330

Selection Procedures for the Means and Variances of Normal Populations When the Sample Sizes are Unequal*

by

Shanti S. Gupta and Deng-Yuan Huang

Purdue University

Department of Statistics
Division of Mathematical Sciences
Mimeograph Series #330

August 1973

*This research was supported by the Office of Naval Research under Contract N00014-67-A-0226-00014 at Purdue University. Reproduction in whole or part is permitted for any purpose of the United States Government.

Selection Procedures for the Means and Variances of Normal Populations When the Sample Sizes are Unequal*

by

Shanti S. Gupta and Deng-Yuan Huang
Purdue University

1. Introduction.

Let $\pi_1,\ldots,\pi_k$ be k independent normal populations with means $\mu_1,\ldots,\mu_k$ and variances $\sigma_1^2,\ldots,\sigma_k^2$, respectively. Our interest is to select a nonempty subset of the k populations containing the best when the populations are ranked in terms of (i) the means μ_i, when $\sigma_i^2 = \sigma^2$, known or unknown, and (ii) the variances σ_i^2, when the μ_i are known or unknown. In most of the earlier work (see, for example, Gupta [4], [7]), it is assumed that the number of observations from each population is the same. Very little work has been done in the case of unequal samples. Sitek [13] proposed a procedure for the normal means; however, his result is shown to be in error by Dudewicz [1]. Recently, Gupta and Huang [8] proposed a procedure which is different from that of Sitek and the one investigated in Section 2 of this paper. Section 3 concerns with a procedure of Gupta and Sobel [11] for selecting the population with the smallest variance based on unequal sample sizes. The exact lower bound for the probability of a correct selection was obtained in [11] only in two special cases. A lower bound is given for the general case in Section 3, and by a similar argument, we discuss a lower bound of the probability of a correct selection of the largest scale

*This research was supported by the Office of Naval Research under Contract N00014-67-A-0226-00014 at Purdue University. Reproduction in whole or part is permitted for any purpose of the United States Government.

parameter for the gamma distribution for a similar rule studied by Gupta [5], with our results being applicable to unequal sample sizes. In Section 4, we propose a rule, which is different from the rule proposed by Gupta and Sobel [10], to select a subset containing all populations better than an unknown control for a common known variance. Sitek [13] has proposed the same type of rule in the case of a common unknown variance. In this section we also discuss and improve the lower bound given by Dunnett [2] for the probability associated with k simultaneous confidence intervals for $\mu_i-\mu_0$, $i=1,2,\dots,k$.

2. Selecting the normal population with the largest mean.

Here we assume that $\sigma_i^2=\sigma^2$, $i=1,\dots,k$. The ordered μ_i's are denoted by $\mu_{[1]} \leq \mu_{[2]} \leq \dots \leq \mu_{[k]}$. It is assumed that there is no prior knowledge of the correct pairing of the ordered and the unordered μ_i's. Let us denote by $\pi_{(i)}$ the population associated with $\mu_{[i]}$, $i=1,\dots,k$. Our goal is to select a non-empty subset of the k populations so as to include the population associated with $\mu_{[k]}$. Defining any such selection as a correct selection, we wish to define a procedure R so that $P(CS|R)$, the probability of a correct selection, is at least a preassigned number $P^*(\frac{1}{k} < p^* < 1)$. We will refer to this requirement as the P*-condition. We will discuss the two cases: (a) σ^2 known and (b) σ^2 unknown.

Case (a): σ^2 known. We assume without any loss of generality that $\sigma^2=1$ and propose the following rule R_1 based on the sample means $\bar{X}_i$, $i=1,\dots,k$.

R_1: Select π_i if and only if

$$(2.1) \qquad \bar{X}_i \geq \max_{1\leq j\leq k} \left(\bar{X}_j - c_1\sqrt{\frac{1}{n_i}+\frac{1}{n_j}}\right),$$

where $c_1=c_1(k,p^*,n_1,\dots,n_k) > 0$ is chosen so as to satisfy the p*-condition.

<u>The expression for $P(CS|R_1)$:</u> Let $\bar{X}_{(i)}$ and $n_{(i)}$ denote the sample mean and the sample size associated with the population $\pi_{(i)}$ with mean $\mu_{[i]}$, $i = 1,2,\ldots, k$. Of course, both $\bar{X}_{(i)}$ and $n_{(i)}$ are unknown. Then

$$P(CS|R_1) = \Pr\{\bar{X}_{(k)} \geq \max_{1\leq j\leq k-1} (\bar{X}_{(j)} - c_1 \sqrt{\frac{1}{n_{(k)}} + \frac{1}{n_{(j)}}})\}$$

$$= \Pr\{(\bar{X}_{(j)} - \bar{X}_{(k)})(\frac{1}{n_{(j)}} + \frac{1}{n_{(k)}})^{-\frac{1}{2}} \leq c_1,\ j = 1,\ldots, k-1\}$$

(2.2)

$$= \Pr\{(\bar{X}_{(j)} - \bar{X}_{(k)} - \mu_{[j]} + \mu_{[k]})(\frac{1}{n_{(j)}} + \frac{1}{n_{(k)}})^{-\frac{1}{2}} \leq c_1 + (\mu_{[k]} - \mu_{[j]})(\frac{1}{n_{(j)}} + \frac{1}{n_{(k)}})^{-\frac{1}{2}},\quad j = 1,\ldots, k-1\}$$

$$= \Pr\{Z_{j,k} \leq c_1 + (\mu_{[k]} - \mu_{[j]})(\frac{1}{n_{(j)}} + \frac{1}{n_{(k)}})^{-\frac{1}{2}},\ j = 1,\ldots, k-1\}.$$

For $\ell = 1,2,\ldots, k$, define

(2.3) $$Z_{r,\ell} = (\bar{X}_{(r)} - \bar{X}_{(\ell)} - \mu_{[r]} + \mu_{[\ell]})(\frac{1}{n_{(r)}} + \frac{1}{n_{(\ell)}})^{-\frac{1}{2}},\ r = 1,\ldots,k;\ r \neq \ell,$$

and

(2.4) $$\rho^{(\ell)}_{r,s} = \rho(Z_{r,\ell}, Z_{s,\ell}) = [(1+ \frac{n_{(\ell)}}{n_{(r)}})(1+ \frac{n_{(\ell)}}{n_{(s)}})]^{-\frac{1}{2}},\ r,s = 1,\ldots, k;\quad r,s \neq \ell;\ r \neq s.$$

Thus $Z_{r,\ell}$, $r \neq \ell$, are standard normal variables with correlation matrix $\{\rho^{(\ell)}_{r,s}\}$. We can write $P(CS|R_1)$ alternatively as

(2.5) $$P(CS|R_1) = \int_{-\infty}^{\infty} \prod_{j=1}^{k-1} \Phi\{\sqrt{\frac{n_{(j)}}{n_{(k)}}}\, y + \delta_{k,j}\sqrt{n_{(j)}} + c_1\sqrt{1+ \frac{n_{(j)}}{n_{(k)}}}\}d\Phi(y),$$

where $\delta_{ij} = \mu_{[i]} - \mu_{[j]}$, $\Phi(\cdot)$ and $\varphi(\cdot)$ denote the cdf and pdf of a standard normal random variable, respectively.

For the evaluation of the infimum of $P(CS|R_1)$ over the parameter space

$$\Omega_1 = \{\underline{\mu} : \underline{\mu} = (\mu_1,\ldots, \mu_k),\ -\infty < \mu_1,\ldots, \mu_k < \infty\},$$

and all possible associations between $(n_1,\ldots, n_k)$ and $(n_{(1)},\ldots, n_{(k)})$, we need the following lemmas. The first one is due to Slepian (See, Gupta [6]) and is stated below without proof.

<u>Lemma 2.1</u>. Let $X_1,\ldots, X_m$ $(Y_1,\ldots, Y_m)$ be standard normal random variables with the correlation matrix $\{\rho_{ij}\}$ $(\{\kappa_{ij}\})$. Let

$\Phi_m(a_1,\ldots, a_m, \{\rho_{ij}\}) = Pr(X_1 < a_1,\ldots, X_m < a_m)$. If $\rho_{ij} \geq \kappa_{ij}$, $i,j=1,2,\ldots,m$, then, for any set of constants $a_1,\ldots, a_m$,

$$\Phi_m(a_1,\ldots, a_m; \{\rho_{ij}\}) \geq \Phi_m(a_1,\ldots, a_m; \{\kappa_{ij}\}). \tag{2.6}$$

Now we state and prove a lemma, which is a direct consequence of the above lemma.

<u>Lemma 2.2</u>. Let $n_1,\ldots, n_k$ be a set of given positive numbers and denote their ordered values by $n_{[1]} \leq \cdots \leq n_{[k]}$. For any ℓ, $1 \leq \ell \leq k$, let

$$\begin{aligned} \kappa_{ij}^{(\ell)} &= [(1+ \frac{n_{[\ell]}}{n_{[i]}})(1+ \frac{n_{[\ell]}}{n_{[j]}})]^{-\frac{1}{2}}, \quad i,j = 1,\ldots,k;\ i,j\neq\ell;\ i\neq j. \\ \kappa_{ii}^{(\ell)} &= 1 \quad , i = 1,2,\ldots, k;\ i \neq \ell. \end{aligned} \tag{2.7}$$

Then, for any set of constants $a_1,\ldots, a_{k-1}$,

$$\begin{aligned} &\Phi_{k-1}(a_1,\ldots, a_{k-1}; \{\kappa_{ij}^{(\ell)}\}) \\ &\geq \Phi_{k-1}(a_1,\ldots, a_{k-1}; \{\kappa_{ij}^{(k)}\}). \end{aligned} \tag{2.8}$$

Proof: The inequality (2.6) follows from Lemma 2.1, if we show that for $\ell < k$ (i) $\kappa_{ij}^{(\ell)} \geq \kappa_{ij}^{(k)}$, $i,j \neq \ell$, $i,j \neq k$; (ii) $\kappa_{k,j}^{(\ell)} \geq \kappa_{\ell,j}^{(k)}$, $j = 1,\ldots, k$; $j \neq \ell,k$. It is easily seen that (i) and (ii) are true, because $n_{[\ell]} \leq n_{[k]}$. And we know $\{\kappa_{ij}^{(\ell)}\}$, $1 \leq \ell \leq k$, is positive definite [3].

We now prove the following theorem regarding the infimum of $P(CS|R_1)$.

Theorem 2.1. For the rule R_1 defined in (2.1),

$$\text{(2.9)} \qquad \min_{n_1,n_2,\ldots,n_k} \inf_{\Omega_1} P(CS|R_1) = \int_{-\infty}^{\infty} \prod_{j=1}^{k-1} \Phi\left[\frac{c_1-\alpha_j u}{(1-\alpha_j^2)^{\frac{1}{2}}}\right] d\Phi(u),$$

where

$$\alpha_i = \left(1 + \frac{n_{[k]}}{n_{[i]}}\right)^{-\frac{1}{2}}, \quad i = 1,\ldots, k-1.$$

Proof. For any given association between $(n_1,\ldots, n_k)$ and $(n_{(1)},\ldots, n_{(k)})$, we can see from (2.5) that the infimum of $P(CS|R_1)$ is attained when $\mu_{[1]} =\ldots= \mu_{[k]}$. Thus the infimum we seek in (2.9) is given by

$$\text{(2.10)} \qquad \min_{1\leq i\leq k} \int_{-\infty}^{\infty} \prod_{\substack{j=1 \\ j\neq i}}^{k} \Phi\left\{\sqrt{\frac{n_{(j)}}{n_{(i)}}}\, y + c_1\sqrt{1+\frac{n_{(j)}}{n_{(i)}}}\right\} d\Phi(y) .$$

Using the alternative form in (2.2), this minimum in (2.10) is equal to

$$\begin{aligned}
&\min_{1\leq \ell\leq k} P\{Z_{r,\ell} \leq c,\ r = 1,\ldots, k;\ r \neq \ell\} \\
&= \min_{1\leq \ell\leq k} \Phi_{k-1}(c, c,\ldots, c;\ \{\rho_{r,s}^{(\ell)}\}) \\
&= \min_{1\leq \ell\leq k} \Phi_{k-1}(c, c,\ldots, c;\ \{\kappa_{r,s}^{(\ell)}\}) \\
&= \Phi_{k-1}(c, c,\ldots, c;\ \{\kappa_{r,s}^{(k)}\}), \text{ by Lemma 2.2.} \\
&= \Pr(V_j \leq c,\ i = 1,2,\ldots, k-1),
\end{aligned}$$

where $V_1, \ldots, V_{k-1}$ are standard normal random variables with correlation $\kappa^{(k)}_{r,s} = \alpha_r \alpha_s$. It is well known that $V_1, \ldots, V_{k-1}$ can be generated from k independent standard normal variates $V'_1, \ldots, V'_{k-1}, V$ by the transformation

$$V_j = (1 - \alpha_j^2)^{\frac{1}{2}} V'_j + \alpha_i V$$

and it follows that the minimum we have obtained above is equal to

$$(2.11) \qquad \int_{-\infty}^{\infty} \prod_{j=1}^{k-1} \Phi\left[\frac{c_1 - \alpha_j u}{(1-\alpha_j^2)^{\frac{1}{2}}}\right] d\Phi(u) \ .$$

This completes the proof of the theorem.

Let S denote the size of the subset selected. Then the expected subset size is given by

$$E(S|R_1) = \sum_{i=1}^{k} \Pr\{\text{Selecting the population } \pi_{(i)} | R_1\}$$

$$(2.12) \qquad = \sum_{i=1}^{k} \Pr\{\overline{X}_{(i)} \geq \max_{1 \leq j \leq k} (\overline{X}_{(j)} - c_1 \sqrt{\frac{1}{n_{(i)}} + \frac{1}{n_{(j)}}})\}$$

$$= \sum_{i=1}^{k} \Pr\{\max_{\substack{1 \leq j \leq k \\ j \neq i}} (\overline{X}_{(j)} - \overline{X}_{(i)}) (\frac{1}{n_{(i)}} + \frac{1}{n_{(j)}})^{-\frac{1}{2}} \leq c_1\}.$$

Theorem 2.2. For the rule R_1

$$(2.13) \qquad \sup_{\Omega_1} E(S|R_1) \leq k\, \Phi(c_1) \ .$$

Proof. Since

$$\Pr\{\max_{\substack{1 \leq j \leq k \\ j \neq i}} (\overline{X}_{(j)} - \overline{X}_{(i)}) (\frac{1}{n_{(i)}} + \frac{1}{n_{(j)}})^{-\frac{1}{2}} \leq c_1\}$$

$$\leq \Pr\{(\overline{X}_{(j)} - \overline{X}_{(i)}) (\frac{1}{n_{(i)}} + \frac{1}{n_{(j)}})^{-\frac{1}{2}} \leq c_1\}, \text{ for any } j \neq i,$$

$$\Pr\{\max_{1\leq j\leq k} (\bar{X}_{(j)} - \bar{X}_{(i)})(\frac{1}{n_{(i)}} + \frac{1}{n_{(j)}})^{-\frac{1}{2}} \leq c_1\} \tag{2.14}$$

$$\leq \frac{1}{k-1} \sum_{\substack{j=1\\ j\neq i}}^{k} \Pr\{(\bar{X}_{(j)} - \bar{X}_{(i)})(\frac{1}{n_{(i)}} + \frac{1}{n_{(j)}})^{-\frac{1}{2}} \leq c_1\}$$

$$= \frac{1}{k-1} \sum_{\substack{j=1\\ j\neq i}}^{k} \Phi[c_1 + (\mu_{[i]} - \mu_{[j]})(\frac{1}{n_{(i)}} + \frac{1}{n_{(j)}})^{-\frac{1}{2}}] .$$

Using (2.14) in (2.12), we have

$$E(S|R_1) \leq \frac{1}{k-1} \sum_{i=1}^{k} \sum_{\substack{j=1\\ j\neq i}}^{k} \Phi[c_1 + \delta_{ij}(\frac{1}{n_{(i)}} + \frac{1}{n_{(j)}})^{-\frac{1}{2}}]$$

$$= \frac{1}{k-1} Q \text{ (say)}. \tag{2.15}$$

Now we show that the supremum of Q over Ω_1 is attained when $\mu_{[1]}=\ldots=\mu_{[k]}$. Towards this end, we consider the configuration

$$\mu_{[1]} =\ldots= \mu_{[m]} = \mu \leq \mu_{[m+1]} \leq \cdots \leq \mu_{[k]},\ 1 \leq m \leq k-1, \tag{2.16}$$

and show that Q is nondecreasing in μ, when $\mu_{[m+1]},\ldots,\mu_{[k]}$ are kept fixed.

For the configuration (2.16), Q can be rewritten as

$$Q = \sum_{i=1}^{m} [(m-1)\ \Phi(c_1) + \sum_{j=m+1}^{k} \Phi\{c_1 - \delta_j\ (\frac{1}{n_{(i)}} + \frac{1}{n_{(j)}})^{-\frac{1}{2}}\}]$$

$$+ \sum_{i=m+1}^{k} [\sum_{j=1}^{m} \Phi\{c_1 + \delta_i\ (\frac{1}{n_{(i)}} + \frac{1}{n_{(j)}})^{-\frac{1}{2}}\} \tag{2.17}$$

$$+ \sum_{\substack{j=m+1\\ j\neq i}}^{k} \Phi\{c_1 + \delta_{ij}\ (\frac{1}{n_{(i)}} + \frac{1}{n_{(j)}})^{-\frac{1}{2}}\}],$$

where $\delta_i = \mu_{[i]} - \mu$.

Interchanging the labels i and j in the sum $\sum_{i=1}^{m}$, $\sum_{j=m+1}^{k}$ and then differentiating with respect to μ and grouping the terms, we have

$$\frac{dQ}{d\mu} = \sum_{i=m+1}^{k} \sum_{j=1}^{m} \left(\frac{1}{n_{(i)}} + \frac{1}{n_{(j)}}\right)^{-\frac{1}{2}} \left[\varphi\left\{c_1 - \delta_i\left(\frac{1}{n_{(i)}} + \frac{1}{n_{(j)}}\right)^{-\frac{1}{2}}\right\}\right.$$

$$\left. -\varphi\left\{c_1 + \delta_i \left(\frac{1}{n_{(i)}} + \frac{1}{n_{(j)}}\right)^{-\frac{1}{2}}\right\}\right] \geq 0.$$

Thus, by successive applications of the above result, with $m = 1,2,\ldots, k-1$, we see that the supremum of Q over Ω_1 is attained when $\mu_{[1]} = \ldots = \mu_{[k]}$ and this gives

$$\sup_{\Omega_1} E(S|R_1) \leq \frac{1}{k-1} \cdot (k-1) \cdot k\Phi(c_1) = k\,\Phi(c_1)\ . \tag{2.18}$$

Remark. For $k = 2$, the constant c_1 obtained to satisfy the P*-condition is given by $\Phi(c_1) = P^*$. Thus, in this case, the bound in (2.18) is $k\,P^*$, which is the exact upper bound in the case of equal sample sizes.

Now we discuss the case of unknown common σ^2.

Case (b): σ^2, unknown. Let s_ν^2 denote the usual pooled estimate of σ^2 on ν degrees of freedom. If the μ_i are unknown, $\nu = \sum_{i=1}^{k} (n_i-1)$. In this case, we propose the rule R_2 defined below.

R_2: Select π_i if and only if

$$\overline{X}_i \geq \max_{1\leq j\leq k} \left(\overline{X}_j - c_2\, s_\nu \sqrt{\frac{1}{n_{(i)}} + \frac{1}{n_{(j)}}}\right), \tag{2.19}$$

where $c_2 = c_2(k, P^*, n_1,\ldots, n_k) > 0$ is to be determined so that the P*-condition is satisfied.

$$P(CS|R_2) = \Pr\{(\bar{X}_{(j)} - \bar{X}_{(k)})(\frac{1}{n_{(j)}} + \frac{1}{n_{(k)}})^{-\frac{1}{2}} \leq c_2 s_\nu, \ j = 1,\dots, k-1\}$$

$$= \Pr\{\frac{Z_{j,k}}{\sigma} \leq \frac{c_2 s_\nu}{\sigma} + \frac{\delta_{k,j}}{\sigma} (\frac{1}{n_{(j)}} + \frac{1}{n_{(k)}})^{-\frac{1}{2}}, \ j = 1,\dots, k-1\}$$

$$\geq \Pr\{\frac{Z_{j,k}}{\sigma} \leq c_2 \frac{s_\nu}{\sigma}, \ j = 1,\dots, k-1\}$$

$$= \int_\sigma^\infty \Pr\{Z'_{j,k} \leq c_2 s, \ j = 1,\dots, k-1\} \, d\, Q_\nu(s),$$

where $Z'_{j,k}$ are standard normal random variables with same correlation matrix as the $Z_{j,k}$ defined earlier and $Q_\nu(s)$ denotes the cdf of a $\chi_\nu/\sqrt{\nu}$ variate. Thus

$$\min_{n_1,\dots,n_k} \inf_{\Omega_2} P(CS|R_2) = \int_0^\infty \min_{n_1,\dots,n_k} \Pr\{Z'_{j,k} \leq c_2 s, j = 1,\dots, k-1\} \, d\, Q_\nu(s)$$

$$= \int_0^\infty \int_{-\infty}^\infty \prod_{j=1}^{k-1} \Phi(\frac{c_2 s - \alpha_i u}{\sqrt{1-\alpha_j^2}}) \, d\Phi(u) dQ_\nu(s),$$

by using the results of Case (a). Thus we obtain the following theorem.

<u>Theorem 2.3</u>. For the rule R_2,

(2.20) $$\min_{n_1,\dots,n_k} \inf_{\Omega_2} P(CS|R_2) = \int_0^\infty \int_{-\infty}^\infty \prod_{i=1}^{k-1} \Phi(\frac{c_2 s - \alpha_i u}{\sqrt{1-\alpha_i^2}}) \, d\Phi(u) \, dQ_\nu(s),$$

where $\Omega_2 = \{\underline{\mu} : \underline{\mu} = (\mu_1,\dots, \mu_k, \sigma^2)\}$.

By similar arguments, we can state the following theorem for the expected size.

<u>Theorem 2.4</u>:

(2.21) $$E(S|R_2) \leq k \int_0^\infty \Phi(c_2 x) \, d\, Q_\nu(x).$$

3. Selecting a Subset Containing the Population with the Smallest Variance and Selecting for the Largest Gamma Scale Parameter.

3.1 Selection for normal variance.

Let $\pi_1, \pi_2, \ldots, \pi_k$ denote k given normal populations with unknown variances $\sigma_1^2, \ldots, \sigma_k^2$, respectively, $(\sigma_i > 0,\ i = 1,2,\ldots, k)$, and with all means known or unknown. The ordered variances are denoted by $\sigma_{[1]}^2 \leq \cdots \leq \sigma_{[k]}^2$. Let $s_{(i)}^2$ denote the (unknown) sample variance that is associated with the i-th smallest population variance, $\sigma_{[i]}^2$; let $\nu_{(i)}$ denote the number of degrees of freedom associated with $s_{(i)}^2$. Gupta and Sobel [11] have considered the following rule:

R_3: Select π_i if and only if

$$s_i^2 \leq \frac{1}{c_3} \min_{1 \leq j \leq k} s_j^2, \quad (0 < c_3 \leq 1).$$

For this rule, they have shown that

$$(3.1) \qquad P(CS|R_3) \geq \min_{1 \leq i \leq k} \int_0^\infty \prod_{\substack{j=1 \\ j \neq i}}^{k} \left[1 - G_{\nu_j}\left(\frac{c_3 \nu_j x}{\nu_i}\right)\right] d\, G_{\nu_i}(x),$$

where $G_\nu(x)$ and $g_\nu(x)$ are chi-square cdf and pdf with ν degrees of freedom, respectively.

The minimum on the right hand side of (3.1) has been obtained by Gupta and Sobel in the two special cases (i) $k = 2$, (ii) all ν_i are equal to 2, except one which is assumed to be any even integer. In the following lemmas we obtain two different lower bounds for the minimum on the right hand side of (3.1)

Lemma 3.1.

$$(3.2)\qquad \min_{1\le i\le k} \int_0^\infty \prod_{\substack{j=1\\ j\ne i}}^{k} \left[1-G_{\nu_j}\left(\frac{c_3\nu_j x}{\nu_i}\right)\right] dG_{\nu_i}(x)$$

$$\ge \int_0^\infty \prod_{j=1}^{k-1} \left[1-G_{\nu_{[j]}}\left(\frac{c_3\nu_{[k]}x}{\nu_{[1]}}\right)\right] dG_{\nu_{[k]}}(x).$$

Note that, $\forall i$

$$\prod_{\substack{j=1\\ j\ne i}}^{k}\left[1-G_{\nu_j}\left(\frac{c_3\nu_j x}{\nu_i}\right)\right] \geq \prod_{j=1}^{k-1}\left[1-G_{\nu_j}\left(\frac{c_3\nu_{[k]}x}{\nu_{[1]}}\right)\right]$$

also Gamma family has MLR hence by the fact [12, p.112] we conclude that

$$\int_0^\infty \prod_{\substack{j=1\\ j\ne i}}^{k}\left[1-G_{\nu_j}\left(\frac{c_3\nu_j x}{\nu_i}\right)\right] dG_{\nu_i}(x)$$

$$\geq \int_0^\infty \prod_{j=1}^{k-1}\left(1-G_{\nu_j}\left(\frac{c_3\nu_{[k]}x}{\nu_{[1]}}\right)\right) dG_{\nu_{[k]}}(x)$$

Proof.

$$\min_{1\le i\le k} \int_0^\infty \prod_{\substack{j=1\\ j\ne i}}^{k} \left[1-G_{\nu_j}\left(\frac{c_3\nu_j x}{\nu_i}\right)\right] dG_{\nu_i}(x)$$

$$= \min_{1\le i\le k} \int_0^\infty \prod_{\substack{j=1\\ j\ne i}}^{k} \left[1-G_{\nu_{[j]}}\left(\frac{c_3\nu_{[j]} x}{\nu_{[i]}}\right)\right] dG_{\nu_{[i]}}(x)$$

$$\ge \min_{1\le i\le k} \int_0^\infty \prod_{\substack{j=1\\ j\ne i}}^{k-1} \left[1-G_{\nu_{[j]}}\left(\frac{c_3\nu_{[k]} x}{\nu_{[1]}}\right)\right]\left[1-G_{\nu_{[k]}}\left(\frac{c_3\nu_{[k]} x}{\nu_{[1]}}\right)\right] dG_{\nu_{[i]}}(x)$$

$$\ge \min_{1\le i\le k} \int_0^\infty \prod_{j=1}^{k-1} \left[1-G_{\nu_{[j]}}\left(\frac{c_3\nu_{[k]} x}{\nu_{[1]}}\right)\right] dG_{\nu_{[i]}}(x), \text{ since}$$

$$1-G_{\nu_{[k]}}\left(\frac{c_3\nu_{[k]}x}{\nu_{[1]}}\right) \ge 1-G_{\nu_{[i]}}\left(\frac{c_3\nu_{[k]}x}{\nu_{[1]}}\right),$$

$$(3.3)\qquad = \int_0^\infty \prod_{j=1}^{k-1} \left[1-G_{\nu_{[j]}}\left(\frac{c_3\nu_{[k]} x}{\nu_{[1]}}\right)\right] dG_{\nu_{[k]}}(x), \text{ using the result in}$$

[12, p. 112] with $\varphi(x) = \prod_{j=1}^{k-1} \left[1-G_{\nu_{[j]}}\left(\frac{c_3\nu_{[k]}x}{\nu_{[1]}}\right)\right]$, and the fact that

$G_{\nu_{[i]}}(x) \ge G_{\nu_{[k]}}(x)$.

Lemma 3.2.

$\therefore \forall i, \quad \prod_{j=1, j\neq i}^{k} \left(1-G_{\nu_j}\left(\frac{c_3\nu_j x}{\nu_i}\right)\right) \geq \prod_{j=2}^{k} \left(1-G_{\nu_{[1]}}\left(\frac{c_3\nu_{[j]}x}{\nu_{[1]}}\right)\right)$

(3.4) $$\min_{1\leq i\leq k} \int_0^\infty \prod_{\substack{j=1\\ j\neq i}}^{k} \left[1-G_{\nu_j}\left(\frac{c_3\nu_j x}{\nu_i}\right)\right] dG_{\nu_i}(x)$$

$$\geq \int_0^\infty \prod_{j=2}^{k} \left[1-G_{\nu_{[1]}}\left(\frac{c_3\nu_{[j]}x}{\nu_{[1]}}\right)\right] dG_{\nu_{[k]}}(x) .$$

Proof.

$$\min_{1\leq i\leq k} \int_0^\infty \prod_{\substack{j=1\\ j\neq i}}^{k} \left[1-G_{\nu_j}\left(\frac{c_3\nu_j x}{\nu_i}\right)\right] dG_{\nu_i}(x)$$

$$= \min_{1\leq i\leq k} \int_0^\infty \prod_{\substack{j=1\\ j\neq i}}^{k} \left[1-G_{\nu_{[j]}}\left(\frac{c_3\nu_{[j]}x}{\nu_{[i]}}\right)\right] dG_{\nu_{[i]}}(x)$$

$$\geq \min_{1\leq i\leq k} \int_0^\infty \prod_{\substack{j=2\\ j\neq i}}^{k} \left[1-G_{\nu_{[1]}}\left(\frac{c_3\nu_{[j]}x}{\nu_{[1]}}\right)\right]\left[1-G_{\nu_{[1]}}(c_3x)\right] dG_{\nu_{[i]}}(x)$$

$$\geq \min_{1\leq i\leq k} \int_0^\infty \prod_{j=2}^{k} \left[1-G_{\nu_{[1]}}\left(\frac{c_3\nu_{[j]}x}{\nu_{[1]}}\right)\right] dG_{\nu_{[i]}}(x), \text{ since}$$

$$1-G_{[1]}(c_3x) \geq 1-G_{\nu[1]}\left(\frac{c_3\nu_{[i]}x}{\nu_{[1]}}\right),$$

$$= \int_0^\infty \prod_{j=2}^{k} \left[1-G_{\nu_{[1]}}\left(\frac{c_3\nu_{[j]}x}{\nu_{[1]}}\right) dG_{[k]}, \text{ by the same}\right.$$

reason as in (3.3).

Using Lemmas 3.1 and 3.2 in (3.1), we obtain the following theorem.

Theorem 3.1.

$$P(CS|R_3) \geq \max\{\int_0^\infty \prod_{j=1}^{k-1} [1-G_{\nu_{[j]}}(\frac{c_3\nu_{[k]}x}{\nu_{[1]}})]\, dG_{\nu_{[k]}}(x),\ \int_0^\infty \prod_{j=2}^{k} [1-G_{\nu_{[1]}}(\frac{c_3\nu_{[j]}x}{\nu_{[1]}})]\, dG_{\nu_{[k]}}(x)\}. \tag{3.5}$$

Remark: To compute c_3, we equate the right hand side of the inequality in (3.5) to P* and solve for c_3.

In the following, we obtain an upper bound for $E(S|R_3)$ under the slippage configurations $\Delta\, \sigma^2_{[1]} = \sigma^2_{[2]} = \ldots = \sigma^2_{[k]}$, $\Delta \geq 1$.

Theorem 3.2. Let

$$\Omega_3 = \{\sigma^2_{[i]} = \Delta\, \sigma^2_{[1]}: i = 2,\ldots, k\},\ \Delta \geq 1 .$$

$$E_{\Omega_3}(S|R_3) \leq \int_0^\infty \prod_{j=1}^{k-1} [1-G_{\nu_{[k]}}(\frac{c_3\nu_{[j]}x}{\nu_{[k]}\Delta})]\, dG_{\nu_{[1]}}(x) + (k-1)\int_0^\infty [1-G_{\nu_{[k]}}(\frac{c_3\nu_{[1]}\Delta}{\nu_{[k]}}x)][1-G_{\nu_{[k]}}(\frac{c_3\nu_{[1]}x}{\nu_{[k]}})]^{k-2}\, dG_{\nu_{[k]}}(x). \tag{3.6}$$

Proof.

$$E_{\Omega_3}(S|R_3) = \sum_{i=1}^{k} P\{s^2_{(i)} \leq \frac{1}{c_3} \min_{\substack{1\leq j\leq k \\ j\neq i}} s^2_{(j)} | \Omega_3\}$$

$$= \sum_{i=1}^{k} P\{\frac{\nu_{(j)}s^2_{(j)}}{\sigma^2_{[j]}} \geq \frac{c_3\nu_{(j)}\sigma^2_{[i]}}{\nu_{(i)}\sigma^2_{[j]}} (\frac{\nu_{(i)}s^2_{(i)}}{\sigma^2_{[i]}}),\quad j=1,\ldots, k \ (j\neq i) | \Omega_3\}$$

$$= P\{\frac{\nu_{(j)}s^2_{(j)}}{\sigma^2_{[j]}} \geq \frac{c_3\nu_{(j)}}{\nu_{(1)}\Delta}\left(\frac{\nu_{(1)}s^2_{(1)}}{\sigma^2_{[1]}}\right),\ j=2,3,\ldots,k\}$$

$$+ \sum_{i=2}^{k} P\{\frac{\nu_{(1)}s^2_{(1)}}{\sigma^2_{[1]}} \geq \frac{c_3\nu_{(1)}\Delta}{\nu_{(i)}}\left(\frac{\nu_{(i)}s^2_{(i)}}{\sigma^2_{[i]}}\right),\ \frac{\nu_{(j)}s^2_{(j)}}{\sigma^2_{[j]}} \geq \frac{c_3\nu_{(j)}}{\nu_{(i)}}\left(\frac{\nu_{(i)}s^2_{(i)}}{\sigma^2_{[i]}}\right),\ j=3,\ldots k\}$$

$$= \int_0^\infty \prod_{j=2}^{k} [1-G_{\nu_{(j)}}\left(\frac{c_3\nu_{(j)}}{\nu_{(1)}\Delta}x\right)]\ dG_{\nu_{(1)}}(x)$$

$$+ \sum_{i=2}^{k} \int_0^\infty [1-G_{\nu_{(1)}}\left(\frac{c_3\nu_{(1)}\Delta}{\nu_{(1)}}x\right)] \prod_{j=3}^{k} [1-G_{\nu_{(j)}}\left(\frac{c_3\nu_{(j)}x}{\nu_{(i)}}\right)]\ dG_{\nu_{(i)}}(x)$$

$$\leq \int_0^\infty \prod_{j=1}^{k-1} [1-G_{\nu_{[k]}}\left(\frac{c_3\nu_{[j]}x}{\nu_{[k]}\Delta}\right)]\ dG_{\nu_{[1]}}(x)$$

$$+ (k-1)\int_0^\infty [1-G_{\nu_{[k]}}\left(\frac{c_3\nu_{[1]}\Delta}{\nu_{[k]}}x\right)][1-G_{\nu_{[k]}}\left(\frac{c_3\nu_{[1]}x}{\nu_{[k]}}\right)]^{k-2}\ dG_{\nu_{[k]}}(x),$$

by using a similar argument as in (3.4).

3.2. Selection for the largest gamma scale parameter.

Let $\pi_1, \pi_2, \ldots, \pi_k$ denote k given gamma populations with density functions

$$\frac{1}{\Gamma(\gamma)\theta_i} e^{-\frac{x}{\theta_i}} \left(\frac{x}{\theta_i}\right)^{\gamma-1},\ x > 0,\ \theta_i > 0,\ i=1,\ldots, k,$$

with a common parameter γ (> 0) which is assumed to be known. The ordered scale parameters θ_i are denoted by $\theta_{[1]} \leq \theta_{[2]} \leq \ldots \leq \theta_{[k]}$. Gupta [5] proposed the following rule.

R_4: Retain π_i in the selected subset

$$\text{if and only if } \bar{X}_i \geq c_4 \max_{1\leq j\leq k} \bar{X}_j,$$

where $c_4 = c_4(k, P^*, 2n_1\gamma, \ldots, 2n_k\gamma)$ is a constant with $0 < c_4 \leq 1$ which is determined in advance of experimentation.

Let $\overline{X}_{(i)}$ denote the (unknown) sample mean that is associated with the i-th smallest population parameter $\theta_{[i]}$; let $\nu_{(i)}$ denote twice the value of the other parameter associated with $\overline{X}_{(i)}$.

The following lemma is given in [5].

<u>Lemma 3.4.</u>

(3.7)
$$P(CS|R_4) = \int_0^\infty \prod_{\alpha=1}^{k-1} [G_{\nu_{(\alpha)}}(\frac{x\,\nu_{(\alpha)}\theta_{[k]}}{c_4\nu_{(k)}\theta_{[\alpha]}})]\, dG_{\nu_{(k)}}(x)$$

$$\geq \min_{1\leq i\leq k} \int_0^\infty \prod_{\substack{j=1\\ j\neq i}}^{k} [G_{\nu_j}(\frac{x\,\nu_j}{c_4\nu_i})]\, dG_{\nu_i}(x),$$

where $G_\nu(x)$ and $g_\nu(x)$ are the cumulative distribution function and the density, respectively, of a standardized gamma random variable (i.e. with $\theta = 1$) and with parameter $\frac{\nu}{2}$.

By a similar argument as in Theorem 3.1 and 3.2, we have the following results.

<u>Theorem 3.3.</u>

(3.8) $P(CS|R_4)$

$$\geq \max\{\int_0^\infty \prod_{j=2}^{k} G_{\nu_{[j]}}(\frac{\nu_{[1]}x}{c_4\nu_{[k]}})\, dG_{\nu_{[1]}}(x),\ \int_0^\infty \prod_{j=1}^{k-1} G_{\nu_{[k]}}(\frac{\nu_{[j]}x}{c_4\nu_{[k]}})\, dG_{\nu_{[1]}}(x)\}.$$

Theorem 3.4.

$$\Omega_4 = \{\theta_{[k]} = \delta\, \theta_{[i]},\ i = 1,2,\ldots, k-1\},\ \delta \geq 1,$$

$$(3.9)\qquad E_{\Omega_4}(S|R_4) \leq \int_0^\infty \prod_{j=2}^{k} G_{\nu_{[1]}}\left(\frac{x\, \nu_{[j]}^{\delta}}{c_4 \nu_{[1]}}\right) dG_{\nu_{[k]}}$$

$$+ (k-1) \int_0^\infty G_{\nu_{[1]}}\left(\frac{x\, \nu_{[k]}}{c_4 \nu_{[1]}^{\delta}}\right) \left[G_{\nu_{[j]}}\left(\frac{x\, \nu_{[j]}^{\delta}}{c_4 \nu_{[k]}}\right)\right]^{k-1} dG_{\nu_{[k]}}(x).$$

4. Selection with Respect to a Control or Standard

4.1. Selecting a subset containing all populations better than a control or standard

Let $\pi_o, \pi_1, \ldots, \pi_k$ be k+1 normal (experimental) populations with unknown means $\mu_o, \mu_1, \ldots, \mu_k$, respectively, and let π_o denote the control population with unknown mean μ_o. Assume that all (k+1) populations have a common known variance $\sigma^2 = 1$. Our goal is to select all experimental populations that are better than the control ($\mu_i \geq \mu_o$). Let $\overline{X}_i$ be the sample mean based on n_i independent observations from π_i, $i = 0, \ldots, k$. Then we propose the following rule.

R_5: Retain the population π_i $(i = 1, \ldots, k)$ in the selected subset if and only if

$$\overline{X}_i \geq \overline{X}_o - \sqrt{\frac{1}{n_i} + \frac{1}{n_o}}\, c_5,\quad c_5 > 0$$

Theorem 4.1.

$$(4.1)\qquad P(CS|R_5) \geq \prod_{i=1}^{k} \int_{-\infty}^{\infty} \Phi\left(-\sqrt{\frac{n_i}{n_o}}\, x + \sqrt{1+\frac{n_i}{n_o}}\, c_5\right) d\Phi(x).$$

Proof is simple and is omitted.

4.2. A multiple comparison procedure for comparing several treatments with a control.

Suppose there are available n_o observations on the control, n_1 observations on the first treatment,..., n_k observations on the k-th treatment. Denote those observations by X_{ij} ($i = 0,1,\ldots, k$; $j = 1,2,\ldots, n_i$) and the i-th treatment mean, $\frac{1}{n_i}\sum_{j=1}^{n_i} X_{ij}$, by $\overline{X}_i$. We make the assumptions that the X_{ij} are independent and normally distributed with common unknown variance σ^2 and mean μ_i. We assume also that there is available an estimate s_ν^2 of σ^2, independent of the $\overline{X}_i$, which is based on ν degrees of freedom, $s_\nu^2 = \frac{1}{\nu}\sum_{i=0}^{k}\sum_{j=1}^{n_i}(X_{ij}-\overline{X}_i)^2$, where $\nu = \sum_{i=0}^{k} n_i-(k+1)$.

The problem is to obtain simultaneous confidence limits for each of the differences $\mu_i - \mu_o$ ($i = 1,2,\ldots, k$) such that the joint confidence coefficient, i.e., the probability that all k confidence intervals will contain the corresponding $\mu_i - \mu_o$, is equal to a preassigned value $P^*(0 < P^* < 1)$.

Let $Z_i = \dfrac{\overline{X}_i-\overline{X}_o-(\mu_i-\mu_o)}{\sqrt{\frac{1}{n_i}+\frac{1}{n_o}}}$ and $t_i = \dfrac{Z_i}{s_\nu}$, $i = 1,\ldots, k$. The Z_i/σ are standard normal variables with correlation $\rho_{ij} = [(1+\frac{n_o}{n_i})(1+\frac{n_o}{n_j})]^{-\frac{1}{2}}$, $i,j = 1,\ldots, k$; $i \neq j$. The r. v.'s t_i, $i = 1,\ldots, k$, have the joint multivariate t-distribution. For this problem, Dunnett[2] proposed the following confidence limits:

(a) lower:

$$\bar{X}_i - \bar{X}_o - d_i\, s_\nu \sqrt{\frac{1}{n_i} + \frac{1}{n_o}}$$

(b) upper:

$$\bar{X}_i - \bar{X}_o + d_i'\, s_\nu \sqrt{\frac{1}{n_i} + \frac{1}{n_o}} \qquad (i = 1, 2, \ldots, k).$$

(c) two-sided:

$$\bar{X}_i - \bar{X}_o \pm d_i''\, s_\nu \sqrt{\frac{1}{n_i} + \frac{1}{n_o}}$$

The constants d_i' and d_i'' satisfy

$$P(t_1 < d_1', \ldots, t_k < d_k') = P^* \tag{4.2}$$

and

$$P(|t_1| < d_1'', \ldots, |t_k| < d_k'') = P^*. \tag{4.3}$$

In order to obtain conservative limits, Dunnett used the inequalities,

$$P(t_1 < d_1', \ldots, t_k < d_k') \geq \prod_{i=1}^{k} p(t_i < d_i') \tag{4.4}$$

and

$$P(|t_1| < d_1'', \ldots, |t_k| < d_k'') \geq \prod_{i=1}^{k} p(|t_i| < d_i'') . \tag{4.5}$$

We give below exact expressions for the probabilities in (4.2) and (4.3).

<u>Theorem 4.2.</u>

$$P(t_1 < d_1', \ldots, t_k < d_k') \tag{4.6}$$

$$= \int_0^\infty \int_{-\infty}^\infty \prod_{i=1}^{k} \Phi\left[\frac{d_i' s - \beta_i u}{\sqrt{1-\beta_i^2}}\right] d\Phi(u)\, dQ_\nu(s)$$

and

(4.7) $$P(|t_1| < d_1'',\ldots, |t_k| < d_k'')$$

$$= \int_0^\infty \int_{-\infty}^\infty \prod_{i=1}^k [\Phi(\frac{d_i''s-\beta_i u}{\sqrt{1-\beta_i^2}}) - \Phi(\frac{-d_i''s-\beta_i u}{\sqrt{1-\beta_i^2}})]\, d\Phi(u)\, dQ_\nu(s),$$

where, $\beta_i = \frac{1}{\sqrt{1+\frac{n_o}{n_i}}}$, $i = 1, 2,\ldots, k$,

and $Q_\nu(s)$ denote the cdf of $\chi_\nu/\sqrt{\nu}$.

Proof.

For (4.2):

$$P(t_1 < d_1',\ldots, t_k < d_k')$$

$$= p(\frac{Z_1}{\sigma} < d_1'\frac{s_\nu}{\sigma},\ldots, \frac{Z_k}{\sigma} < d_k'\frac{s_\nu}{\sigma}$$

$$= \int_0^\infty P(\frac{Z_1}{\sigma} < d_1's,\ldots, \frac{Z_k}{\sigma} < d_k's)\, dQ_\nu(s)$$

$$= \int_0^\infty \int_{-\infty}^\infty \prod_{i=1}^k \Phi[\frac{d_i's-\beta_i u}{\sqrt{1-\beta_i^2}}]\, d\Phi(u)\, dQ_\nu(s),$$ the same argument as in (2.11).

Similarly, for (4.3):

$$p(|t_1| < d_1'',\ldots, |t_k| < d_k'')$$

$$= \int_0^\infty P(-d_1''s < \frac{Z_1}{\sigma} < d_1''s,\ldots, -d_k''s < \frac{Z_k}{\sigma} < d_k''s)\, d\, Q_\nu(s)$$

$$= \int_0^\infty \int_{-\infty}^\infty \prod_{i=1}^k [\Phi(\frac{d_i''s-\beta_i u}{\sqrt{1-\beta_i^2}}) - \Phi(\frac{-d_i''s-\beta_i u}{\sqrt{1-\beta_i^2}})]\, d\Phi(u)\, dQ_\nu(s)\ .$$

5. Numerical Values and Examples

5.1. Suppose $n_{[1]} = \ldots = n_{[k-1]} = \alpha n_{[k]}$, i.e.

$$\alpha_i = [1 + \frac{1}{\alpha}]^{-\frac{1}{2}} = (\frac{\alpha}{1+\alpha})^{\frac{1}{2}}, \; i = 1,2,\ldots,k-1.$$

We note that (2.9) can be rewritten as

$$\int_{-\infty}^{\infty} \prod_{i=1}^{k-1} \Phi[\frac{c_1-\alpha_i x}{\sqrt{1-\alpha_i^2}}] d\Phi(x) = \int_{-\infty}^{\infty} \prod_{i=1}^{k-1} \Phi[\frac{c_1+\alpha_i x}{\sqrt{1-\alpha_i^2}}] d\Phi(x)$$

$$= \int_{-\infty}^{\infty} \Phi^{k-1}(\frac{\rho^{\frac{1}{2}} x+c_1}{\sqrt{1-\rho}}) d\Phi(x), \text{ with } c_1 = \rho^{\frac{1}{2}} c, \; \rho = \frac{\alpha}{1+\alpha} .$$

Equating the above integral to P*, Gupta, Nagel and Panchapakesan [9] have solved for c_1 for special values of ρ = 0.100, 0.125, 0.200, 0.250, 0.300, $\frac{1}{3}$, 0.375, 0.400, $\frac{1}{2}$, 0.600, 0.625, $\frac{2}{3}$, 0.700, 0.750, 0.800, 0.875, 0.900, and k = 2(1)11(2)51, and P* = 0.99, 0.975, 0.95, 0.90, 0.75. For example, $\alpha = 0.5$, k=5, P* = 0.90, we have $\rho = \frac{0.5}{1+0.5} = \frac{1}{3}$, then $c_1 = 1.8886$ and for (2.13), $k\Phi(c_1) = 5\Phi(1.8886) = 4.85$.

5.2 When $k = 2\ell$ and $n_{[1]} = \ldots = n_{[\ell]} = \alpha n_{[\ell+1]} = \ldots = \alpha n_{[k]}$.

$$\alpha_i = \begin{cases} \dfrac{1}{1 + \dfrac{n_{[k]}}{n_{[i]}}} = (\dfrac{\alpha}{1+\alpha})^{\frac{1}{2}}, & i = 1,2,\ldots,\ell \\ \dfrac{1}{\sqrt{2}}, & i = \ell+1,\ldots,k , \end{cases}$$

we have

$$\int_{-\infty}^{\infty} \prod_{j=1}^{k-1} \Phi[\frac{c_1+\alpha_i x}{\sqrt{1-\alpha_i^2}}] \; d\Phi(x)$$

$$= \int_{-\infty}^{\infty} \Phi^{\ell-1}(x+\sqrt{2}\, c_1)\Phi^{\ell}(\sqrt{\alpha}x + \sqrt{1+\alpha}c_1) d\Phi(x) .$$

For special values of $k = 4, 6$ and $\alpha = \frac{1}{2}, \frac{1}{4}, \frac{3}{4}, \frac{9}{10}$, the c_1-value is tabulated in Table 1.

5.3. When $k = 3\ell$ and $n_{[1]} = \ldots = n_{[\ell]} = \alpha n_{[\ell+1]} = \ldots = \alpha n_{[2\ell]} = \beta n_{[2\ell+1]} = \ldots = \beta n_{[k]}$.

$$\alpha_i = \begin{cases} \dfrac{\sqrt{\beta}}{\sqrt{1+\beta}} , & i = 1,\ldots,\ell , \\[2ex] \dfrac{1}{\sqrt{1+\alpha/\beta}} , & i = \ell+1,\ldots,2\ell , \\[2ex] \dfrac{1}{\sqrt{2}} , & i = 2\ell+1,\ldots,3\ell . \end{cases}$$

Then

$$\int_{-\infty}^{\infty} \prod_{j=1}^{k-1} \Phi\left[\frac{c_1+\alpha_i x}{\sqrt{1-\alpha_i^2}}\right] d\Phi(x)$$

$$= \int_{-\infty}^{\infty} \Phi^{\ell-1}(x+\sqrt{2}\, c_1)\Phi^{\ell}(\sqrt{\beta}\, x+\sqrt{1+\beta}\, c_1)\Phi^{\ell}\left(\sqrt{\frac{\beta}{\alpha}}\, x+\sqrt{1+\frac{\beta}{\alpha}}\, c_1\right)d\Phi(x) .$$

For special values of $k = 3, 6$, and $\alpha = \frac{1}{2}, \frac{1}{4}$, $\beta = \frac{1}{2}, \frac{3}{4}$, the c_1-value is tabulated in Table 3.

Table 1

c_1-value of rule R_1 for special values of k, α and P*

k	P* \ α	$\frac{1}{4}$	$\frac{1}{2}$	$\frac{3}{4}$	$\frac{9}{10}$
4	0.75	1.266	1.233	1.208	1.196
	0.90	1.783	1.763	1.747	1.739
	0.95	2.100	2.086	2.074	2.068
	0.99	2.717	2.707	2.699	2.695
6	0.75	1.488	1.446	1.415	1.400
	0.90	1.981	1.955	1.935	1.924
	0.95	2.285	2.266	2.251	2.242
	0.99	2.882	2.869	2.858	2.852

The entry is the smallest value c_1 (to 3 decimals of accurracy) satisfying

$$\int_{-\infty}^{\infty} \Phi^{\frac{k}{2}-1}(x + \sqrt{2}\, c_1)\Phi^{\frac{k}{2}}(\sqrt{\alpha}\, x + \sqrt{1+\alpha}\, c_1)d\Phi(x) = P^* .$$

Table 2

c_1-value of rule R_1 for special values of k, α, β and P*

k	α / β / P*	1/4, 1/2	1/4, 3/4	1/2, 3/4
3	0.75	1.046	1.041	1.027
	0.90	1.598	1.595	1.586
	0.95	1.932	1.930	1.924
	0.99	2.573	2.571	2.568
6	0.75	1.431	1.415	1.409
	0.90	1.941	1.929	1.929
	0.95	2.254	2.243	2.246
	0.99	2.859	2.851	2.854

The entry is the smallest value of c_1 (3 decimals of accuracy) satisfying

$$\int_{-\infty}^{\infty} \Phi^{\frac{k}{3}-1}(x+c_1)\Phi^{\frac{k}{3}}(\sqrt{\beta}\, x + \sqrt{1+\beta}\, c_1)\Phi^{\frac{k}{3}}\left(\sqrt{\frac{\beta}{\alpha}}\, x + \sqrt{1+\frac{\beta}{\alpha}}\, c_1\right)d\Phi(x) = P^* .$$

6. Acknowledgement.

The authors acknowledge gratefully the assistance of Professor S. Panchapakesan who read a first draft of this paper and made many valuable suggestions to improve it. The authors also wish to thank Dr. W. T. Huang for some helpful conversations.

References

[1] Dudewicz, E. J. A note on selection procedures with unequal observation numbers. Zastosowania Mathematyki 14 (to appear). 1974, No. 1, P. 39 —

[2] Dunnett, C. W. (1955). A multiple comparison procedure for comparing several treatments with a control. J. Amer. Statist. Assoc. 50, 1096-1121.

[3] Dunnett, C. W. and Sobel, M. (1955). Approximations to the probability integral and certain percentage points of a multivariate analogue of Student's t distribution. Biometrika 42, 258-260.

[4] Gupta, S. S. (1956). On a decision rule for a problem in ranking means. Mimeo Series No. 150, Inst. of Statist., Univ. of North Carolina, Chapel Hill, NC.

[5] Gupta, S. S. (1963). On a selection and ranking procedure for gamma populations. Ann. Inst. Statist. Math. 14, 199-216.

[6] Gupta, S. S. (1963). Probability integrals of multivariate normal and multivariate t. Ann. Math. Statist. 34, 792-828.

[7] Gupta, S. S. (1965). On some multiple decision (selection and ranking) rules. Technometrics 7, 225-245.

[8] Gupta, S. S. and Huang, W. T. (1973). A note on selecting a subset of normal populations with unequal sample sizes. (to appear in Sankhyā).

[9] Gupta, S. S.; Nagel, K. and Panchapakesan, S. (1973). On the order statistics from equally correlated normal random variables. Biometrika 60, 403-413.

[10] Gupta, S. S. and Sobel, M. (1958). On selecting a subset containing all populations better than a standard. Ann. of Math. Statist. 29, 235-244.

[11] Gupta, S. S. and Sobel, M. (1962). On selecting a subset containing the population with the smallest variance. Biometrika 49, 495-507.

[12] Lehmann, E. L. (1959). Testing Statistical Hypothesis. Wiley, New York.

[13] Sitek, M. (1972). Application of the selection procedure R to unequal observation numbers. Zastosowania Mathematyki 12, 355-363.

Abstract:

Let $\pi_1, \ldots, \pi_k$ be k independent normal populations with means $\mu_1, \ldots, \mu_k$ and variances $\sigma_1^2, \ldots, \sigma_k^2$, respectively. Our interest is to select a non-empty subset of the k populations containing the best when the populations are ranked in terms of (i) the means μ_i, when $\sigma_i^2 = \sigma^2$, known or unknown, and (ii) the variance σ_i^2, when the μ_i are known or unknown. Procedures and results are derived for the case when sample sizes are unequal. We also discuss gamma populations with scale parameter, and selection for normal means that are better than control.

Key words:

Selection procedures, normal means and variances, gamma distribution, scale parameters, better than control.

Unclassified
Security Classification

DOCUMENT CONTROL DATA - R&D

(Security classification of title, body of abstract and indexing annotation must be entered when the overall report is classified)

1. ORIGINATING ACTIVITY (Corporate author) Purdue University	2a. REPORT SECURITY CLASSIFICATION Unclassified
	2b. GROUP

3. REPORT TITLE

Selection Procedures for the Means and Variances of Normal Populations Wher the Sample Sizes are Unequal

4. DESCRIPTIVE NOTES (Type of report and inclusive dates)

Technical Report, August 1973

5. AUTHOR(S) (Last name, first name, initial)

Gupta, Shanti S. and Huang, Deng-Yuan

6. REPORT DATE	7a. TOTAL NO. OF PAGES	7b. NO. OF REFS
August 1973	24	13

8a. CONTRACT OR GRANT NO. N00014-67-A-0226-00014	9a. ORIGINATOR'S REPORT NUMBER(S) Mimeo Series #330
b. PROJECT NO.	
c.	9b. OTHER REPORT NO(S) (Any other numbers that may be assigned this report)
d.	

10. AVAILABILITY/LIMITATION NOTICES

Distribution of this document is unlimited.

11. SUPPLEMENTARY NOTES	12. SPONSORING MILITARY ACTIVITY
	Office of Naval Research Washington, DC

13. ABSTRACT

Let $\pi_1,\ldots,\pi_k$ be k independent normal populations with means $\mu_1,\ldots,\mu_k$ and variances $\sigma_1^2,\ldots,\sigma_k^2$, respectively. Our interest is to select a non-empty subset of the k populations containing the best when the populations are ranked in terms of (i) the means μ_i, when $\sigma_i^2 = \sigma^2$, known or unknown, and (ii) the variance σ_i^2, when the μ_i are known or unknown. Procedures and results are derived for the case when sample sizes are unequal. We also discuss gamma populations with scale parameter, and selection for normal means that are better than control.

DD FORM 1 JAN 64 1473

Unclassified
Security Classification