On Swanepoel's Elimination Procedures

by

Raymond J. Carroll*

Department of Statistics
University of North Carolina at Chapel Hill

Institute of Statistics Mimeo Series #1074

June, 1976

* This research was supported by the Air Force Office of Scientific Research under Grant No. AFOSR-75-2796.

On Swanepoel's Elimination Procedures

by

Raymond J. Carroll*
University of North Carolina at Chapel Hill

SUMMARY

Swanepoel (1976) has developed classes of sequential ranking and selection procedures which have the elimination property and which asymptotically satisfy certain probability requirements. Under a variety of circumstances, this note obtains the asymptotic distributions of his stopping times. In the indifference zone formulation, it is shown that his procedures are infinitely more efficient than those appearing in the literature except possibly in circumstances closely related to the least favorable configuration.

* This research was supported by the Air Force Office of Scientific Research under Grant No. AFOSR-75-2796.

Introduction

Swanepoel (1976) has developed classes of sequential ranking and selection procedures (indifference zone and subset selection formulations) which have the elimination property and which asymptotically satisfy certain probability requirements. Under a variety of circumstances, this note obtains the asymptotic distributions of a class of stopping times which includes the Swanepoel procedures. We show that his procedures are infinitely more efficient than those appearing in the literature except possibly in circumstances where the least favorable configuration obtains.

Elimination rules are of obvious importance in the problem of selecting the best of k populations because they allow the experimenter the option of removing obviously inferior populations (compare the non-elimination procedures in Geertsema (1972) and Bechhofer, Kiefer and Sobel (1968)). Swanepoel's paper is noteworthy in that it is the first paper to adequately treat this topic, the most practical one in sequential ranking theory. However, he does not treat the topic of how many observations the user can expect to save by using the elimination feature (i.e., the distributions of the stopping times), and it is to this problem that the current paper is addressed.

We deal only with the case of selecting the better of two populations for two reasons. First, the extension to three or more populations is straightforward (one of the advantages of studying elimination procedures) but notationally cumbersome. Second, the savings available from an early elimination of the inferior populations when $k = 2$ will be illustrative and quite striking. This is especially in comparison to the Robbins, Sobel and Starr (1968) and Geertsema (1972) procedures which continue to take observations even when the decision is obvious.

In the general problem, we have k populations $\pi_1, \ldots, \pi_k$ and a sample $X_{i1}, X_{i2}, \ldots$ available from π_i. One forms location-scale equivariant statistics T_{in} from π_i based on $X_{i1}, \ldots, X_{in}$ (i.e., $T_{in}(\xi(X_{i1}-\theta), \ldots, \xi(X_{in}-\theta)) = \xi(T_{in}-\theta)$) and location invariant, scale equivariant variance estimates σ_{in}^2. The assumptions are the existence of constants $\mu_1, \ldots, \mu_k, \sigma_1^2, \ldots, \sigma_k^2$ such that

(1.1a) $T_{in} - \mu_i = n^{-1} \sum_{j=1}^{n} Y_{ij} + R_{in}$, where $Y_{i1}, \ldots, Y_{in}$ are i.i.d. mean zero, variance σ_i^2 random variables with $E|Y_{i1}^3| < \infty$

(1.1b) $$n^{\frac{1}{2}} R_{in} \to 0 \quad \text{(a.s.)}$$

(1.1c) $$\sigma_{in}^2 = \sigma_i^2 + o\left((\log_2 n)^{-\frac{1}{2}}\right) \quad \text{(a.s.)}.$$

Assume without loss of generality that $\mu_1 \le \ldots \le \mu_k$ and that it is desired to select the population with the largest value of μ_i. The indifference zone approach (we consider subset selection in the last section) assumes $\mu_k - \mu_{k-1} \ge \Delta$ and defines $r = r(\Delta) \to \infty$ as $\Delta \to 0$ (Swanepoel chooses $r = [1/\Delta]$, where $[\circ]$ is the greatest integer function). Let g be a given nonnegative function. Then Swanepoel's procedure eliminates population π_i at the n^{th} stage if there is a population π_j still in contention at the n^{th} stage for which

$$T_{jn} - T_{in} \ge r^{\frac{1}{2}} g(n/r) (\sigma_{in}^2 + \sigma_{jn}^2)^{\frac{1}{2}}/n - \Delta.$$

Sampling continues until only one population is left. If $W(t)$ is the standard Wiener process, and CS denotes a correct selection, Swanepoel shows

$$(1.2)\quad \liminf_{\Delta\to 0} \Pr\{CS\} \geq 1 - \Pr\{W(t) \geq g(t) \text{ for some } t > 0\}(k-1).$$

If $k = 2$, the stopping time N_1 becomes

$$N_1 = \inf\left\{n\geq 1: \begin{array}{c} T_{2n}-T_{1n}-(\mu_2-\mu_1) \geq r^{\frac12}g(n/r)(\sigma_{1n}^2+\sigma_{2n}^2)^{\frac12}/n-\Delta-(\mu_2-\mu_1) \\ \text{or} \\ T_{2n}-T_{1n}-(\mu_2-\mu_1) \leq -r^{\frac12}g(n/r)(\sigma_{1n}^2+\sigma_{2n}^2)^{\frac12}/n+\Delta-(\mu_2-\mu_1) \end{array}\right\}$$

Now define $T_n = T_{2n} - T_{1n} - (\mu_2-\mu_1)$, $\mu = \mu_2 - \mu_1 \geq \Delta$ and $\sigma_n^2 = \sigma_{1n}^2 + \sigma_{2n}^2$. Then the stopping rule takes the form

$$N = \inf\left\{n\geq 1: \begin{array}{c} T_n - r^{\frac12}g(n/r)\sigma_n/n + (\mu+\Delta) \geq 0 \\ \text{or} \\ T_n + r^{\frac12}g(n/r)\sigma_n/n + (\mu-\Delta) \leq 0 \end{array}\right\},$$

where it is assumed in this paper that

$$(1.3a)\quad T_n = n^{-1}\sum_{i=1}^{n} Y_i + R_n, \quad n^{\frac12}R_n \to 0 \text{ (a.s.)}, \quad EY_1 = 0,$$

$$EY_1^2 = \sigma^2 \text{ and } E|Y_1^3| < \infty.$$

$$(1.3b)\quad \hat{\sigma}_n^2 = \sigma^2 + \left((\log_2 n)^{-\frac12}\right) \text{ (a.s.)}.$$

In Section 2, we investigate the stopping time N as $\Delta \to 0$ $(r \to \infty)$, relate this back to the selection problem in Section 3, and discuss subset selection in Section 4.

Asymptotic Distributions

To begin with we will write $Y_i = \psi(X_i)$ to emphasize the dependence on the original sample $X_1, X_2, \ldots$. It is not surprising that since the $\Pr\{CS\}$ is related to the Wiener process W, the distribution of N is also related to that of W. The proof of Theorem 1 gets its inspiration from work of Lai (1975) and, while self-contained, depends heavily on the approximations of Swanepoel.

Theorem 1 Suppose that $g(0) > 0$ and that g is a nondecreasing continuous function for $t \geq 0$ satisfying

(2.1z) $$t^{-\frac{1}{2}}g(t) \uparrow \infty \text{ as } t \downarrow 0 \text{ and as } t \uparrow \infty.$$

(2.1b) $$\int_{0^+}^{\infty} t^{-3/2}g(t)\exp\{-g^2(t)/32t\}dt < \infty.$$

Then, for $t > 0$, if $r^{\frac{1}{2}}(\mu+\Delta) \to \infty$, $\Pr\{N/r>t\} \to 0$. If $r^{\frac{1}{2}}(\mu+\Delta) \to c_1$, $r^{\frac{1}{2}}(\mu-\Delta) \to c_0$ $(0 \leq c_0 \leq c_1 < \infty)$,

$$\Pr\{N/r > t\} \to H(t) = \Pr\left\{\begin{array}{l} \max\limits_{0\leq s\leq t} \left(W(s)-g(s)+sc_1/\sigma\right) \leq 0 \text{ and} \\ \\ \min\limits_{0\leq s\leq t} \left(W(s)+g(s)+sc_0/\sigma\right) \geq 0 \end{array}\right\}.$$

Proof of Theorem 1 Let $a > 0$ be arbitrarily small.

$$\Pr\{N/r>t\} = \Pr\left\{\begin{array}{ll}\max\limits_{n\le[rt]}\left(T_n - r^{\frac12}g(n/r)\sigma_n/n+(\mu+\Delta)\right) \le 0 & \text{and} \\ \min\limits_{n\le[rt]}\left(T_n + r^{\frac12}g(n/r)\sigma_n/n+(\mu-\Delta)\right) \ge 0 & \end{array}\right\}$$

$$= \Pr\left\{\begin{array}{ll}\max\limits_{0\le s\le t}\left(T_{[rs]} - r^{\frac12}g([rs]/r)\sigma_{[rs]}/[rs]+(\mu+\Delta)\right) \le 0 & \text{and} \\ \min\limits_{0\le s\le t}\left(T_{[rs]} + r^{\frac12}g([rs]/r)\sigma_{[rs]}/[rs]+(\mu-\Delta)\right) \ge 0 & \end{array}\right\}$$

$$\le \Pr\left\{\begin{array}{ll}\max\limits_{a\le s\le t}\left(T_{[rs]} - r^{\frac12}g([rs]/r)\sigma_{[rs]}/[rs]+(\mu+\Delta)\right) \le 0 & \text{and} \\ \min\limits_{a\le s\le t}\left(T_{[rs]} + r^{\frac12}g([rs]/r)\sigma_{[rs]}/[rs]+(\mu-\Delta)\right) \ge 0 & \end{array}\right\}.$$

Let $W_r(s) = \sigma^{-1}r^{-\frac12}[rs]T_{[rs]}$. Then the last expression is

$$\Pr\left\{\begin{array}{ll}\max\limits_{a\le s\le t}\left(W_r(s) - \sigma_{[rs]}g([rs]/r)/\sigma+[rs](\mu+\Delta)\sigma^{-1}r^{-\frac12}\right) \le 0 & \text{and} \\ \min\limits_{a\le s\le t}\left(W_r(s) + \sigma_{[rs]}g([rs]/r)/\sigma+[rs](\mu-\Delta)\sigma^{-1}r^{-\frac12}\right) \ge 0 & \end{array}\right\}.$$

Since $W_r \Rightarrow W$, $\sigma_{[rs]}/\sigma \to 1$ (a.s.), and g is continuous, the Theorem is proved if $r^{\frac12}(\mu+\Delta) \to \infty$, while if $r^{\frac12}(\mu+\Delta) \to c_1$, $r^{\frac12}(\mu-\Delta) \to c_0$, the last probability must converge for fixed $a > 0$ to

$$H(a,t) = \Pr\left\{\begin{array}{ll}\max\limits_{a\le s\le t}\left(W(s) - g(s)+sc_1/\sigma\right) \le 0 & \text{and} \\ \min\limits_{a\le s\le t}\left(W(s) + g(s)+sc_0/\sigma\right) \ge 0 & \end{array}\right\}.$$

Since $H(\cdot,t)$ is continuous in a, $\overline{\lim}\Pr\{N/r>t\} \le H(0,t)$. Hence, it suffices to show that as $r \to \infty$ and $a \to 0$, if $r^{\frac12}(\mu\pm\Delta)$ have finite limits,

$$\Pr\left\{\begin{array}{ll}\max\limits_{n\le[ra]}\left(T_n - r^{\frac12}g(n/r)\sigma_n/n+(\mu+\Delta)\right) \le 0 & \text{and} \\ \min\limits_{n\le[ra]}\left(T_n + r^{\frac12}g(n/r)\sigma_n/n+(\mu-\Delta)\right) \ge 0 & \end{array}\right\} \to 1.$$

We only show that

$$G(a) = \Pr\{\max_{1\le n\le ra} \left(T_n - r^{\frac{1}{2}}g(n/r)\sigma_n/n + (\mu+\Delta)\right) > 0\} \to 0,$$

as the other case is much easier. Now, for all $\varepsilon > 0$ there exists n_0 such that $\Pr\{\sigma_n^2 < n_0$ for all $n \ge 1\} > 1 - \varepsilon$. Taking $n_0 = 1$ without loss of generality, it thus suffices to show that

$$G(a) = \Pr\{\max_{1\le n\le ra} \left(\sigma_n^{-1}T_n - r^{\frac{1}{2}}g(n/r)/n + (\mu+\Delta)\right) > 0\} \to 0.$$

Now

$$\begin{aligned} G(a) \le\ & \Pr\{\sigma_n^{-1}\textstyle\sum_1^n\psi(X_i) - W(n) \ge \left(r^{\frac{1}{2}}g(n/r) - n(\mu+\Delta)\right)/4 \text{ for some } 1 \le n \le ra\} \\ & + \Pr\{\sigma_n^{-1}nR_n \ge \left(r^{\frac{1}{2}}g(n/r) - n(\mu+\Delta)\right)/4 \text{ for some } 1 \le n \le ra\} \\ & + \Pr\{W(n) \ge \left(r^{\frac{1}{2}}g(n/r) - n(\mu+\Delta)\right)/2 \text{ for some } 1 \le n \le ra\} \\ =\ & P_2 + P_3 + P_4 \quad \text{(say)}. \end{aligned}$$

We can make the substitution $n(\Delta+\mu) \sim nc_1/r^{\frac{1}{2}}$ and note that for $1 \le n \le ra$

$$\begin{aligned} r^{\frac{1}{2}}g(n/r) - n(\mu+\Delta) &\sim r^{\frac{1}{2}}\left(g(n/r) - nc_1/r\right) \\ &\ge r^{\frac{1}{2}}\left(g(n/r) - ac_1\right). \end{aligned}$$

Since g is increasing and $g(0) > 0$, we then obtain

$$r^{\frac{1}{2}}g(n/r) - n(\mu+\Delta) \ge r^{\frac{1}{2}}g(n/r)/2 \tag{2.2}$$

for a sufficiently small. Thus we can replace the left side of (2.2) by its right side in P_2, P_3 and P_4. The rest of the proof follows from Swanepoel's equations (8) - (12) which we will sketch. First note that from Jain, Jogdeo and Stout (1975)

$$\sigma^{-1} \sum_1^n \psi(X_i) - W(n) = o(n^{\frac{1}{2}}) \quad \text{(a.s.)}.$$

Since $x^{-\frac{1}{2}}g(x) \uparrow \infty$ as $x \to 0$, we have

$$\begin{aligned} P_2 &\le \Pr\{(\sigma_n^{-1}\textstyle\sum_1^n\psi(X_i)-W(n))/n^{\frac{1}{2}} \ge a^{-\frac{1}{2}}g(a)/16 \text{ for some } 1\le n\le ar\} \\ &\le \Pr\{(\sigma^{-1}\textstyle\sum_1^n\psi(X_i)-W(n))/n^{\frac{1}{2}} \ge a^{-\frac{1}{2}}g(a)/32 \text{ for some } n\ge 1\} \\ &\quad + \Pr\{(\sigma_n^{-1}-\sigma^{-1})\textstyle\sum_1^n\psi(X_i)/n^{\frac{1}{2}} \ge a^{-\frac{1}{2}}g(a)/32 \text{ for some } n\ge 1\} \\ &\to 0 \quad (\text{as } a \to 0). \end{aligned}$$

Also, since $n^{\frac{1}{2}}R_n \to 0$ (a.s.),

$$\begin{aligned} P_3 &\le \Pr\{\sigma_n^{-1}n^{\frac{1}{2}}R_n \ge a^{-\frac{1}{2}}g(a)/16 \text{ for some } n\ge 1\} \\ &\to 0 \quad (\text{as } a \to 0). \end{aligned}$$

Finally, since $\alpha^{-\frac{1}{2}}W(\alpha t)$ is a standard Wiener process and applying Kolmogorov's test (see Itô and McKean (1965), page 34) we get

$$\begin{aligned} P_4 &\le \Pr\{W(t) \ge r^{\frac{1}{2}}g(t/n)/4 \text{ for some } 0<t<ar\} \\ &= \Pr\{W(t) \ge g(t)/4 \text{ for some } 0<t<a\} \\ &\to 0 \quad (\text{as } a \to 0). \end{aligned}$$

<u>Corollary 1</u> Suppose that Theorem 1 holds and in addition, that

(2.3a) $\qquad g(s)s^{-1+\varepsilon} \to 0$ as $s \to \infty$ for some $\frac{1}{2} > \varepsilon > 0$.

(2.3b) $\qquad g(s)(s \log|\log s|)^{-\frac{1}{2}} \to \infty$ as $s \to 0$.

Here log is the natural logarithm. Then $H(t) \to 1$ as $t \to 0$ and, if $c_1 > 0$, $H(t) \to 0$ as $t \to \infty$. If (2.3b) holds when $s \to \infty$, then if $c_0 = c_1 = 0$,

$$\liminf_{t\to\infty} H(t) > 0.$$

The proof of Corollary 1 is a consequence of the Law of the Iterated Logarithm for the Wiener process (Breiman (1968), page 266). The conclusion is that $1 - H(t)$ is a distribution function only if $c_1 > 0$. This has important consequences in Swanepoel's procedure $(r = 1/\Delta)$ when $\mu = \Delta$, consequences which are discussed in detail in Section 3.

While it is clear that the proof of Theorem 1 breaks down $(N/r \xrightarrow{P} 0)$ if $r^{\frac{1}{2}}(\mu+\Delta) \to \infty$, it is not clear why this should be so. The key is that if $r^{\frac{1}{2}}\Delta = O(1)$ (true in all our applications), then $r^{\frac{1}{2}}(\mu+\Delta) \to \infty$ implies that $r^{\frac{1}{2}}(\mu-\Delta) \to \infty$ and one begins to make correct selections with probability one and the problem becomes that of a *one*-sided stopping rule as opposed to the original two-sided problem.

<u>Lemma 1</u> Suppose $r^{\frac{1}{2}}(\mu-\Delta) \to \infty$. Then $\Pr\{CS\} \to 1$. If for some $\varepsilon > 0$, $r^{\frac{1}{2}-\varepsilon}(\mu+\Delta) \to \infty$, then $M(\mu+\Delta)/r^{\frac{1}{2}} \xrightarrow{P} \sigma g(0)$.

<u>Proof of Lemma 1</u> (μ fixed) It is clear that $N \to \infty$ (a.s.) implies $\Pr\{CS\} \to 1$. This means that $\Pr\{N \neq M\} \to 0$, where

$$M = \inf\{n : T_n \geq r^{\frac{1}{2}} g(n/r)\sigma_n/n - (\mu+\Delta)\}.$$

Since $T_n \to 0$ (a.s.), $\sigma_n \to \sigma$ (a.s.) and $M/r \xrightarrow{P} 0$ by Theorem 1, an analysis along the lines of Chow and Robbins (1965) yields $r^{\frac{1}{2}} g(M/r)\sigma_M/M(\mu+\Delta) \xrightarrow{P} 1$. Since $M/r \xrightarrow{P} 0$ and g is continuous, the proof is complete.

<u>Proof of Lemma 2</u> $(\mu \to 0)$. By inspection or by Swanepoel's Theorem 3.1,

$\Pr\{\text{Incorrect Selection}\} \leq \Pr\{-T_n \geq r^{\frac{1}{2}} g(n/r)\sigma_n/n + (\mu-\Delta) \text{ for some } n \geq 1\}$.

Now, for all $\varepsilon > 0$ there exists $n_0 > 0$ such that $\Pr\{\sigma_n^2 \leq n_0 \text{ for all } n\geq 1\} \geq 1 - \varepsilon$, so to prove $\Pr\{CS\} \to 1$ it suffices to show

$$\Pr\{-\sigma_n^{-1} T_n \geq r^{\frac{1}{2}} g(n/r)/n + (\mu-\Delta)/n_0 \text{ for some } n\geq 1\} \to 0.$$

Setting $n_0 = 1$, from the argument of Theorem 1 we see that as $a_1 \to 0$, $a_2 \to \infty$ the above probability is

$$\Pr\{-\sigma_n^{-1} T_n \geq r^{\frac{1}{2}} g(n/r)/n + (\mu-\Delta)/n_0 \text{ for some } a_1 r \leq n \leq a_2 r\} + o(1).$$

Fix a_1, a_2 so that the $o(1)$ term above is arbitrarily small. Then the probability becomes

$$\Pr\{\sigma_{[rt]}^{-1}[rt] r^{-\frac{1}{2}} T_{[rt]} + g\left(\frac{[rt]}{r}\right) + \frac{(\mu-\Delta)[rt]}{\sigma_{[rt]} r^{\frac{1}{2}}} \leq 0 \text{ for some } a_1 \leq t \leq a_2\} \to 0$$

$$(\text{as } r \to \infty)$$

since $r^{\frac{1}{2}}(\mu-\Delta) \to \infty$. Thus, $\Pr\{CS\} \to 1$ and as before it suffices to consider M. It will be helpful to denote the dependence on μ by $M(\mu)$. We first want to show that $\log_2 M(\mu)/(M(\mu)\mu^2) \xrightarrow{p} 0$; this would imply that $T_{M(\mu)}/(\mu+\Delta) \xrightarrow{p} 0$ and hence that

$$\frac{r^{\frac{1}{2}} g(M(\mu)/r)\sigma}{M(\mu)(\mu+\Delta)} \xrightarrow{p} 1.$$

Since $M(\mu)/r \xrightarrow{p} 0$, this argument would then yield $(\mu+\Delta)M(\mu)/r^{\frac{1}{2}} \xrightarrow{p} \sigma g(0)$. We have already shown that $M(1)/r^{\frac{1}{2}} \xrightarrow{p} \sigma g(0)$ since $(1+\Delta)M(1)/r^{\frac{1}{2}} \xrightarrow{p} \sigma g(0)$. Define $\eta_1 = (\mu+\Delta)^{\frac{1}{4}-\varepsilon_1/2}$, where $\varepsilon_1 > 0$ is very small. $M(\eta_1) \geq M(1)$ so that

$$\Pr\{2M(\eta_1)\eta_1^2 \geq \sigma g(0) r^{\frac{1}{2}} (\mu+\Delta)^{\frac{1}{2}-\varepsilon_1}\} \to 1.$$

Since $\log_2 M(\eta_1) \leq \log_2 r$, this gives

$$\frac{\log_2 M(\eta_1)}{M(\eta_1)\eta_1^2} \xrightarrow{p} 0.$$

Hence, $\eta_1 M(\eta_1)/r^{\frac{1}{2}} \xrightarrow{p} \sigma g(0)$. Now let $\eta_2 = (\mu+\Delta)^{\frac{1}{2}-\varepsilon_2/2}$. Then $M(\eta_2) \geq M(\eta_1)$ so that

$$\Pr\{2M(\eta_2)\eta_2^2 \geq \sigma g(0) r^{\frac{1}{2}} (\mu+\Delta)^{3/4-\varepsilon_2+\varepsilon_1/2}\} \to 1.$$

One continues the process, picking $\eta_3 = (\mu+\Delta)^{3/4-\varepsilon_3/2}$ and carefully choosing $\varepsilon_1, \varepsilon_2, \ldots$.

We note that for technical reasons the approach of Lemma 1 cannot handle the case $\mu + \Delta = r^{-\frac{1}{2}}(\log_2 r)$, and we can see no way to get around this. However, in obtaining the asymptotic distribution of N, only the result of Lemma 1 is necessary. Recall again that if $r^{\frac{1}{2}}(\mu-\Delta) \to \infty$, we are essentially dealing with a one-sided stopping rule, and it was this that enabled us to show $M(\mu+\Delta)/r^{\frac{1}{2}} \xrightarrow{p} \sigma g(0)$. From now on, let us assume

(2.4) $n^{\frac{1}{2}}(\sigma_n^2-\sigma^2)$ has a limiting distribution F and is uniformly continuous in probability (Anscombe (1952)).

<u>Theorem 2</u> Assume Lemma 1 holds and that as $x \to 0$, $x^{-\frac{1}{2}}\big(g(x)-g(0)\big) \to 0$. Further assume that ψ is bounded and that $n^{\frac{1}{2}}(\sigma_n^2-\sigma_{n-1}^2) \to 0$ (a.s.). Defining $x_0^2(r,\mu) = r^{\frac{1}{2}}\sigma g(0)/(\mu+\Delta)$, we have

$$\{r^{\frac{1}{2}}\sigma g(0) - (\Delta+\mu)N\}/x_0(r,\mu) = N^{\frac{1}{2}}\big(T_N - (\mu+\Delta)(\sigma_N^2-\sigma^2)/2\sigma^2\big) + o_p(1).$$

<u>Proof of Theorem 2</u> Recall that $(\mu+\Delta)M(\mu)/r^{\frac{1}{2}} \to \sigma g(0)$. Since $\Pr\{CS\} \to 1$, it again suffices to consider M. Then

$$M^{\frac{1}{2}}T_M \geq (M/r)^{-\frac{1}{2}}g(M/r)\sigma_M - (\Delta+\mu)M^{\frac{1}{2}}$$

$$= (M/r)^{-\frac{1}{2}}\big(g(M/r)-g(0)\big)\sigma_M + (M/r)^{-\frac{1}{2}}g(0)(\sigma_M-\sigma) + (M/r)^{-\frac{1}{2}}g(0)\sigma - (\Delta+\mu)M^{\frac{1}{2}} .$$

Now, $r/M \xrightarrow{p} \infty$ so that

$$M^{\frac{1}{2}}T_M - (r/M)^{\frac{1}{2}}g(0)(\sigma_M-\sigma) \geq \big(r^{\frac{1}{2}}g(0)\sigma - (\Delta+\mu)M\big)/M^{\frac{1}{2}} + o_p(1).$$

Now, from (2.4) and Anscombe (1952), $M^{\frac{1}{2}}(\sigma_M^2-\sigma^2)$ and hence $M^{\frac{1}{2}}(\sigma_M-\sigma)$ have limiting normal distributions. Thus,

$$g(0)(\sigma_M-\sigma)\{(r/M)^{\frac{1}{2}} - (\mu+\Delta)M^{\frac{1}{2}}/\sigma g(0)\} \xrightarrow{p} 0.$$

Similarly,

$$(\mu+\Delta)M^{\frac{1}{2}}\{(\sigma_M-\sigma)/\sigma - (\sigma_M^2-\sigma^2)/2\sigma^2\} \xrightarrow{p} 0,$$

which shows that

$$(2.5) \quad M^{\frac{1}{2}}\big(T_M - (\mu+\Delta)(\sigma_M^2-\sigma^2)/2\sigma^2\big) \geq \big(r^{\frac{1}{2}}g(0)\sigma - (\Delta+\mu)M\big)/x_0(r,\mu) + o_p(1).$$

For the opposite inequality, note that

$$M^{\frac{1}{2}}T_{M-1} < (Mr)^{\frac{1}{2}}g\big((M-1)/r\big)\sigma_{M-1}/(M-1) - (\Delta+\mu)M^{\frac{1}{2}} .$$

Then, since ψ is bounded,

$$M^{\frac{1}{2}}(T_{M-1} - T_M) = M^{-\frac{1}{2}}T_{M-1} - M^{-\frac{1}{2}}\psi(X_M) + o_p(1) = o_p(1).$$

By assumption, $M^{\frac{1}{2}}(\sigma_M^2-\sigma_{M-1}^2) = o_p(1)$, so that the opposite to (2.5) is obtained, completing the proof.

To summarize this section, we have found the asymptotic distribution of N for the two population indifference zone selection problem in the two cases $r^{\frac{1}{2}}(\mu_2-\mu_1+\Delta) \to c_1 > 0$ and $r^{\frac{1}{2}-\varepsilon}(\mu_2-\mu_1+\Delta) \to \infty$. These turn out to be mathematically different because the latter case reduces to a one-sided problem. If $r^{\frac{1}{2}}(\mu_2-\mu_1+\Delta) \to 0$, partial results have been obtained.

Asymptotic Comparisons

The rule proposed by Geertsema (1972), which like that of Swanepoel is essentially nonparametric but does not possess the elimination feature and assumes $\sigma_1^2 = \sigma_2^2 = \sigma_0^2$, reduces for $k = 2$ to

$$N^* = \inf\{n \geq 5 : n \geq \beta(P^*)(\sigma_{1n}^2+\sigma_{2n}^2)/2\Delta^2\},$$

where $\beta(P^*)$ are constants defined for general k by

$$\int \Phi^{k-1}\big(x + \beta(P^*)\big)d\Phi(x) = P^*,$$

and if $k = 2$, $\beta(P^*)^2 \sim -4 \log(1-P^*)$ as $P^* \to 1$. One shows that, *independent of* $\mu_2 - \mu_1$,

$$\Delta^2 N^*/\beta(P^*) \xrightarrow{p} \sigma_0^2 .$$

Let us first assume that $r^{\frac{1}{2}}(\mu_2-\mu_1) \to d_0$ $(0 < d_0 \leq \infty)$ and that, as is done by Swanepoel, $r = 1/\Delta$. Then Lemma 1 and Theorem 1 show that, for any $\varepsilon > 0$, $N/r^{1+\varepsilon} = \Delta^{1+\varepsilon}N \xrightarrow{p} 0$, leading for fixed P^* to the conclusion that the elimination rule is infinitely more efficient, i.e.,

$$N^*/N \xrightarrow{p} \infty.$$

In support of this conclusion, Swanepoel has found a function $g_*(t)$ given by

$$g_*(t) = (t+\tfrac{1}{4})^{\frac{1}{2}}\{\log(t+\tfrac{1}{4}) - 2\log(1-P^*)\}^{\frac{1}{2}}$$

for which as $P^* \to 1$, $N^*/N \xrightarrow{P} 4$ if $\mu_2 - \mu_1 = \Delta$, while $N^*/N \xrightarrow{P} \infty$ if $P^* \to 1$ and then $(\mu_2-\mu_1)/\Delta \to \infty$.

It is in the case $(\mu_2-\mu_1)/\Delta^{\frac{1}{2}} \to 0$ that the comparison of the elimination rule $(r = 1/\Delta)$ with Geertsema's rule is not clear; this is due to the fact that $1 - H(t)$ is not then a probability distribution function. One might conjecture as in Lemma 1 that $(\mu_2-\mu_1+\Delta)N/r^{\frac{1}{2}}$ has a limiting distribution, but we have been unable to show this in the least favorable configuration $\mu_2 - \mu_1 = \Delta = r^{-1}$.

Now suppose, however, that we define $r = ([1/\Delta])^2$. Then, since $\mu_2 - \mu_1 \geq \Delta$ the case $r^{\frac{1}{2}}(\mu_2-\mu_1) \to 0$ is impossible and there are only two important cases (i) $r^{\frac{1}{2}}(\mu_2-\mu_1) \to d_0$ $(1 \leq d_0 < \infty)$, i.e., $\mu_2 - \mu_1 \sim d_0\Delta$ and (ii) $r^{\frac{1}{2}}(\mu_2-\mu_1) \to \infty$. In the first case, Theorem 1 and Corollary 1 tell us that $N/r \sim \Delta^2 N$ has a limiting distribution and, as $\Delta \to 0$,

$$\Pr\{N^*<N\} \sim \Pr\{\Delta^2 N > \beta(P^*)\sigma_0^2\}$$
$$\to \Pr\left\{\begin{array}{l} \max\limits_{0\leq s\leq\beta(P^*)\sigma_0^2} \left(W(s) - g(s) + s(1+d_0)/2\sigma_0^2\right) \leq 0 \quad \text{and} \\ \min\limits_{0\leq s\leq\beta(P^*)\sigma_0^2} \left(W(s) + g(s) + s(d_0-1)/2\sigma_0^2\right) \geq 0 \end{array}\right\}.$$

As σ_0^2 becomes large, this approaches

$$1 - \Pr\{-g(s) \leq W(s) \leq g(s) \quad \text{for all} \quad 0 \leq s < \infty\},$$

and when one recalls that

$$\Pr\{W(s) \le g(s) \quad \text{for all} \quad 0 \le s < \infty\} = P^*,$$

we see that the elimination procedure N would have high probability of beating the noneliminating procedure N^* even in the least favorable configuration. In case (ii) above, a simple analysis shows that $N^*/N \xrightarrow{p} \infty$ as before. Similarly, for fixed μ and Δ, as $P^* \to 1$, $T_N \to 0$ (a.s.) so that

$$\frac{r^{\frac{1}{2}} g(N/r)}{N} \to \frac{(\mu+\Delta)}{\sqrt{2}\,\sigma_0} .$$

Using g_* suggested above, this leads to $N/\beta(P^*) \xrightarrow{p} \sigma_0^2(\mu+\Delta)^{-2}$ so that as $P^* \to 1$, $N^*/N \xrightarrow{p} 4$ as as $P^* \to 1$ and then $\mu/\Delta \to \infty$ we obtain $N^*/N \xrightarrow{p} \infty$. We note in passing that for this choice of r, the correct selection probabilities are asymptotically bounded below by P^*.

Truncated Subset Selection

Swanepoel proposes a subset selection rule (see Gupta (1965)) in cases where at most r observations can be taken. Here $\Delta = 0$ and sampling proceeds as in the indifference zone problem with the exception that at the r^{th} stage sampling is discontinued and all remaining populations are declared "best". Swanepoel shows

$$\liminf_{r\to\infty} \Pr\{CS\} \ge 1 - \Pr\{W(t) \ge g(t) \quad \text{for some} \quad 0 < t < 1\}(k-1).$$

Theorem 1 holds for $0 \le t \le 1$, while $\Pr\{N>r\} = 0$, and both Lemma 1 and Theorem 2 hold as stated.

Acknowledgement The author thanks Professor Swanepoel for sending a preprint of his paper, and for some interesting discussions.

REFERENCES

ANSCOMBE, J. F. (1952). Large sample theory of sequential estimation. *Proc. Camb. Phil. Soc.*, *48*, 600-617.

BECHHOFER, R. E., KIEFER, J. and SOBEL, M. (1968). *Sequential Identification and Ranking Procedures*. University of Chicago Press.

BREIMAN, L. (1968). *Probability*. Addison-Wesley.

CHOW, Y. S. and ROBBINS, H. (1965). On the asymptotic theory of fixed-width sequential confidence intervals for the mean. *Ann. Math. Statist.*, *36*, 457-462.

GEERTSEMA, J. C. (1972). Nonparametric sequential procedures for selecting the best of k populations. *J. Am. Statist. Assoc.*, *67*, 614-616.

GUPTA, S. S. (1965). On some multiple decision rules. *Technometrics*, *7*, 225-245.

JAIN, N. C., JODGEO, K. and STOUT, W. F. (1975). Upper and lower functions for martingales and mixing processes. *Ann. Prob.*, *3*, 119-145.

ITÔ, K., and MCKEAN, H. P. (1965). *Diffusion Processes and Their Sample Paths*. Springer.

LAI, T. L. (1975). A note on first exist times with applications to sequential analysis. *Ann. Statist.*, *3*, 999-1005.

ROBBINS, H., SOBEL, M. and STARR, N. (1968). A sequential procedure for selecting the largest of k means. *Ann. Math. Statist.*, *39*, 88-92.

SWANEPOEL, J. W. H. (1976). Nonparametric elimination selection procedures based on robust estimators. Unpublished manuscript.