COMMUN. STATIST.-THEOR. METH., A7(8), 799-811 (1978)

ON TWO-STAGE SELECTION PROCEDURES AND RELATED PROBABILITY-INEQUALITIES

Yosef Rinott

Hebrew University of Jerusalem
Israel

Key Words and Phrases: normal means; unknown variances; probability of correct selection.

ABSTRACT

In this paper we discuss a modification of the Dudewicz-Dalal procedure for the problem of selecting the population with the largest mean from k normal populations with unknown variances. We derive some inequalities and use them to lower-bound the probability of correct selection. These bounds are applied to the determination of the second-stage sample size which is required in order to achieve a prescribed probability of correct selection. We discuss the resulting procedure and compare it to that of Dudewicz and Dalal (1975).

1. INTRODUCTION

Let X_{ij} be normal and independent random variables from population π_i with mean μ_i and variance σ_i^2 ($i = 1,\ldots,k$, $j = 1, 2,\ldots$). We assume that the μ_i and σ_i are unknown. The ordered μ_i's are denoted by $\mu_{[1]} \leq \cdots \leq \mu_{[k]}$. Our goal is to select the best population, namely the population associated with the largest mean $\mu_{[k]}$. Let

$$\Omega(\delta^*) = \{(\mu_1,\ldots,\mu_k): \ \mu_{[k]} - \mu_{[k-1]} \geq \delta^*\}.$$

The problem is to find a rule for which the probability of correct selection (denoted by $P(CS)$) will be greater than or equal to a prescribed number P^* throughout $\Omega(\delta^*)$, i.e. our probability requirement is

$$P(CS) \geq P^* \ \text{ whenever } \ (\mu_1,\ldots,\mu_k) \in \Omega(\delta^*). \tag{1}$$

A single-stage procedure cannot satisfy this requirement. (Dudewicz (1971)).

Two-stage procedures for this problem which are generalizations of Stein's (1945) two-stage procedures have been given by Dudewicz and Dalal (1975). Dudewicz and Dalal developed the following procedure $P_E(h)$:

Take an initial sample of N_0 observations from each population π_i and a second sample of size $N_i - N_0$ observations from each π_i where

$$N_i = \max\{N_0 + 1, \ \ [(\tfrac{h}{\delta^*})^2 S_i^2]\},$$

h is a constant to be discussed later and S_i^2 is the usual unbiased estimate of the variance σ_i^2 based on the first N_0 observations from π_i, $i = 1,\ldots,k$. (Here $[y]$ denotes the smallest integer $\geq y$.) Choose a_{ij} $(j = 1,\ldots,N_i)$ that satisfy

$$\sum_{j=1}^{N_i} a_{ij} = 1, \ \ a_{i1} = \ldots = a_{iN_0}, \ \text{ and } \ S_i^2 \sum_{j=1}^{N_i} a_{ij}^2 = (\frac{\delta^*}{h})^2$$

for $i = 1,\ldots,k$. Set $\tilde{X}_i = \Sigma_{j=1}^{N_i} a_{ij}X_{ij}$ and select as the "best" the population which gave rise to the largest of the generalized sample means $\tilde{X}_i$, $i = 1,\ldots,k$. They show that for any $\sigma_1^2,\ldots,\sigma_k^2$

$$P(CS|P_E) \geq \int_{-\infty}^{\infty} G^{k-1}(t+h)g(t)dt \tag{2}$$

provided $(\mu_1,\ldots,\mu_k) \in \Omega(\delta^*)$, with equality in (2) holding whenever $\mu_{[1]} = \cdots = \mu_{[k-1]} = \mu_{[k]} - \delta^*$ (to be called the least favorable configuration of $\mu_1,\ldots,\mu_k$), where G and g denote the distribution function and density, respectively of the t distribution with $N_0 - 1$ degrees of freedom. Thus probability requirement (1) is satisfied if we apply the procedure $P_E(h)$ with h determined by the integral equation

$$\int_{-\infty}^{\infty} G^{k-1}(t+h)g(t)dt = P^* . \tag{3}$$

Following Stein (1945) one may expect that a similar procedure based on ordinary sample means may be more appealing than the above procedure. The following procedure called $P_R(h)$ is discussed by Dudewicz and Dalal (1975) and generalizes Stein's considerations in a natural way. The first stage in $P_R(h)$ is the same as in $P_E(h)$ above. Then in the second stage take $N_i - N_0$ additional observations from each π_i where now

$$N_i = \max\{N_0, [(\tfrac{h}{\delta^*})^2 S_i^2]\}.$$

Set $\overline{X}_i = (1/N)\Sigma_{j=1}^{N_i} X_{ij}$ and select as "best" the population which yields the largest of the overall sample means $\overline{X}_i$, $i = 1,\ldots,k$. The question whether it is generally true that

$$P(CS|P_R(h)) \geq P(CS|P_E(h)) \tag{4}$$

was raised by Dudewicz and Dalal and remained open (except for the case $k = 2$ in which they showed that (4) holds so that P_R is uniformly better than P_E) and thus it is not clear how h should be determined so that P_R will guarantee the probability requirement (1). We show that (4) does not hold in general and provide a new integral equation (13) from which h^* can be determined so that $P_R(h^*)$ will guarantee (1).

We discuss one- and two-stage procedures based on sample means showing that a procedure which intuitively may seem most reasonable is sometimes worse than guessing, i.e. may yield $P(CS) < 1/k$ for certain configurations of the variances and suitable δ^*. Furthermore a decrease in the variances of the statistics involved may cause a decrease in $P(CS)$ contrary to statistical intuition.

We prove some integral inequalities relating (13) and (3) and compare $P_R(h^*)$ and $P_E(h)$ from the points of view of sample size and probability of correct selection. Both P_R and P_E have some undesirable properties but it appears to us that unless one is interested in very small samples and small P^*, P_R might be the more efficient procedure, having a larger $P(CS)$ under configurations that differ from the least favorable one.

2. THE PROCEDURE $P_R(h^*)$

In this section we show that the procedure $P_R(h^*)$ (defined in Section 1) satisfies our probability requirement.

For easy reference we first state the following result due to Slepian (1962) which is basic in many of our considerations.

If $(X_1,\ldots,X_n)$ has the multivariate normal distribution with nonsingular covariance matrix $\Sigma = ||\sigma_{ij}||_{i,j=1}^n$, then for any constants $c_1,\ldots,c_n$ the probability $\Pr(X_1 \leq c_1,\ldots,X_n \leq c_n)$ is strictly increasing as a function of each σ_{ij} for $i \neq j$. In particular if $\sigma_{ij} > 0$, $i, j = 1,\ldots,n$ then

$$\Pr(X_1 \leq c_1,\ldots,X_n \leq c_n) > \prod_{i=1}^{n} \Pr(X_i \leq c_i).$$

Proposition 1. *The procedure* $P_R(h^*)$ *with* h^* *defined by* (13) *below satisfies* $P(CS) > P^*$ *for all* $(\mu_1,\ldots,\mu_k) \in \Omega(\delta^*)$ *and all values of* $\sigma_1^2,\ldots,\sigma_k^2$.

Proof. It is easy to see that the $P(CS)$ of P_R is minimized over $\Omega(\delta^*)$ when $\mu_{[1]} = \cdots = \mu_{[k-1]} = \mu_{[k]} - \delta^*$ (the so-called least favorable configuration); we shall consider this configuration in the following probability calcualtions. Denote by $\pi_{(i)}$ the population having mean $\mu_{[i]}$, $(i = 1,\ldots,k)$. Let $\sigma_{(i)}^2$ be the variances of $\pi_{(i)}$ and denote by $S^2_{(i)}$ the sample variances based on the first N_0 observations from $\pi_{(i)}$. Let $N_{(i)}$ be the total number of observations from $\pi_{(i)}$ and $\overline{X}_{(i)}$ the overall sample mean from $\pi_{(i)}$ so that $E\overline{X}_{(i)} = \mu_{[i]}$, $i = 1,\ldots,k$. We then have

$$P(CS) = \Pr(\overline{X}_{(i)} < \overline{X}_{(k)},\ i = 1,\ldots,k-1)$$

$$= \Pr\left\{\frac{\overline{X}_{(i)} - (\overline{X}_{(k)} - \delta^*)}{[(\sigma^2_{(i)}/N_{(i)}) + (\sigma^2_{(k)}/N_{(k)})]^{1/2}} < \frac{\delta^*}{[(\sigma^2_{(i)}/N_{(i)}) + (\sigma^2_{(k)}/N_{(k)})]^{1/2}},\ i = 1,\ldots,k-1\right\}. \tag{5}$$

Denote

$$Z_{(i)} = \frac{\overline{X}_{(i)} - (\overline{X}_{(k)} - \delta^*)}{[(\sigma^2_{(i)}/N_{(i)}) + (\sigma^2_{(k)}/N_{(k)})]^{1/2}}$$

and

$$Q_{(i)} = \frac{h^*}{[(\sigma^2_{(i)}/S^2_{(i)}) + (\sigma^2_{(k)}/S^2_{(k)})]^{1/2}},\ i = 1,\ldots,k-1.$$

Since $N_{(i)} \geq (\frac{h^*}{\delta^*})^2 S^2_{(i)}$ it follows that

$$\frac{\delta^*}{[(\sigma^2_{(i)}/N_{(i)}) + (\sigma^2_{(k)}/N_{(k)})]^{1/2}} \geq Q_{(i)}$$

so that (5) implies

$$P(CS) \geq \Pr(Z_{(i)} < Q_{(i)},\ i = 1,\ldots,k-1). \tag{6}$$

By standard arguments it follows that under the least favorable configuration the conditional distribution of $Z_{(1)},\ldots,Z_{(k-1)}$ given $S^2_{(1)},\ldots,S^2_{(k)}$ is multivariate normal with means equal to zero and covariance matrix determined by $\operatorname{Var} Z_{(i)} = 1$, $i = 1,\ldots,k-1$, and for $i \neq j$, $i,j = 1,\ldots,k-1$

$$\operatorname{Cov}(Z_{(i)},Z_{(j)}) = \frac{\sigma^2_{(k)}/N_{(k)}}{[(\sigma^2_{(i)}/N_{(i)})+(\sigma^2_{(k)}/N_{(k)})]^{1/2}[(\sigma^2_{(j)}/N_{(j)})+(\sigma^2_{(k)}/N_{(k)})]^{1/2}}. \tag{7}$$

Since the variances are constant, the covariances positive, and since the $Q_{(i)}$ are functions of $S^2_{(1)},\ldots,S^2_{(k)}$ and may therefore be regarded as constants when we condition on $S^2_{(1)},\ldots,S^2_{(k)}$, we have by Slepian's inequality

$$\Pr(Z_{(i)} < Q_{(i)},\ i = 1,\ldots,k-1 | S^2_{(1)},\ldots,S^2_{(k)})$$
$$> \prod_{i=1}^{k-1} \Pr(Z_{(i)} < Q_{(i)} | S^2_{(1)},\ldots,S^2_{(k)}). \tag{8}$$

Let Φ denote the distribution function of a $N(0,1)$ variable i.e. $\Phi(t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{t} e^{-x^2/2}dx$. Note that the marginal conditional distribution of each $Z_{(i)}$ given $S^2_{(1)},\ldots,S^2_{(k)}$ is $N(0,1)$ so that each $Z_{(i)}$ is independent of $S^2_{(1)},\ldots,S^2_{(k)}$. Thus

$$\Pr(Z_{(i)} < Q_{(i)} | S^2_{(1)},\ldots,S^2_{(k)}) = \Phi(Q_{(i)}). \tag{9}$$

Combining (8) and (9) we obtain

$$\Pr(Z_{(i)} < Q_{(i)},\ i = 1,\ldots,k-1)$$
$$= E\{\Pr(Z_{(i)} < Q_{(i)},\ i = 1,\ldots,k-1 | S^2_{(1)},\ldots,S^2_{(k)})\}$$
$$> E\left\{\prod_{i=1}^{k-1} \Phi(Q_{(i)})\right\} = E\left\{\prod_{i=1}^{k-1} \Phi\left(\frac{h^*}{[(\sigma^2_{(i)}/S^2_{(i)})+(\sigma^2_{(k)}/S^2_{(k)})]^{1/2}}\right)\right\}. \tag{10}$$

The variables $Y_{(i)}$ defined by $Y_{(i)} = (N_0-1)S^2_{(i)}/\sigma^2_{(i)}$, $i = 1,\ldots,k$, are independent χ^2 variables with $N_0 - 1$ degrees of freedom, and by (6) and (10) we have

$$P(CS) > E\left\{\prod_{i=1}^{k-1} \Phi\left(\frac{h^*}{\{(N_0-1)[(1/Y_{(i)}) + (1/Y_{(k)})]\}^{1/2}}\right)\right\}. \tag{11}$$

Note that the expression on the right side of (11) is independent of $\sigma^2_{(1)},\ldots,\sigma^2_{(k)}$. Let f denote the density of the χ^2 distribution with $N_0 - 1$ degrees of freedom. Applying the independence of the $Y_{(i)}$, $i = 1,\ldots,k$, we can simplify the above expression to

$$E\left\{\prod_{i=1}^{k-1} \Phi\left(\frac{h^*}{\{(N_0-1)[(1/Y_{(i)}) + (1/Y_{(k)})]\}^{1/2}}\right)\right\}$$

$$= E\left\{E\left[\prod_{i=1}^{k-1} \Phi\left(\frac{h^*}{\{(N_0-1)\ [(1/Y_{(i)}) + (1/Y_{(k)})]\}^{1/2}}\right) \middle| Y_{(k)}\right]\right\}$$

$$= \int_0^\infty\left[\int_0^\infty \Phi\left(\frac{h^*}{\{(N_0-1)[(1/x) + (1/y)]\}^{1/2}}\right)f(x)dx\right]^{k-1} f(y)dy. \tag{12}$$

Thus (11) and (12) imply $P(CS) > P^*$ if h^* is determined by the integral equation

$$\int_0^\infty\left[\int_0^\infty \Phi\left(\frac{h^*}{\{(N_0-1)[(1/x) + (1/y)]\}^{1/2}}\right)f(x)dx\right]^{k-1} f(y)dy = P^*. \tag{13}$$

3. PROPERTIES OF P_R

In this section we point out some undesirable properties of selection procedures based on sample means. We also answer the question of comparison of P_E and P_R raised by Dudewicz and Dalal (1975) (see eq. (4)) and extend some of their results.

Consider a one-stage procedure for the problem of identifying the normal population having largest mean, variances unknown, where N observations are taken from each population and the population yielding the largest of the sample means $\overline{X}_1,\ldots,\overline{X}_k$ is selected as "best". Assume $\mu_{[1]} = ,\ldots, = \mu_{[k-1]} = \mu_{[k]} - \delta^*$ where δ^* is small to be specified below, and consider a configuration of the variances such that $\sigma^2_{(1)} = ,\ldots, = \sigma^2_{(k-1)} > \sigma^2_{(k)}$, $\sigma^2_{(i)}$ denoting as before the variance associated with the population having mean $\mu_{[i]}$. Denote $\theta = \sigma^2_{(i)}/\sigma^2_{(k)} > 1$. For $k > 2$ we have:

Proposition 2. Under these conditions the procedure performs worse than guessing without any observations, i.e. $P(CS) < \frac{1}{k}$.

Proof. With notation as above

$$P(CS) = \Pr(\overline{X}_{(i)} < \overline{X}_{(k)},\ \ i = 1,\ldots,k-1)$$

$$= \Pr\left\{Z_{(i)} < \frac{\delta^*}{[(\sigma^2_{(i)}/N) + (\sigma^2_{(k)}/N)]^{1/2}},\ \ i = 1,\ldots,k-1\right\}, \tag{14}$$

and for $\delta^* = 0$

$$P(CS) = \Pr(Z_{(i)} < 0, \quad i = 1,\dots,k-1),$$

where $(Z_{(1)},\dots,Z_{(k-1)})$ is multivariate normally distributed with means equal to zero and covariance matrix determined by

$$\operatorname{Var} Z_{(i)}=1, \; \operatorname{Cov}(Z_{(i)},Z_{(j)}) = \left(\frac{\sigma^2_{(i)}}{\sigma^2_{(k)}} + 1\right)^{-1/2}\left(\frac{\sigma^2_{(j)}}{\sigma^2_{(k)}} + 1\right)^{-1/2} \tag{15}$$

$$= (\theta+1)^{-1} < 1/2, \quad 1 \leq i \neq j \leq k-1.$$

Let $(U_1,\dots,U_{k-1})$ be a $(k-1)$-variate normal random vector such that $E(U_i) = 0$, $\operatorname{Var}(U_i) = 1$ and $\operatorname{Cov}(U_i, U_j) = 1/2$, $1 \leq i \neq j \leq k-1$. Then we have

$$\Pr(Z_{(i)} < 0, \quad i = 1,\dots,k-1) < \Pr(U_i < 0, \quad i = 1,\dots,k-1)$$

$$= \int_{-\infty}^{\infty} \Phi^{k-1}(x)d\Phi(x) = 1/k,$$

where the strict inequality follows from Slepian's (1962) inequality. By continuity of $P(CS)$ in δ^* we have for some $\delta^* > 0$ and the configuration described before Proposition 2 $P(CS) < 1/k$. Moreover letting $\sigma^2_{(i)} \to \infty$ and $\sigma^2_{(k)}$ bounded or $\sigma^2_{(k)} \to \infty$ in such a way that $\theta = \sigma^2_{(i)}/\sigma^2_{(k)} \to \infty$, (14) and (15) imply

$$\Pr(\overline{X}_{(i)} < \overline{X}_{(k)}, \quad i = 1,\dots,k-1) \to (1/2)^{k-1}$$

so that for large values of the variances we can have

$$P(CS) \leq (1/2)^{k-1} + \varepsilon \quad \text{for any} \quad \varepsilon > 0$$

and clearly $(1/2)^{k-1} < \frac{1}{k}$ for $k > 2$.

Remark 1. For $k > 2$ there exist values of h and configurations of the parameters such that $P(CS|P_R(h)) < P(CS|P_E(h))$, showing that (4) is not universally true.

Proof: First note that (4) and (2) would imply

$$P(CS|P_R(h)) \geq \int_{-\infty}^{\infty} G^{k-1}(t+h)g(t)dt \geq \frac{1}{k}\,.$$

As $h \to 0$ the probability of drawing a second sample in $P_R(h)$

tends to 0 and hence $P(CS|P_R(h))$ will tend to the $P(CS)$ of the one-stage procedure described above, and it follows from the preceding discussion that for some $h > 0$ and certain configurations of parameters we have $P(CS|P_R(h)) < \frac{1}{k}$. This contradicts the above inequality and we conclude that (4) does not hold for sufficiently small values of h when $k > 2$. Note that the argument holds even when $P_R(h)$ and $P_E(h)$ are compared with different initial sample sizes.

Remark 2. Assume $\mu_{[1]} = \cdots = \mu_{[k-1]} = \mu_{[k]} - \delta^*$ and consider two configurations of variances $\sigma_{(1)},\ldots,\sigma_{(k)}$ and $\sigma'_{(1)},\ldots,\sigma'_{(k)}$ such that $\sigma_{(i)} = \sigma'_{(i)}$, $i = 1,\ldots,k-1$, and $\sigma_{(k)} < \sigma'_{(k)}$. Then for some δ^* and h we have $P(CS|P_R(h)) < P'(CS|P_R(h))$ where the probability on the left-hand side is under $\sigma_{(1)},\ldots,\sigma_{(k)}$ and on the right-hand side under $\sigma'_{(1)},\ldots,\sigma'_{(k)}$ (i.e. increasing a variance sometimes increases the probability of correct selection). This follows from the fact that the covariances in (15) increase in $\sigma_{(k)}$ and thus by Slepian's result increasing $\sigma_{(k)}$ will increase the $P(CS)$ of the aforementioned one-stage procedure for small values of δ^*. Letting $h \to 0$, $P(CS|P_R(h))$ tends to $P(CS)$ of the one-stage procedure implying that for some small δ^* and h, $P(CS|P_R(h)) < P'(CS|P_R(h))$.

4. INEQUALITIES AND THEIR APPLICATIONS TO A COMPARISON OF P_R AND P_E

We compare the procedure P_R with h determined by (13) and P_E both taken with initial sample of N_0 observations from each population.

4.1 Comparison of h

Recall that h is the constant appearing in the determination of the second sample size in P_R and P_E.

Proposition 3. *The value of* h^* *determined by (13) to guarantee the probability requirement (1) with the procedure* $P_R(h^*)$ *is larger than the value of* h *determined by (3) needed to guarantee*

<u>the probability requirement (1) by</u> $P_E(h)$ <u>for any</u> P^*, $k > 2$. <u>For</u> $k = 2$ <u>the two values coincide</u>.

<u>Proof</u>. Comparing the integral expressions in (13) and (3) from which the values of h are determined this proposition is equivalent to the relation:

$$\int_{-\infty}^{\infty} G^{k-1}(t+h)g(t)dt \geq \int_0^\infty \left[\int_0^\infty \Phi\left(\frac{h}{\{(N_0-1)(\frac{1}{x}+\frac{1}{y})\}^{1/2}}\right)f(x)dx\right]^{k-1} f(y)dy \tag{16}$$

with equality holding for $k = 2$, and strict inequality for $k > 2$. In order to prove (16) define Z_i to be $N(0,1)$ variables and S_i^2 to be χ^2 variables with $N_0 - 1$ degrees of freedom all independently distributed, $i = 1,\dots,k$. Setting $\tilde{Z}_i = Z_i/[S_i^2/(N_0-1)]^{1/2}$, $i = 1,\dots,k$, the $\tilde{Z}_i$ are independent t variables with $N_0 - 1$ degrees of freedom. We have

$$\begin{aligned}
&\int_{-\infty}^{\infty} G^{k-1}(t+h)g(t)dt \\
&\quad = \Pr(\tilde{Z}_i < \tilde{Z}_k + h,\ i = 1,\dots,k-1) \\
&\quad = E\left\{\Pr\left(\frac{\tilde{Z}_i - \tilde{Z}_k}{\{(N_0-1)[(1/S_i^2)+(1/S_k^2)]\}^{1/2}} \right.\right. \\
&\qquad \left.\left. < \frac{h}{\{(N_0-1)[(1/S_i^2)+(1/S_k^2)]\}^{1/2}},\ i = 1,\dots,k-1 \,\middle|\, S_1^2,\dots,S_k^2\right)\right\} \\
&\quad \geq E \prod_{i=1}^{k-1} \Phi\left(\frac{h}{\{(N_0-1)[(1/S_i^2)+(1/S_k^2)]\}^{1/2}}\right),
\end{aligned}$$

where the last inequality follows from Slepian's inequality with strict inequality for $k > 2$ and clearly equality holding for $k = 2$. as in (12) the last expression reduces to the right-hand side of (16).

In the next Proposition 4 we give a lower bound to the integral of equation (13). The lower bound will be applied to the procedure $P_R(h)$ as stated in Proposition 5.

Proposition 4.

$$\int_0^\infty\left[\int_0^\infty \Phi\left(\frac{h}{\{(N_0-1)[(1/x)+(1/y)]\}^{1/2}}\right)f(x)\,dx\right]^{k-1} f(y)\,dy$$

$$\geq \left(\int_{-\infty}^{\infty} G(t+h)g(t)\,dt\right)^{k-1} .$$

Proof. This follows from Jensen's inequality (in the form $EX^n \geq (EX)^n$) and the case $k = 2$ equality of (16).

Propositions 1 and 4 imply that

$$P(CS|P_R(h)) \geq \left(\int_{-\infty}^{\infty} G(t+h)g(t)\,dt\right)^{k-1}$$

thus proving

Proposition 5. The probability requirement (1) will be satisfied by $P_R(h)$ if h is determined from the equation

$$\int_{-\infty}^{\infty} G(t+h)g(t)\,dt = (P^*)^{1/(k-1)} . \tag{17}$$

This determination of h will of course yield larger values of h than those of h^* obtained by (13), the difference becoming less significant for large value of P^*. Dudewicz and Dalal (1975) presented tables for the integral in (17) and by Proposition 5 these tables can be applied for a conservative determination of h in $P_R(h)$ for any k.

Numerical comparisons indicate that for large values of P^* ($P^* \geq 0.75$) the difference between the values of h as determined from (3), (13) and (17) is small and thus the difference in the second-stage sample sizes between P_R and P_E due to the different determination of h in these procedures may not be very significant.

4.2 Comparison of Sample Size

Denote by $N_i^{(E)}$ and $N_i^{(R)}$ the sample sizes in P_E and P_R, respectively. Here $N_i^{(E)}$ is determined with h satisfying

(3), and $N_i^{(R)}$ is determined with h^* satisfying (13). By Proposition 3 we have $h^* \geq h$. Recall that

$$N_i^{(E)} = \max\{N_0 + 1, [(\frac{h}{\delta^*})^2 S_i^2]\}$$

and

$$N_i^{(R)} = \max\{N_0, [(\frac{h}{\delta^*})^2 S_i^2]\}, \quad i = 1,\ldots,k.$$

If the same N_0 initial observations in each population are used for both procedures we have $N_i^{(R)} + 1 \geq N_i^{(E)}$ (that is $P_R(h^*)$ cannot save more than one observation per population over P_E). We thus have

Proposition 6. $EN_i^{(R)} + 1 \geq EN_i^{(E)}$, $i = 1,\ldots,k$.

The relation between $EN_i^{(R)}$ and $EN_i^{(E)}$ depends on the unknown values of the variances. If N_0 remains fixed, letting $\sigma_i^2 \to 0$ we have $N_i^{(R)} \to N_0$ and $N_i^{(E)} \to N_0 + 1$ and thus for small values of σ_i^2 we have

$$EN_i^{(R)} < EN_i^{(E)}$$

Letting $\sigma_i \to \infty$ we see that the relation $h^{(R)} \geq h^{(E)}$ implies

$$EN_i^{(R)} \geq EN_i^{(E)}$$

for large values of σ_i^2.

4.3 Comparison of P(CS)

We compare the probability of correct selection of P_R and P_E under the least favorable configuration. For the procedure P_E we have $P(CS|P_E) = P^*$ for any values of the variances, while as follows clearly from previous sections $P(CS|P_R) \to 1$ as $\sigma_i \to 0$, $i = 1,\ldots,k$, for any fixed $\delta^* > 0$ so that the P(CS) of P_R will exceed that of P_E when variances are small.

Similarly when $N_0 \to \infty$, $\delta^* > 0$ fixed then $P(CS|P_R) \to 1$ and we conclude that for large N_0, P_R will have a larger P(CS) than P_E for any fixed configuration of the rest of the parameters.

ACKNOWLEDGEMENT

This research was supported in part by NSF Grant GP 24438 at Cornell University.

The author is grateful to Professor R.E. Bechhofer for introducing him to the subject and to the referees for many useful comments.

BIBLIOGRAPHY

Dudewicz, E.J. (1971). Nonexistence of a single-sample selection procedure whose P(CS) is independent of the variances. S. Afric. Statist. J. 5, 37-39.

Dudewicz, E.J. and Dalal, S.R. (1975). Allocation of observations in ranking and selection with unknown variances. Sankhyā B 37, 28-78.

Slepian, D. (1962). The one-sided barrier problem for Gaussian noise. Bell System Tech. J. 41, 463-501.

Stein, C. (1945). A two-sample test for a linear hypothesis whose power is independent of the variance. Ann. Math. Statist. 16, 243-58.

Received December, 1977; Retyped March, 1978.

Refereed by A. C. Tamhane, Northwestern University, Illinois and E. J. Dudewicz, Ohio State University, Columbus Ohio.

Recommended by S. S. Gupta, Purdue University, W. Lafayette Indiana.