUE
15 CONTINUE
DO 40 J=1,JTOT
40 IFLAG=0
I=1
35 IF (SVAR(I,J).LE.SVAR(I+1,J)) GO TO 30
TEMP=SVAR(I,J)
SVAR(I,J)=SVAR(I+1,J)
SVAR(I+1,J)=TEMP
ITEMP=INB(I,J)
INB(I,J)=INB(I+1,J)
INB(I+1,J)=ITEMP
IFLAG=1
30 I=I+1
IF (I.LT.P) GO TO 35
IF (IFLAG.EQ.1) GO TO 40
DO 45 I=1,B
K=INB(I,J)
IF (K.EQ.P) NENT(J)=NENT(J)+1
Program SIM1
(Continued)

N(K,J)=TVAR(K,J)*6+1
N(K,J)=MAX0(NO,N(K,J))
NTEMP=N(K,J)-NO
NSUM(K,J)=NSUM(K,J)+NTEMP
NSUMSQ(K,J)=NSUMSQ(K,J)+NTEMP**2
TEMP=N(K,J)
XBAR=((SIGMA(K,J)*R(K))/SQRT(TEMP))+MEAN(K,J)
IF (I.NE.1) GO TO 50
KSTAR=K
XMAX=XBAR
GO TO 45
50 IF (XBAR.LT.XMAX) GO TO 45
XMAX=XBAR
KSTAR=K
45 CONTINUE
IF (KSTAR.EQ.P) NCS(J)=NCS(J)+1
25 CONTINUE
100 CONTINUE
WRITE (6,1) P,A,B,DELTA,THETA,NO,H
1 FORMT((3X,'P=",I1,3X,'A=",I1,3X,'B=",I1,3X,'DELTA=",F4.2,
13X,'THETA=",F4.2,3X,'NO=",I2,3X,'H=",F6.3)
XTOT=FLOAT(NTOT)
DO 55 J=1,NTOT
WRITE (6,2)((I,MEAN(I,J)),I=1,P)
2 FORMT((3X,10("MEAN",I1,"=",F4.2,3X))
WRITE (6,3)((I,VAR(I,J)),I=1,P)
3 FORMT((3X,10("VAR",I1,"=",F4.2,3X))
PCS=FLOAT(NCS(J))/XTOT
PENT=FLOAT(NENT(J))/XTOT
PCSSD=SQRT(PCS*(1-PCS)/XTOT)
PENTSD=SQRT(PENT*(1-PENT)/XTOT)
WRITE (6,4)PCS,PCSSD,PENT,PENTSD
4 FORMT((3X,"PCS=",F7.4,3X,"PCSSD=",F7.4,3X,"PENT=",F7.4,3X,
1,"PENTSD=",F7.4)
DO 60 I=1,P
ASN(I)=FLOAT(NSUM(I,J))/XTOT
ASNSD(I)=SQRT((NSUMSQ(I,J)-XTOT*ASN(I)**2)/(XTOT-1))
ASN(I)=ASN(I)+NO
60 CONTINUE
WRITE (6,6)((I,ASN(I),I,ASNSD(I)),I=1,P)
55 CONTINUE
STOP
END
Program SIM2

DIMENSION MEAN(10,12),SIGMA(10,12),VAR(10,12),SVAR(10,12),
IN(12,12),INB(10,12),NSUM(10,12),NSUMSQ(10,12),ASN(12),
2,NCSO(12),R(10),W(50),ASNSD(12)
DIMENSION INA(10,12),INC(10,12),XBAR(10,12),YBAR(10,12)
DOUBLE PRECISION DSEED
INTEGER P,A,B
REAL MEAN
READ (5,*) P,A,B,DELTA,THETA,,NO,H
READ (5,*) JTOT,NTOT
DO 5 J=1, JTOT
READ (5,*) (MEAN(I,J),I=1,P)
READ (5,*) (VAR(I,J),I=1,P)
NCS(J)=0
NCSO(J)=0
DO 10 I=1, P
SIGMA(I,J)=SQRT(VAR(I,J))
NSUM(I,J)=O
NSUMSQ(I,J)=0
10 CONTINUE
5 CONTINUE
G=(H/DELTA)**2
NU=NO-1
DSEED=121146.DO
DO 100 ITOT=1, NTOT
CALL GGNML(DSEED,P,R)
DO 15 I=1, P
CALL G6CHS(DSEED,NU,W,CHISQ)
DO 20 J=1, JTOT
INA(I,J)=I
INB(I,J)=1
SVAR(I,J)=(VAR(I,J)*CHISQ)/NU
N(I,J)=SVAR(I,J)*G+1
N(I,J)=MAX0(NO,N(I,J))
NTEMP=N(I,J)-NO
NSUM(I,J)=NSUM(I,J)+NTEMP
NSUMSQ(I,J)=NSUMSQ(I,J)+NTEMP**2
TEMP=N(I,J)
XBAR(I,J)=SIGMA(I,J)*R(I)/SQRT(TEMP)+MEAN(I,J)
YBAR(I,J)=XBAR(I,J)
20 CONTINUE
15 CONTINUE
DO 25 J=1, JTOT
40 IFLAG=0
I=1
35 IF (SVAR(I,J).LE.SVAR(I+1,J)) GO TO 30
TEMP=SVAR(I,J)
SVAR(I,J)=SVAR(I+1,J)
SVAR(I+1,J)=TEMP
ITEMP=INB(I,J)
INB(I,J)=INB(I+1,J)
INB(I+1,J)=ITEMP
IFLAG=1
Program SIM2
(Continued)

30 I=I+1
   IF (I.LT.P) GO TO 35
   IF (IFLAG.EQ.1) GO TO 40
41 IFLAG=0
   I=1
36 IF (YBAR(I,J).LE.YBAR(I+1,J)) GO TO 31
    TEMP=YBAR(I,J)
    YBAR(I,J)=YBAR(I+1,J)
    YBAR(I+1,J)=TEMP
    ITEMP=INA(I,J)
    INA(I,J)=INA(I+1,J)
    INA(I+1,J)=ITEMP
    IFLAG=1
31 I=I+1
   IF (I.LT.P) GO TO 36
   IF (IFLAG.EQ.1) GO TO 41
   K=INB(I,J)
   IC=0
   DO 45 I1=1,A
     K=INB(P-I1+1,J)
     DO 50 I2=1,B
       IF (INB(I2,J).NE.K) GO TO 50
       IC=IC+1
       INC(IC,J)=K
45 CONTINUE
   CONTINUE
   IF (IC.GT.0) GO TO 55
   NCS0(J)=NCS0(J)+1
   GO TO 25
55 KSTAR=INC(1,J)
   XMAX=XBAR(KSTAR,J)
   IF (IC.EQ.1) GO TO 60
   DO 65 I=2,IC
     K=INC(I,J)
     IF (XMAX.GT.XBAR(K,J)) GO TO 65
     KSTAR=K
     XMAX=XBAR(K,J)
65 CONTINUE
60 IF (KSTAR.EQ.P) NCS(J)=NCS(J)+1
25 CONTINUE
100 CONTINUE
   WRITE (6,1) P,A,B,DELTA,THETA,N0,H
   1 FORMAT (3X,'P=',I1,3X,'A=',I1,3X,'E=',I1,3X,'DELTA=',F4.2,
     13X,'THETA=',F4.2,3X,'N0=',I2,3X,'H=',F6.3)
   XTOT=FLOAT(NTOT)
   DO 70 J=1,JWTOT
      WRITE (6,2) ((I,MEAN(I,J)),I=1,P)
      2 FORMAT (3X,10('MEAN',I1,'=',F5.2,3X))
      WRITE (6,3) ((I,BAR(I,J)),I=1,P)
Program SIM2
(Continued)

3 FORMAT (3X,10('VAR',I1,'=',F4.2,3X))
   PCS=FLOAT(NCS(J))/XTOT
   PCS0=FLOAT(NCS0(J))/XTOT
   PCSSD=SQR((PCS-(1-PCS)/XTOT)
   PCS0SD=SQR((PCS0-(1-PCS0)/XTOT)
   WRITE (6,4) PCS,PCSSD,PCS0,PCS0SD
4 FORMAT (3X,'PCS=',F7.4,3X,'PCSSD=',F7.4,3X,'PCS0=',F7.4,3X,
   1'PCS0SD=',F7.4)
   DO 75 I=1,P
   ASN(I)=FLOAT(NSUM(I,J))/XTOT
   ASNSD(I)=SQR((NSUMSQ(I,J)-XTOT*ASN(I)**2)/(XTOT-1))
   ASN(I)=ASN(I)+NO
75 CONTINUE
   WRITE (6,6) ((I,ASN(I),I,ASNSD(I)),I=1,P)
6 FORMAT (3X,10('EN',I1,'=',F6.2,2X,'SD',I1,'=',F6.2,2X))
70 CONTINUE
STOP
END