SOME LOCALLY OPTIMAL SUBSET SELECTION RULES FOR COMPARISON WITH A CONTROL*

by

Deng-Yuan Huang, National Taiwan Normal University
S. Panchapakesan, Southern Illinois University
Sheng-Tsaing Tseng, National Taiwan Institute of Technology

Technical Report #82-30

Department of Statistics
Purdue University

March 1983

*This research of the first two authors was supported by the Office of Naval Research Contract N00014-75-C-0455 at Purdue University. Reproduction in whole or in part is permitted for any purpose of the United States Government.

SOME LOCALLY OPTIMAL SUBSET SELECTION RULES FOR COMPARISON WITH A CONTROL*

by

Deng-Yuan Huang, National Taiwan Normal University
S. Panchapakesan, Southern Illinois University
and
Sheng-Tsaing Tseng, National Taiwan Institute of Technology

1. INTRODUCTION

Let $\pi_0, \pi_1, \ldots, \pi_k$ be k + 1 independent populations where π_i has the associated distribution function $F(x, \theta_i)$ and density $f(x, \theta_i)$ with the unknown parameter θ_i belonging to an interval Θ of the real line R. We are generally interested in two types of goals. Goal I is to select a subset (preferably small in size) of the k populations $\pi_1, \ldots, \pi_k$ that will contain the best (suitably defined) among them and Goal II is to select from $\pi_1, \ldots, \pi_k$ (k experimental treatments) those populations, if any, that are better (to be defined) than π_0 which is the control or standard population.

In the recent years, several authors have investigated construction of optimal subset selection rules and also established optimality properties of known selection rules for specific cases. Some of the important papers in these directions are Berger and Gupta [1], Bickel and Yahav [2], Bjørnstad [3], Chernoff and Yahav [4], Goel and Rubin [5], Gupta and Hsiao [6], Gupta and Hsu [7], Gupta and Huang [8,9], Gupta and Kim [12,13], Gupta and Miescke [14], and Miescke [17]. These investigations generally deal with the symmetric case of equal sample

*The research of the first two authors was supported by the Office of Naval Research Contract N00014-75-C-0455 at Purdue University.

sizes. There have been some investigations in the unequal sample sizes case but these are concerned with ad hoc and heuristic procedures and are not generally successful in establishing the least favorable configuration (LFC) for the probability of a correct decision. For many classical procedures in the literature for selecting a subset containing the best, the LFC turns out to be $\theta_1 = \ldots = \theta_k$. This provides the motivation for seeking rules which are optimal in some suitable sense in a neighborhood of every equiparameter point. When we are comparing these populations with π_0, the local optimality is related to the ability to reject the populations that are inferior to π_0 and select those that are superior in a suitable neighborhood of selected parametric configurations.

Though some early investigation of locally optimal subset selection rules based on ranks appeared in Nagel [18], such rules were not investigated further until recently. Some locally optimal subset selection rules based on ranks were derived by Gupta, Huang and Nagel [11] and Huang and Panchapakesan [16]. The first of these papers considered Goal I whereas the latter considered both Goals I and II. Though these rules were based on ranks it was assumed that the functional form of the density is known; the justification for seeking rules based on ranks comes from the usual robustness considerations in that the ranks are insensitive to outliers and there could be possible deviations from the model. Gupta, Huang and Nagel [11] maximized the probability of a correct selection in the neighborhood of any equiparameter point whereas Huang and Panchapakesan [16] used for Goal I the criterion of strong monotonicity in the same type of neighborhood and for Goal II a local optimality condition which reflects the sensitivity of the rule when all but one population are not distinctly superior and the remaining one is close to others but distinctly superior.

The present paper derives a locally optimal subset selection rule for Goal II based on parametric inference type statistics with no assumption of equal sample sizes. In Section 2, we state a local optimality condition and prove the main theorem giving the construction of a locally optimal rule. Section 3 deals with the applications of the general result to the following special cases: (a) normal means comparison - common known variance, (b) normal means comparison - common unknown variance, (c) gamma scale parameters comparison - known (unequal) shape parameters and (d) regression slopes. In all these cases, the locally optimal rule is obtained based on samples of unequal sizes.

Finally, for detailed discussions on optimality and other aspects of multiple decision problems and general decision theoretic approach, the reader is referred to Gupta and Huang [10] and Gupta and Panchapakesan [15].

2. DERIVATION OF THE RULE

Let $\{X_{ij}\}$, $j=1,\ldots,n_i$, denote the random sample from π_i, $i=0,1,\ldots,k$. Our goal is to construct a rule to select all populations that are better than the control. The selection rule will depend upon the observations through the statistics T_{i0}, $i=1,\ldots,k$, where T_{ij} is suitably defined to indicate the difference between π_i and π_j. For fixed $n_0,n_1,\ldots,n_k$, we assume that T_{ij} has a density function $g_{\tau_{ij}}(t_{ij})$ depending on the parameter τ_{ij}. The parameter τ_{ij} is a measure of the 'separation' of π_i from π_j. Also, τ_{ii} is the same for all i and this common value is denoted by τ^*. For example, in the location case, we can take $\tau_{ij} = \theta_i - \theta_j$ giving $\tau^* = 0$. On the other hand, if θ is a scale parameter, we can take $\tau_{ij} = \theta_i/\theta_j$ so that $\tau^* = 1$.

Now, we define population π_i to be superior to the control if $\tau_{i0} > \tau^*$ and inferior otherwise. Let

$$\Omega = \{\underline{\tau} | \underline{\tau} = (\tau_{10}, \ldots, \tau_{k0}),\ \tau_{i0} \in R,\ i=1,\ldots,k\} ,$$

$$\Omega_0 = \{\underline{\tau} | \tau_{10} = \ldots = \tau_{k0} = \tau^*\} ,$$

$$\Omega_i = \{\underline{\tau} | \tau_{i0} > \tau^* = \tau_{j0},\ j \neq i\} .$$

Define $\underline{\delta} = (\delta_1, \ldots, \delta_k)$, where $\delta_i (1 \leq i \leq k)$ is a test function for $H_0: \underline{\tau} \in \Omega_0$ vs. $H_i: \underline{\tau} \in \Omega_i$.

2.1. Optimality Requirement. Let $0 < \gamma_i < 1$ $(1 \leq i \leq k)$ be specified. Define

(2.1) $$S = \{\underline{\delta} | E_{\underline{\tau}^*}(\delta_i) = \gamma_i,\ 1 \leq i \leq k\} .$$

We wish to derive a rule $\underline{\delta} \in S$ which

(2.2) $$\text{maximizes} \sum_{i=1}^{k} \frac{\partial}{\partial \tau_{i0}} E_{\underline{\tau}}(\delta_i) \Big|_{\underline{\tau}=\underline{\tau}^*}$$

among all the rules in S. We note that $\underline{\delta} \in S$ means that the error probabilities are controlled and that the condition (2.2) amounts to maximizing the efficiency in a certain sense of the rule in picking out the superior population in the direction of each component at $\underline{\tau}^* = (\tau^*, \ldots, \tau^*)$.

Let $h_{\underline{\tau}}(\underline{t})$ denote the joint density of $\underline{T} = (T_{10}, \ldots, T_{k0})$ with respect to a σ-finite measure μ. Let $h_{\underline{\tau}^*}(\underline{t})$ denote the density $h_{\underline{\tau}}(\underline{t})$ when $\underline{\tau} = \underline{\tau}^*$ and $h^{(i)}_{\underline{\tau}^*}(\underline{t})$ denote the partial derivative $\frac{\partial}{\partial \tau_{i0}} h_{\underline{\tau}}(\underline{t})$ evaluated at $\underline{\tau} = \underline{\tau}^*$ Finally, we need to assume certain regularity conditions, namely, that $h_{\underline{\tau}}(\underline{t})$ is continuously differentiable with respect to each component of $\underline{\tau}$ and $\frac{\partial}{\partial \tau_{i0}} h_{\underline{\tau}}(\underline{t})$ is integrable. Under these regularity conditions, it is easy to see that

$$\frac{\partial}{\partial \tau_{i0}} E_{\underline{\tau}}(\delta_i) \Big|_{\underline{\tau}=\underline{\tau}^*} = \int_{\mathcal{T}} \delta_i(\underline{t}) h^{(i)}_{\underline{\tau}^*}(\underline{t}) d\mu(\underline{t})$$

where $\mathcal{T}$ is the sample space of $\underline{T} = (T_{10},\dots,T_{k0})$.

Summarizing the above discussion, we are seeking a rule $\underline{\delta}\ (\underline{t})$ such that

(2.3) $$\int_{\mathcal{T}} \delta_i(\underline{t}) h_{\underline{\tau}^*}(\underline{t}) d\mu(\underline{t}) = \gamma_i,\ i=1,\dots,k$$

and it maximizes, among all rules satisfying (2.3), the expression

(2.4) $$\sum_{i=1}^{k} \int_{\mathcal{T}} \delta_i(\underline{t}) h^{(i)}_{\underline{\tau}^*}(\underline{t}) d\mu(\underline{t}) .$$

<u>2.2. A Locally Optimal Rule.</u> We now state and prove the main theorem of this section.

<u>Theorem 2.1.</u> Under all the assumptions stated previously, a rule $\underline{\delta}^0(\underline{t})$ which satisfies (2.3) and maximizes (2.4) among all rules satisfying (2.3) is given by

(2.5) $$\delta_i^0(\underline{t}) = \begin{cases} 1 & > \\ \lambda_i \quad \text{if} \quad h^{(i)}_{\underline{\tau}^*}(\underline{t}) & = \quad c_i h_{\underline{\tau}^*}(\underline{t}) \\ 0 & < \end{cases}$$

where λ_i and c_i such that

(2.6) $$\int_{\mathcal{T}} \delta_i^0(\underline{t}) h_{\underline{\tau}^*}(\underline{t}) d\mu(\underline{t}) = \gamma_i,\ i=1,\dots,k .$$

<u>Proof.</u> The proof is straightforward by noting that for any $\underline{\delta}$ satisfying (2.3) we have

(2.7) $$\sum_{i=1}^{k} \int_{\mathcal{T}} (\delta_i^0(\underline{t}) - \delta_i(\underline{t}))(h^{(i)}_{\underline{\tau}^*}(\underline{t}) - c_i h_{\underline{\tau}^*}(\underline{t})) d\mu(\underline{t}) \geq 0.$$

<u>Remark 2.1</u>. It is easy to see that the rule $\underline{\delta}^0(\underline{t})$ given by (2.5) actually maximizes $\sum_{i=1}^{k} a_i \frac{\partial}{\partial \tau_{io}} E_{\underline{\tau}}(\delta_i)\Big|_{\underline{\tau}=\underline{\tau}^*}$ for any set of positive constants a_i among all rules satisfying (2.3).

3. SPECIAL CASES

In this section, we apply the result of Section 2 to several special cases.

Case A: Normal Means Comparison; Common Known Variance. Here π_i is $N(\theta_i, \sigma_0^2)$, where σ_0^2 is known. Let $\bar{X}_i$ be the sample mean based on n_i independent observations from π_i, $i=0,1,\ldots,k$. Take $T_{i0} = (\bar{x}_i - \bar{x}_0)/\sigma_0$ and $\tau_{i0} = (\theta_i - \theta_0)/\sigma_0$. Of course, $\tau^* = 0$. Then the joint density $h_{\underline{\tau}}(\underline{t})$ of $\underline{T} = (T_{10},\ldots,T_{k0})$ is a multivariate normal distribution with mean vector $\underline{\tau} = (\tau_{10},\ldots,\tau_{k0})$ and covariance matrix $\Lambda = (\lambda_{ij})$ given by

$$\text{(3.1)} \qquad \lambda_{ij} = \begin{cases} \frac{1}{n_0}\left(1 + \frac{1}{a_i}\right) & \text{for } i = j \\ \frac{1}{n_0} & \text{for } i \neq j \end{cases}$$

where $a_i = n_i/n_0$, $i=1,\ldots,k$. Thus

$$\text{(3.2)} \qquad h_{\underline{\tau}}(\underline{t}) = (2\pi)^{-k/2}|\Lambda|^{-1/2}\exp[-\tfrac{1}{2}(\underline{t}-\underline{\tau})'\Lambda^{-1}(\underline{t}-\underline{\tau})].$$

It is easy to verify that $h_{\underline{\tau}^*}^{(i)}(\underline{t})/h_{\underline{\tau}^*}(\underline{t})$

$$= (\Lambda^{-1}\underline{t})_i$$

$$= n_0\left(\frac{a_i}{1+\sum_{\ell} a_\ell}\right)\{t_i(1+\sum_{\ell\neq i} a_\ell) - \sum_{j\neq i} a_j t_j\} ,$$

where $(\Lambda^{-1}\underline{t})_i$ denotes the ith coordinate of the vector $(\Lambda^{-1}\underline{t})$ and in summations involving the a_ℓ such as $\sum_{\ell\neq i} a_\ell$, the subscript ranges from 1 to k subject to any exceptions stated. The locally optimal rule $\underline{\delta}^0$ is now given by

(3.3) $$\delta_i^0(\underline{t}) = \begin{cases} 1 & \text{if} \quad t_i - \dfrac{1}{1+\sum_{\ell\neq i} a_\ell} \sum_{j\neq i} a_j t_j \geq c_i^* \\ 0 & \text{otherwise} \end{cases}$$

where c_i^* is determined by

(3.4) $$P_{\underline{\tau}^*}\left(T_i - \frac{1}{1+\sum_{\ell\neq i} a_\ell} \sum_{j\neq i} a_j T_j \geq c_i^*\right) = \gamma_i \ .$$

Now, using the fact that, when $\underline{\tau} = \underline{\tau}^*(=\underline{0})$, $(\Lambda^{-1}\underline{T})_i$ is normally distributed with mean zero and variance $n_0 a_i(1+\sum_{\ell\neq i} a_\ell)/(1+\Sigma a_\ell)$, it is easy to see that

(3.5) $$c_i^* = \Phi^{-1}(1-\gamma_i)\sqrt{(1+\Sigma a_\ell)/n_0 a_i(1+\sum_{\ell\neq i} a_\ell)}$$

$$= \Phi^{-1}(1-\gamma_i)\sqrt{N/n_i(N-n_i)}$$

where $N = n_0 + n_1 + \ldots + n_k$ and Φ denotes the standard normal distribution function.

Remark 3.1. The individual selection probability $\delta_i^0(\underline{t})$ in (3.3) can also be expressed as

$$\delta_i^0(\underline{t}) = \begin{cases} 1 & \text{if} \quad t_i - \dfrac{1}{N-n_i} \sum_{\substack{j=1 \\ j\neq i}}^{k} n_j t_j \geq c_i \\ 0 & \text{otherwise} \end{cases}$$

In this form, it can be recognized as a weighted average type rule. When the sample sizes are equal, it is the usual average type rule and selects π_i if and only if $t_i - \frac{1}{k} \sum_{j\neq i} t_j \geq c_i^*$.

Case B: Normal Means Comparison; Common Unknown Variance. Unless stated otherwise, the notations of Case A will apply here. Let S_i^2 denote the sample variance (divisor $n_i - 1$) based on the sample from π_i. Then

$S_p^2 = \sum_{i=0}^{k} (n_i-1)S_i^2 / \sum_{i=0}^{k} (n_i-1)$ is the usual pooled estimator of the common unknown variance σ^2 on $\nu = N - k - 1$ degrees of freedom. Define $Y_i = (\bar{X}_i - \bar{X}_0)$ and $T_{i0} = Y_i/S_p$, $i=1,\ldots,k$. Then $\underline{T} = (T_{10},\ldots,T_{k0})$ has a multivariate normal t distribution and its density $h_{\underline{\tau}}(\underline{t})$ can be written in an integral form as follows:

$$(3.6) \qquad h_{\underline{\tau}}(\underline{t}) = A_0 \int_0^\infty e^{-\frac{1}{2}(\sqrt{\frac{w}{\nu}}\,\underline{t}-\underline{\tau})'\Lambda^{-1}(\sqrt{\frac{w}{\nu}}\,\underline{t}-\underline{\tau}) - \frac{w}{2}}\, w^{\frac{\nu+k}{2}-1}\, dw ,$$

where $\underline{\tau} = (\tau_{10},\ldots,\tau_{k0})$, $\tau_{i0} = (\theta_i-\theta_0)/\sigma$, $i=1,\ldots,k$, A_0 is the appropriate constant, and $\Lambda = (\lambda_{ij})$ is the same as in Case A given by (3.1). Now, letting $C_{\nu,k} = \sqrt{2}\,\Gamma(\frac{\nu+k+1}{2})/\Gamma(\frac{\nu+k}{2})$, it is easy to verify that

$$h_{\underline{\tau}^*}^{(i)}(\underline{t})/h_{\underline{\tau}^*}(\underline{t}) = C_{\nu,k}(\Lambda^{-1}\underline{t})_i/\sqrt{\nu+\underline{t}'\Lambda^{-1}\underline{t}}$$

$$= C_{\nu,k}(\Lambda^{-1}\underline{y})_i/\sqrt{\nu s_p^2+\underline{y}'\Lambda^{-1}\underline{y}}$$

$$= C_{\nu,k}\,\frac{n_i(N-n_i)}{N}\,\frac{y_i - \frac{1}{N-n_i}\sum_{\substack{j=1\\j\neq i}}^{k} n_j y_j}{\sqrt{\nu s_p^2+\underline{y}'\Lambda^{-1}\underline{y}}}$$

$$= C_{\nu,k}\,\frac{n_i(N-n_i)}{N}\,\psi_i(\underline{y},s_p^2), \text{ say ,}$$

where $\underline{y} = (y_1,\ldots,y_k)$.
Thus, the locally optimal rule $\underline{\delta}^0$ is given by

$$(3.7) \qquad \delta_i^0(\underline{y},s_p^2) = \begin{cases} 1 & \text{if } \psi_i(\underline{y},s_p^2) \geq c_i^* \\ 0 & \text{otherwise} \end{cases}$$

where c_i^* is determined by

$$(3.8) \qquad P_{\underline{\tau}^*}(\psi_i(\underline{Y},S_p^2) \geq c_i^*) = \gamma_i \ .$$

Since, under $\underline{\tau}=\underline{0}$, $Y_i - \frac{1}{N-n_i}\sum_{j\neq i} n_j Y_j$ has a normal distribution with mean zero and variance $\sigma^2 N/n_i(N-n_i)$ and $(\nu S_p^2+\underline{Y}'\Lambda^{-1}\underline{Y})/\sigma^2$ has a chi-square distribution with N - 1 degrees of freedom, it follows that $\psi_i(\underline{Y},S_p^2)/\sqrt{\frac{N}{n_i(N-n_i)(N-1)}}$ has a t-distribution with N - 1 degrees of freedom. So we see from (3.8) that

$$c_i^* = t_{\gamma_i,N-1}\sqrt{\frac{N}{n_i(N-n_i)(N-1)}} \tag{3.9}$$

where $t_{\gamma_i,N-1}$ denotes the upper $100\gamma_i$ percent point of the t-distribution with N - 1 degrees of freedom.

<u>Case C: Gamma Scale Parameters Comparison: Unequal (Known) Shape Parameters.</u> Let π_i (i=0,...,k) be a gamma population with density

$$f(x;\theta_i,\nu_i) = \frac{x^{\nu_i-1}}{\Gamma(\nu_i)\theta_i^{\nu_i}}\exp\{-x/\theta_i\},\ x \geq 0,\ \theta_i > 0, \tag{3.10}$$

where the shape parameters ν_i are known. We take $\tau_{i0} = \theta_i/\theta_0$ so that $\tau^* = 1$. Let X_{ij}, j=1,...,n_i, be independent observations from π_i and define $T_{i0} = \bar{X}_i/\bar{X}_0$, i=1,...,k. The joint density of $T_{10},\ldots,T_{k0}$ is easily derived to be

$$h_{\underline{\tau}}(\underline{t}) = \Gamma(M)\prod_{i=0}^{k}\left\{\left(\frac{n_i}{n_0\tau_{i0}}\right)^{n_i\nu_i}\cdot\frac{1}{\Gamma(n_i\nu_i)}\right\} \times \frac{\prod_{i=1}^{k} t_i^{n_i\nu_i-1}}{\left[1+\sum_{i=1}^{k}\frac{n_i t_i}{n_0\tau_{i0}}\right]^M}, \tag{3.11}$$

where $M = \sum_{i=0}^{k} n_i\nu_i$. From this we get

$$(3.12) \qquad \frac{h^{(i)}_{\underline{\tau}^*}(\underline{t})}{h_{\underline{\tau}^*}(\underline{t})} = \frac{Mn_i t_i}{n_0 + \sum_{j=1}^{k} n_j t_j} - n_i \nu_i$$

$$= \frac{Mn_i \bar{x}_i}{\sum_{j=0}^{k} n_j \bar{x}_j} - n_i \nu_i \, .$$

Thus the locally optimal rule $\underline{\delta}^0$ is given by

$$(3.13) \qquad \delta_i^0(\bar{x}_0,\ldots,\bar{x}_k) = \begin{cases} 1 & \text{if} \quad n_i\bar{x}_i / \sum_{j=0}^{k} n_j \bar{x}_j \geq c_i^* \\ 0 & \text{otherwise} \end{cases}$$

where the constant c_i^* is determined by

$$(3.14) \qquad P_{\underline{\tau}^*}\left(\frac{n_i \bar{X}_i}{\sum_{j=0}^{k} n_j \bar{X}_j} \geq c_i^* \right) = \gamma_i \, .$$

When $\underline{\tau} = \underline{\tau}^*$, $n_j\bar{X}_j$ has a gamma distribution with parameters θ and $n_i\nu_i$. However, the probability in (3.14) is independent of θ. It is known that $n_i\bar{X}_i / \sum_{j=0}^{k} n_j\bar{X}_j$ has a beta distribution with parameters $n_i\nu_i$ and $M - n_i\nu_i$, denoted by $B(n_i\nu_i, M-n_i\nu_i)$. Thus c_i^* is the upper 100γ percentage point of $B(n_i\nu_i, M-n_i\nu_i)$ and can be obtained from tables of incomplete beta function.

Remark 3.2. It should be first pointed out that the above problem includes as a special case the problem of comparing normal variances based on samples of unequal sizes. It also includes the problem of comparing scale parameters θ_i of Weibull populations which has a common known shape parameter β. If X_{ij}, $j=1,\ldots,n_i$, are the sample observations from π_i, then we can transform these by $Y_{ij} = X_{ij}^{\beta}$ into sample observations from an exponential distribution with mean $\lambda_i = \theta_i^{\beta}$. Thus it is a special case of the gamma problem.

Case D: Comparison of Regression Slopes. Let π_i denote a simple linear regression model

$$Y = \alpha_i + \beta_i x + \varepsilon \tag{3.15}$$

where $\varepsilon \sim N(0,\sigma^2)$ and σ^2 is unknown. Let $\{x_{ij}, Y_{ij}\}$, $j=1,\ldots,n_i$, denote the sample data from (3.15). Define $\tau_{i0} = (\beta_i - \beta_0)/\sigma$, $i=1,\ldots,k$. The least squares estiamtors of α_i and β_i are given by $\hat{\alpha}_i = \bar{Y}_i - \hat{\beta}_i\bar{x}_i$ and $\hat{\beta}_i = S_{x_iY}/S^2_{x_i}$, where $\bar{x}_i = \Sigma x_{ij}/n_i$, $\bar{Y}_i = \Sigma Y_{ij}/n_i$, $S^2_{x_i} = \Sigma(x_{ij}-\bar{x}_i)^2$, and $S_{x_iY} = \Sigma(x_{ij}-\bar{x}_i)(Y_{ij}-\bar{Y}_i)$. All the summations are over j going from 1 to n_i. The usual pooled unbiased estimator of σ^2 is $S_p^2 = \sum_{i=0}^{k}\sum_{j=1}^{n_i}(y_{ij}-\hat{\alpha}_i-\hat{\beta}_i x_{ij})^2/\sum_{i=0}^{k}(n_i-2)$. It is well-known that $Q = \nu S_p^2/\sigma^2$ has a chi-square distribution with $\nu = \sum_{i=0}^{k}(n_i-2)$ degrees of freedom. Define $Z_i = \hat{\beta}_i - \hat{\beta}_0$ and $T_{i0} = Z_i/S_p$, $i=1,\ldots,k$. Then the joint density $h_{\underline{T}}(\underline{t})$ of $\underline{T} = (T_{10},\ldots,T_{k0})$ is the same as the expression in (3.6) but with $\tau_{i0} = (\beta_i-\beta_0)/\sigma$ and $\Lambda = (\lambda_{ij})$ given by

$$\lambda_{ij} = \begin{cases} \dfrac{1}{S^2_{x_0}}\left(1+\dfrac{1}{a_i}\right), & i = j \\[2ex] \dfrac{1}{S^2_{x_0}}, & i \neq j \end{cases} \tag{3.16}$$

where $a_i = S^2_{x_i}/S^2_{x_0}$, $i=1,\ldots,k$. It is now easy to see that the locally optimal rule $\underline{\delta}^0$ is given by

$$(3.17) \qquad \delta_i^0(\underline{z}) = \begin{cases} 1 & \text{if } \psi_i(\underline{z}; S_{x_0}^2, \ldots, S_{x_k}^2) \geq c_i^* \\ 0 & \text{otherwise} \end{cases}$$

where $\underline{z} = (z_1, \ldots, z_k)$ and

$$(3.18) \qquad \psi_i(\underline{z}; S_{x_0}^2, \ldots, S_{x_k}^2) = \frac{z_i - \dfrac{\sum_{j \neq i} S_{x_j}^2 z_j}{S_{x_0}^2 + \sum_{j \neq i} S_{x_j}^2}}{\sqrt{\nu s_p^2 + \underline{z}' \Lambda^{-1} \underline{z}}} .$$

The constant c_i^* is determined by

$$(3.19) \qquad P_{\underline{\tau}^*}(\psi_i(\underline{Z}; S_{x_0}^2, \ldots, S_{x_k}^2) \geq c_i^*) = \gamma_i$$

and by appealing to the result of Case B, we get

$$(3.20) \qquad c_i^* = t_{\gamma_i, N-k-2} \frac{1}{S_{x_i}} \sqrt{\left(1 + \frac{S_{x_i}^2}{S_{x_0}^2 + \sum_{j \neq i} S_{x_j}^2}\right) / (N+k-2)} .$$

REFERENCES

[1] Berger, R. L. and Gupta, S. S. (1980). Minimax subset selection rules with applications to unequal variance (unequal sample size) problems. Scand. J. Statist. 7, 21-26.

[2] Bickel, P. J., and Yahav, J. A. (1977). On selecting a subset of good populations. Statistical Decision Theory and Related Topics-II (Eds. S. S. Gupta and D. S. Moore), New York: Academic Press, pp 37-55.

[3] Bjørnstad, J. F. (1981). A decision-theoretic approach to subset selection. Commun. Statist.-Theor. Meth. A10, 2411-2433.

[4] Chernoff, H., and Yahav, J. A. (1977). A subset selection problem employing a new criterion. Statistical Decision Theory and Related Topics-II (Eds. S. S. Gupta and D. S. Moore), New York: Academic Press, pp. 93-119.

[5] Goel, P. K. and Rubin, H. (1977). On selecting a subset containing the best population - a Bayesian approach. Ann. Statist. 5, 969-983.

[6] Gupta, S. S. and Hsiao, P. (1981). On Γ-minimax, minimax, and Bayes procedures for selecting populations close to a control. Sankhyā Ser. B 43, 291-318.

[7] Gupta, S. S. and Hsu, J. C. (1978). On the performance of some subset selection procedures. Commun. Statist.-Simula. Computa. B7, 561-591.

[8] Gupta, S. S. and Huang, D. Y. (1977). On some Γ-minimax subset selection and multiple comparison procedures. Statistical Decision Theory and Related Topics-II (Eds. S. S. Gupta and D. S. Moore), New York: Academic Press, pp.139-155.

[9] Gupta, S. S. and Huang, D. Y. (1980). A note on optimal subset selection procedures. Ann. Statist. 8, 1164-1167.

[10] Gupta, S. S. and Huang, D. Y. (1981). Multiple Statistical Decision Theory: Recent Developments. Lecture Notes in Statistics, Vol. 8, New York: Springer-Verlag.

[11] Gupta, S. S., Huang, D. Y., and Nagel, K. (1979). Locally optimal subset selection procedures based on ranks. Optimizing Methods in Statistics (Ed. J. Rustagi), New York: Academic Press, pp.251-262.

[12] Gupta, S. S. and Kim, W. C. (1980). Γ-minimax and minimax decision rules for comparison of treatments with a control. Recent Developments in Statistical Inference and Data Analysis (Ed. K. Matusita), Amsterdam: North-Holland Publishing Co., pp. 55-71.

[13] Gupta, S. S. and Kim, W. C. (1981). On the problem of selecting good populations. Commun. Statist.-Theor. Meth. A10, 1043-1077.

[14] Gupta, S. S. and Miescke, K.-J. (1981). Optimality of subset selection procedures for ranking means of three normal populations. Sankhyā Ser. B 43, 1-17.

[15] Gupta, S. S. and Panchapakesan, S. (1979). Multiple Decision Procedures: Theory and Methodology of Selecting and Ranking Populations, New York: John Wiley.

[16] Huang, D. Y. and Panchapakesan, S. (1982). Some locally optimal subset selection rules based on ranks. Statistical Decision Theory and Related Topics-III, Vol. 2, (Eds. S. S. Gupta and J. Berger), New York: Academic Press, pp. 1-14.

[17] Miescke, K.-J. (1979). Bayesian subset selection for additive and linear loss functions. Commun. Statist.-Theor. Meth. A8, 1205-1226.

[18] Nagel, K. (1970). On subset selection rules with certain optimality properties. Ph.D. Thesis. Also Mimeo. Ser. No. 222, Dept. of Statist., Purdue University, W. lafayette, Indiana.

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER Technical Report #82-30	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) SOME LOCALLY OPTIMAL SUBSET SELECTION RULES FOR COMPARISON WITH A CONTROL		5. TYPE OF REPORT & PERIOD COVERED Technical
		6. PERFORMING ORG. REPORT NUMBER Technical Report #82-30
7. AUTHOR(s) Deng-Yuan Huang, S. Panchapakesan, and Sheng-Tsaing Tseng		8. CONTRACT OR GRANT NUMBER(s) N00014-75-C-0455
9. PERFORMING ORGANIZATION NAME AND ADDRESS Purdue University Department of Statistics West Lafayette, IN 47907		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS
11. CONTROLLING OFFICE NAME AND ADDRESS Office of Naval Research Washington, DC		12. REPORT DATE August 1982
		13. NUMBER OF PAGES 14
14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office)		15. SECURITY CLASS. (of this report) Unclassified
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release, distribution unlimited.

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

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Subset selection, locally optimal, comparison with control, normal means, gamma scale parameters, regression slopes.

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

Let $\pi_0, \pi_1, \ldots, \pi_k$ be k+1 independent populations. For $i = 0,1,\ldots,k$, π_i has the density $f(x,\theta_i)$, where the (unknown) parameter θ_i belongs to an interval of the real line. Our goal is to select from $\pi_1,\ldots,\pi_k$ (experimental treatments) those populations, if any, that are better (suitably defined) than π_0 which is the control population. A locally optimal rule is derived in the class of rules for

DD FORM 1 JAN 73 1473

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered) which $Pr(\pi_i$ is selected) $= \gamma_i$, $i = 1,\ldots,k$, when $\theta_0 = \theta_1 = \ldots = \theta_k$. The criterion used for local optimality amounts to maximizing the efficiency in a certain sense of the rule in picking out the superior populations for specific configurations of $\underline{\theta} = (\theta_0,\ldots,\theta_k)$ in a neighborhood of an equiparameter configuration. The general result is then applied to the following special cases: (a) normal means comparison - common known variance, (b) normal means comparison - common unknown variance, (c) gamma scale parameters comparison - known (unequal) shape parameters, and (d) comparison of regression slopes. In all these cases, the rule is obtained based on samples of unequal sizes.