AN ASYMPTOTIC COMPARISON OF SUBSET SELECTION PROCEDURES*

MILTON SOBEL AND S. P. YEN

(Received Feb. 16, 1971)

A procedure R for selecting a subset of k populations containing at least one of the t best populations was introduced in [2]. For normal populations we put π_i in the selected subset if and only if $\bar{X}_i \geqq \bar{X}_{[k-s+1]} - a_s$, where $\bar{X}_i$ is the sample mean from π_i and the ordered sample means are $\bar{X}_{[1]} \leqq \cdots \leqq \bar{X}_{[k]}$. Under procedure R both $s \geqq 1$ and $a_s \geqq 0$ are determined so that the probability of a correct subset $P\{CS\} \geqq P^*$ (specified), whenever the minimum difference between any one of the t largest population means and any one of the $k-t$ smallest is at least δ^* (specified). For $t=1$ and $\delta^*=0$ the goal is the same as that considered by Gupta [1], but his procedure R_G is not the same as procedure R. In [2] many exact and asymptotic comparisons are made for $t \geqq 1$ and $\delta^* \geqq 0$ but the emphasis there is on the equal parameter (EP) configuration, where the expected subset size is maximized. Moreover for $t=1$ and $\delta^*=0$ the value of the expected subset sizes $E\{S|\text{EP}\}$ is the same, namely kP^*, for both procedures and hence this criteria does not lead to any clear preference in this special case. It was shown in [2] that if either $t>1$ or $\delta^*>0$ then asymptotically ($P^* \to 1$) the value of $E\{S|\text{EP}\}$ is smaller for procedure R than for the natural generalization R_M (cf. [2]) of procedure R_G for $t>1$ and $\delta^*>0$; in fact, the ratio approaches zero as $P^* \to 1$.

In this note we consider only the special case $t=1$ and $\delta^*=0$ and make asymptotic ($P^* \to 1$) comparisons of $E\{S|\boldsymbol{\theta}\}$ for any k-vector $\boldsymbol{\theta}$ of true parameter values. Let $\theta_{[1]} \leqq \cdots \leqq \theta_{[k]}$ denote the ordered parameter values and let $\delta_{ij} = \theta_{[i]} - \theta_{[j]}$. In terms of the differences δ_{ij}, we find the exact sets S_R and S_G of vectors $\boldsymbol{\theta}$ which have a smaller asymptotic ($P^* \to 1$) value for $E\{S|\boldsymbol{\theta}\}$ under procedure R and R_G, respectively. Since both S_R and S_G are non-empty, it follows that neither of these two procedures is uniformly better than the other in the sense of this criterion. We assume normal populations with a common variance σ^2, which we can take to be unity. Let $\bar{X}_{(i)}$ denote the sample mean which

* This research was supported by National Science Foundation Grant GP-28922X.

has parameter value $\theta_{[i]}$ $(i=1, 2, \cdots, k)$.

Under procedure R with $t=1$, we set $s=k-1$ for P^* close to one and obtain for $\delta^*=0$

$$
\begin{aligned}
(1)\qquad E\{S|\boldsymbol{\theta}, R\} &= \sum_{i=1}^{k} P\{\bar{X}_{(i)} \geqq \bar{X}_{[2]} - a_{k-1}\} \\
&= \sum_{i=1}^{k} \Big[1 - \sum_{\substack{j=1 \\ j\neq i}}^{k} P\{\bar{X}_{(i)} + a_{k-1} < \bar{X}_{(j)},\ \bar{X}_{(j)} = \bar{X}_{[2]}\}\Big] \\
&= k - \sum_{i=1}^{k} \sum_{\substack{j=1 \\ j\neq i}}^{k} \int \Phi(x+\lambda_{ji}-A) \prod_{\substack{\alpha=1 \\ \alpha\neq i,j}}^{k} [1-\Phi(x+\lambda_{j\alpha})] d\Phi(x)\ ,
\end{aligned}
$$

where $A=a_{k-1}\sqrt{n}$, $\lambda_{ij}=\delta_{ij}\sqrt{n}$ and $\Phi(x)$, $\varphi(x)$ are used to denote the standard normal c.d.f. and density, respectively. As in Section 8 of [2] we use the Laplace-Feller expansion for the 'tail' of the normal c.d.f. in (1), drop the denominator and then 'complete the square.' Neglecting the error term, $o(\exp\{-\theta^2(A-\lambda_{k1})^2/2\})$, we obtain

$$
\begin{aligned}
(2)\qquad k - E\{S|\boldsymbol{\theta}, R\} &\approx \frac{C}{A} \sum_{i} \sum_{j\neq i} \int [\varphi(\chi+A-\lambda_{ji})\varphi(x)] \prod_{\alpha=1}^{k} \Phi(x-\lambda_{j\alpha}) dx \\
&\approx \frac{C}{A} \sum_{i} \sum_{j\neq i} \phi\Big(\frac{A-\lambda_{ji}}{\sqrt{2}}\Big) \int \prod_{\substack{\alpha=1 \\ \alpha\neq i,j}}^{k} \Phi\Big(\frac{y}{\sqrt{2}}+D_\alpha\Big) d\Phi(y) \\
&\approx \frac{C}{A^{k-1}} \sum_{i} \sum_{j\neq i} \varphi\Big(\frac{A-\lambda_{ji}}{\sqrt{2}}\Big) \int \Big[\prod_{\substack{\alpha=1 \\ \alpha\neq i,j}}^{k} \varphi\Big(\frac{y}{\sqrt{2}}+D_\alpha\Big)\Big] \varphi(y) dy \\
&\approx \frac{C}{A^{k-1}} \sum_{i} \sum_{j\neq i} \varphi\Big(\frac{A-\lambda_{ji}}{\sqrt{2}}\Big) \\
&\quad \cdot \exp\Big\{-\Big[\sum_{\alpha\neq i,j} D_\alpha^2 - \frac{1}{k}\Big(\sum_{\alpha\neq i,j} D_\alpha\Big)^2\Big]\Big/2\Big\}
\end{aligned}
$$

where $D_\alpha=(\lambda_{ji}-A-2\lambda_{j\alpha})/2$. Collecting the factors of the form $\exp\{-CA^2\}$ and $\exp\{CA\}$, we use the fact that

$$
(3)\qquad (k-1)\lambda_{ji} - \sum_{\alpha\neq i,j} \lambda_{j\alpha} = \lambda_{ji} + \sum_{\alpha\neq j} \lambda_{\alpha i} = \sum_{\alpha} \lambda_{\alpha i}
$$

does not depend on j, and obtain from (2)

$$
\begin{aligned}
(4)\qquad k - E\{S|\boldsymbol{\theta}, R\} &\approx \frac{C}{A^{k-1}} \sum_{i} \sum_{j\neq i} \exp\Big\{-\Big(\frac{k-1}{2k}\Big)(A-\lambda_{ji})^2\Big\} \\
&\quad \cdot \exp\Big\{\frac{1}{k}(\lambda_{ji}-A) \sum_{\alpha\neq i,j} \lambda_{j\alpha}\Big\} \\
&\approx \frac{C}{A^{k-1}} \exp\Big\{-\Big(\frac{k-1}{2k}\Big)A^2\Big\} \sum_{i} \exp\Big\{\frac{A}{k}\sum_{\alpha} \lambda_{\alpha i}\Big\}\ .
\end{aligned}
$$

The maximum term in (4) for large A is obtained by maximizing over i the sum in braces and this clearly occurs for $i=1$. Hence

$$k-E\{S|\boldsymbol{\theta}, R\} \approx \frac{C}{A^{k-1}} \exp\left\{-\left(\frac{k-1}{2k}\right)A^2+\frac{A}{k}\sum_{\alpha}\lambda_{\alpha 1}\right\}. \tag{5}$$

It is shown in (8.11) of [2] that for $P^* \to 1$

$$A \approx \sqrt{2\left(\frac{k}{k-1}\right) ln\left(\frac{1}{1-P^*}\right)} \tag{6}$$

and applying this to (5) gives the final form for procedure R

$$k-E\{S|\boldsymbol{\theta}, R\} = \frac{C}{A^{k-1}}(1-P^*) \exp\left\{\left(\sum_{\alpha}\lambda_{\alpha 1}\right)\sqrt{\frac{2}{k(k-1)} ln\left(\frac{1}{1-P^*}\right)}\right\}. \tag{7}$$

For procedure R_G it is easily shown as in Gupta [1] that

$$\begin{aligned} E\{S|\boldsymbol{\theta}, R_G\} &= \sum_{i=1}^{k} \int \prod_{j\neq i} \Phi(x+A+\lambda_{ij}) d\Phi(x) \\ &\approx \sum_{i} \int \left\{1-\sum_{j\neq i}[1-\Phi(x+A+\lambda_{ij})]\right\} d\Phi(x) \\ &= k-\sum_{i}\sum_{j\neq i}\left[1-\Phi\left(\frac{A+\lambda_{ij}}{\sqrt{2}}\right)\right]. \end{aligned} \tag{8}$$

Hence

$$\begin{aligned} k-E\{S|\boldsymbol{\theta}, R_G\} &= \frac{C}{A}\sum_{i}\sum_{j\neq i}\varphi\left(\frac{A+\lambda_{ij}}{\sqrt{2}}\right) \\ &\approx \frac{C}{A}e^{-A^2/4}\sum_{i}\sum_{j\neq i}e^{A\lambda_{ij}/2}. \end{aligned} \tag{9}$$

The maximum term for large A is obtained by setting $i=k$ and $j=1$; hence this gives

$$k-E\{S|\boldsymbol{\theta}, R_G\} \approx \frac{C}{A}\exp\left\{-\frac{A^2}{4}+\frac{A}{2}\lambda_{k1}\right\}. \tag{10}$$

In (8.11) of [2] we set $s=t=1$ to obtain the A-value for procedure R_G, namely

$$A \approx 2\sqrt{ln\left(\frac{1}{1-P^*}\right)} \tag{11}$$

and applying this to (10) gives the final form

$$k-E\{S|\boldsymbol{\theta}, R_G\} \approx \frac{C}{A}(1-P^*) \exp\left\{\lambda_{k1}\sqrt{ln\left(\frac{1}{1-P^*}\right)}\right\}. \tag{12}$$

It follows from (7) and (12) that $E\{S|\boldsymbol{\theta}, R\}$ is smaller than $E\{S|\boldsymbol{\theta}, R_G\}$ for P^* close to one when

$$\sqrt{\frac{2}{k(k-1)}} \sum_{\alpha} \lambda_{\alpha 1} > \lambda_{k1} , \tag{13}$$

and it is larger when the inequality is reversed. For $k=2$ the procedures are identical and (13) is vacuous. For $k=3$ the inequality in (13) holds when

$$\delta_{21} > (1+\sqrt{3})\delta_{32} = (2.732\ldots)\delta_{32} . \tag{14}$$

If we define the configuration C_j ($j=1, 2, \cdots, k$) by setting

$$\theta_{[1]} = \cdots = \theta_{[k-j]} ; \qquad \theta_{[k-j+1]} = \cdots = \theta_{[k]} \tag{15}$$

then (13) takes the form

$$j > \sqrt{\frac{k(k-1)}{2}} \tag{16}$$

and we note that (13) always holds for C_{k-1} for $k>2$. On the other hand, for all $k>2$ the inequality in (13) is reversed for C_1 and also for the configuration in which adjacents parameters are equally spaced.

A table of values for $E\{S|C_j\}$ ($j=1, 2, 3, 4$) for $k=5$ is included in [3] and it illustrates numerically the results proved above.

UNIVERSITY OF MINNESOTA, MINNEAPOLIS, MINNESOTA

REFERENCES

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

[2] Sobel, M. (1969). Selecting a subset containing at least one of the t best populations. *Multivariate Analysis Vol. II, Proceedings of an International Symposium*. Ed. by P. R. Krishnaiah, Academic Press, New York.

[3] Yen, S. P. (1969). *Efficiency Comparisons for Two Subset Selection Procedures*, University of Minnesota, Technical Report, 119.