DISCUSSION PAPER NO. 185

# ASYMPTOTIC EFFICIENCY OF BECHHOFER'S SINGLE-STAGE PROCEDURE RELATIVE TO A TWO-STAGE PROCEDURE FOR SELECTING THE LARGEST NORMAL MEAN

by

Ajit C. Tamhane*/

November 25, 1975

*/ Department of Industrial Engineering and Management Science
Northwestern University
Evanston, Illinois

# TABLE OF CONTENTS

| | | Page |
|---|---|---|
| | Abstract | i |
| 1. | Introduction | 1 |
| 2. | Statement of the Problem | 3 |
| 3. | Preliminary Results | 7 |
| 4. | Main Results | 13 |
| 5. | Discussion | 18 |
| 6. | Suggestions for Future Research | 20 |
| | Acknowledgements | 21 |
| | Bibliography | 22 |

Abstract

For the problem of selecting the normal population having the largest mean, the relative efficiency of Bechhofer's [2] single-stage procedure is studied with respect to a permanent elimination type two-stage procedure [1,6]; the relative efficiency being defined as the ratio of the expected total sample size required by the latter procedure to that required by the former to guarantee a specified probability requirement. The two-stage procedure is considered with two different design criteria: a minimax criterion and a restricted minimax criterion. It is shown that the asymptotic (as $P^*$, the specified probability of a correct selection, tends to 1) relative efficiency of the single-stage procedure w.r.t. the minimax two-stage procedure is 1 at the equal means configuration. It is also shown that the asymptotic relative efficiency of the single-stage procedure w.r.t. the restricted minimax two-stage procedure is 1/4 at the least favorable configuration of means. The latter result implies that the unrestricted minimax two-stage procedure asymptotically possesses a Wald-Wolfowitz optimum property for the two population case.

Keywords: ranking and selection problems, single-stage and two-stage procedures, asymptotic relative efficiency, normal means problem.

IMS Subject Classifications: 62F07, 62L05

# 1. INTRODUCTION

Consider the problem of selecting the population associated with the largest mean from several normal populations which have a common known variance and suppose that the problem is to be formulated using the so-called indifference-zone approach. (All the wavy underlined terms are defined precisely in Section 2). A single-stage procedure which guarantees a specified requirement on the probability of a correct selection was proposed by Bechhofer [2]. A permanent elimination type two-stage procedure was proposed by Alam [1] for the same problem. Alam suggested that the design constants necessary to implement his two-stage procedure should be chosen to minimize the maximum of the expected total sample size associated with it; the maximum being taken over a specified subset (the so-called preference-zone) of the parameter space. In [6] we pointed out that if the true parameter configuration lies in the indifference-zone of the parameter space then Alam's suggestion might lead to an undesirable situation where the expected total sample size for the two-stage procedure exceeds the fixed total sample size required by Bechhofer's single-stage procedure to guarantee the same probability requirement. We recommended that the design constants for the two-stage procedure should be chosen to minimize the maximum taken over the entire parameter space of the expected total sample size associated with it. This ensures that for any parameter configuration, the expected total sample size for the two-stage procedure does not exceed the fixed total sample size required by Bechhofer's single-stage procedure to guarantee the same probability requirement.

In the present paper we are concerned with the study of the relative performances of the three procedures ([1],[2],[6]) discussed above. To compare the performances of two competing procedures, say $P_1$ and $P_2$, we consider the ratio of the expected total sample size required by $P_1$ to that required by $P_2$, both guaranteeing a specified probability requirement. Here the expected total sample sizes associated with $P_1$ and $P_2$ are computed at some specified parameter configuration of interest. This ratio is referred to as the relative efficiency (RE) of $P_1$ w.r.t. $P_2$. We define the asymptotic relative efficiency (ARE) of $P_1$ w.r.t. $P_2$ as the limit of the corresponding RE as $P^*$, the specified probability of a correct selection, tends to 1.

A detailed study of the ARE of Bechhofer's single-stage procedure w.r.t the two-stage procedures proposed in [1] and [6] is made in the present paper. Some numerical results for the case of small samples are also included. A very important result of our study is that the Alam's two-stage procedure and the open-ended fully sequential procedure (BKS) proposed by Bechhofer, Kiefer and Sobel [4] are asymptotically (as $P^* \to 1$) equivalent in the sense that the RE of Alam's procedure w.r.t BKS tends to 1, when the parameter configuration of interest is the so-called least favorable configuration (LFC). In particular, we show that for a two-population problem the two-stage procedure of Alam possesses a Wald-Wolfowitz optimal property (see [7]) asymptotically as $P^* \to 1$.

Although our calculations indicate that in the indifference-zone, the performance of Alam's procedure becomes gradually poorer compared to Bechhofer's single-stage procedure as $P^* \to 1$, we could not obtain a definitive result as obtained by Bechhofer [3] regarding a similar undesirable property associated with the Wald sequential probability ratio test (WSPRT).

## 2. STATEMENT OF THE PROBLEM

Let $\Pi_1, \Pi_2, \cdots, \Pi_k$ be $k \geqq 2$ normal populations with unknown means $\mu_1, \mu_2, \cdots, \mu_k$ and a common known variance $\sigma^2$. We define the parameter space $\Omega$ as the collection of all parameter vectors $\underline{\mu} = (\mu_1, \mu_2, \cdots, \mu_k)'$. Let $\mu_{[1]} \leqq \mu_{[2]} \leqq \cdots \leqq \mu_{[k]}$ denote the ordered values of the $\mu_i$ and let $\Pi_{(i)}$ be the population associated with $\mu_{[i]} (1 \leq i \leq k)$. We assume that the experimenter has no prior knowledge concerning the correct pairing between $\Pi_i$ and $\Pi_{(j)} (1 \leq i, j \leq k)$. If $\mu_{[k]} > \mu_{[k-1]}$ then we regard $\Pi_{(k)}$ as the "best" population. The experimenter's goal is to select the best population. The event of selection of the best population is referred to as a correct selection and is denoted by CS.

In the indifference-zone approach to this selection problem, we assume that two constants $\{\delta^*, P^*\}$ can be preassigned where $\delta^* > 0$ and $1/k < P^* < 1$. The preference-zone $\Omega(\delta^*)$ is defined by $\Omega(\delta^*) = \{\underline{\mu} \in \Omega | \mu_{[k]} - \mu_{[k-1]} \geqq \delta^*\}$. The experimenter restricts consideration to only those procedures R which guarantee the probability requirement

$$P_{\underline{\mu}}(CS|R) \geq P^* \;\forall \underline{\mu} \in \Omega(\delta^*) \tag{2.1}$$

Bechhofer's single-stage procedure $R_0$ consists of taking $n_0$ independent observations $X_{ij} (1 \leq j \leq n_0)$ from each $\Pi_i$, computing the sample means $\bar{X}_i$ and asserting that the population associated with $\max_{1 \leq i \leq k} \bar{X}_i$ is best. To guarantee (2.1), the common sample size $n_0$ is chosen to be the smallest integer $\geqq (d_0 \sigma/\delta^*)^2$ where $d_0$ is the solution to the equation

$$\int_{-\infty}^{\infty} \Phi^{k-1}(x + d_0) d\Phi(x) = P^*, \tag{2.2}$$

and where $\Phi(\cdot)$ represents the standard normal cdf.

The basic two-stage permanent elimination type two-stage procedure $R_1$ proposed in [1] and [6] consists of the following steps:

1. In the first stage take $n_1$ independent observations $X_{ij}^{(1)}$ $(1 \le j \le n_1)$ and compute the sample means $\bar{X}_i^{(1)}$ $(1 \le i \le k)$. Choose a subset I $\subseteq \{1,2, \cdots ,k\}$ where

$$I = \{i \mid \bar{X}_i^{(1)} \geqq \max_{1 \le j \le k} \bar{X}_j^{(1)} - h\} \tag{2.3}$$

and $h > 0$.

2a. If I consists of a single population stop sampling and assert that, that population is best.

2b. If not, go onto the second stage. Take $n_2$ additional independent observations $X_{ij}^{(2)}$ $(1 \le j \le n_2)$ from each $\Pi_i$ for $i \in I$. Compute the cumulative sample means $\bar{X}_i = \left(\sum_{j=1}^{n_1} X_{ij}^{(1)} + \sum_{j=1}^{n_2} X_{ij}^{(2)}\right)/(n_1 + n_2)$ and assert that the population associated with $\max_{i \in I} \bar{X}_i$ is best. In the above $(n_1,n_2,h)$ are determined prior to the experiment so as to guarantee (2.1).

Define $T = 0$ if $|I| = 1$ and $T = |I|$ if $|I| > 1$, where $|I|$ denotes the cardinality of set I defined by (2.3). Then the total sample size N (random) associated with $R_1$ is given by

$$N = kn_1 + Tn_2.$$

For the selection of the design constants $(n_1,n_2,h)$ of $R_1$, Alam essentially proposed the following design criterion: for specified $\{\delta^*,P^*\}$ choose $(n_1,n_2,h)$ to minimize $\sup_{\underline{\mu} \in \Omega(\delta^*)} E_{\underline{\mu}}(N \mid R_1)$ subject to (2.1). We refer to this as the restricted minimax or the <u>R-minimax criterion</u>. We denote by $(\tilde{n}_1,\tilde{n}_2,\tilde{h})$ a solution to the above optimization problem and the corresponding $R_1$ by $\widetilde{R}_1$.

In [6] we pointed out certain drawbacks associated with the R-minimax criterion and proposed that $(n_1,n_2,h)$ be chosen to minimize $\sup_{\underline{\mu} \in \Omega} E_{\underline{\mu}}(N|R_1)$ instead of $\sup_{\underline{\mu} \in \Omega(\delta^*)} E_{\underline{\mu}}(N|R_1)$ subject to (2.1). We refer to this as the unrestricted minimax or simply as the <u>minimax criterion</u>. We denote by $(\hat{n}_1,\hat{n}_2,\hat{h})$ a solution to the above optimization problem and the corresponding $R_1$ by $\hat{R}_1$.

We define the RE of $R_0$ w.r.t. $R_1$ by

$$RE_{\underline{\mu}}(\delta^*,P^*,k;R_1/R_0) = E_{\underline{\mu}}(N|R_1)/kn_0, \tag{2.4}$$

where $R_0$ and $R_1$ both guarantee (2.1). We note that $R_0$ is a special case of $R_1$ for $h = 0$ and $h = \infty$. Hence $RE_{\underline{\mu}}(\delta^*,P^*,k;\widetilde{R}_1/R_0) \leqq 1 \ \forall \underline{\mu} \in \Omega(\delta^*)$ and $RE_{\underline{\mu}}(\delta^*,P^* k;\hat{R}_1/R_0) \leqq 1 \ \forall \underline{\mu} \in \Omega$.

For $\delta \geqq 0$ we denote by $\underline{\mu}(\delta)$ any $\underline{\mu}$ such that $\mu_{[1]} = \ldots = \mu_{[k-1]} = \mu_{[k]} - \delta$; $\underline{\mu}(\delta)$ is referred to as a slippage configuration. In [6] it is shown that for any $R_1, E_{\underline{\mu}}(N|R_1)$ is a non-increasing function of $\delta_{ki} (1 \le i \le k - 1)$. Therefore for any $R_1$ we have

$$E_{\underline{\mu}}(N|R_1) \leqq E_{\underline{\mu}(\delta^*)}(N|R_1) \ \forall \underline{\mu} \in \Omega(\delta^*),$$

and

$$E_{\underline{\mu}}(N|R_1) \leqq E_{\underline{\mu}(0)}(N|R_1) \quad \forall \underline{\mu} \in \Omega.$$

Thus the parameter configurations that are of interest to us are $\underline{\mu}(\delta^*)$ and $\underline{\mu}(0)$. Our objective in the present paper is to study the RE as $P^* \to 1$, $\delta^*$ and k being kept fixed, of $\hat{R}_1$ and $\widetilde{R}_1$ at $\underline{\mu}(\delta^*)$ and $\underline{\mu}(0)$. We define

$$ARE_{\underline{\mu}}(\delta^*,k;R_1/R_0) = \lim RE_{\underline{\mu}}(\delta^*,P^*,k;R_1/R_0) \tag{2.5}$$

where (and hereafter in the present paper) it is understood that the limiting operation is performed as $P^* \to 1$.

To evaluate the AREs we make the following assumptions:

A1: The conjecture (see[1]) that for $k > 2$, $\underline{\mu}(\delta^*)$ is a least favorable configuration (LFC) for $R_1$ is true, i.e., $P_{\underline{\mu}(\delta^*)}(CS|R_1) = \inf_{\underline{\mu} \in \Omega(\delta^*)} P_{\underline{\mu}}(CS|R_1)$. It is known that this conjecture is true for $k = 2$.

A2: All the limits encountered in the present paper exist but may take values $\pm \infty$.

It would appear that both the assumptions would be true in practice although we could not prove them rigorously.

For the sake of simplicity we shall henceforth denote $\underline{\mu}(\delta^*)$ by LFC and $\underline{\mu}(0)$ by EMC (the equal means configuration).

## 3. PRELIMINARY RESULTS

For further analysis, we reparametrize the design constants associated with $R_1$ as follows:

$$c = \frac{h\sqrt{n_1}}{\sigma}\ ,\ d_1 = \frac{\delta^*\sqrt{n_1}}{\sigma}\ ,\ d_2 = \frac{\delta^*\sqrt{n_2}}{\sigma}\ . \tag{3.1}$$

We shall regard c, $d_1$, $d_2$ and also $n_0$ as nonnegative continuous variables. It is known (see [6]) that the value of $P_{\underline{\mu}}(CS|R_1)$ at the conjectured LFC depends on $\delta^*$, $\sigma$, $n_1$, $n_2$ and h only through c, $d_1$ and $d_2$. We denote the corresponding value of $P_{LFC}(CS|R_1)$ by $\Psi(c,d_1,d_2;k)$; an expression for $\Psi$ involving multivariate normal integrals is given in [6]. The objective functions for the optimization problems associated with finding the design constants for $\hat{R}_1$ and $\tilde{R}_1$ can also be written in terms of c, $d_1$ and $d_2$ as indicated below in (3.5) and (3.6). Thus a major advantage of the reparametrization and regarding the variables defined in (3.1) as continuous is that, the solutions in $(c,d_1,d_2)$ of the optimization problems associated with $\hat{R}_1$ and $\tilde{R}_1$ do not depend on $\delta^*$ and $\sigma$ which facilitates the tabling of these constants for the purpose of implementation of the procedures.

In [6] it is shown that an expression for the RE of $R_0$ w.r.t.$R_1$ can be written as follows:

$$RE_{\underline{\mu}}(\delta^*,P^*,k;R_1/R_0) = \frac{1}{kd_0^2}\Big[kd_1^2 + d_2^2\sum_{i=1}^{k}\int_{-\infty}^{\infty}\Big\{\prod_{\substack{j=1\\j\neq i}}^{k}\Phi\Big[x+\frac{\delta_{ij}d_1}{\delta^*}+c\Big]$$

$$-\prod_{\substack{j=1\\j\neq i}}^{k}\Phi\Big[x+\frac{\delta_{ij}d_1}{\delta^*}-c\Big]\Big\}\,d\Phi(x)\Big], \tag{3.2}$$

where $\delta_{ij} = \mu_{[i]} - \mu_{[j]}$ $(1 \le i \ne j \le k)$, $d_0$ is given by (2.2), and $c$, $d_1$, and $d_2$ satisfy $\Psi(c,d_1,d_2;k) = P^*$.

Remark: From (3.2) it may be noted that for any $R_1$, the RE depends on $\delta^*$ only through the ratios $\delta_{ij}/\delta^*$ $(1 \le i \ne j \le k)$. For the EMC all the ratios equal 0 and for the LFC all the ratios equal 0 or 1. Thus the $RE_{EMC}$ and $RE_{LFC}$ are independent of $\delta^*$. Further since $R_0$ is always used as a basis for comparison, we shall omit the dependence of the corresponding REs on $R_0$ from the notation. Thus we shall simply write $RE_{EMC}(P^*,k;R_1)$ and $RE_{LFC}(P^*,k;R_1)$. We write below the expressions for these two quantities for later reference:

$$RE_{EMC}(P^*,k;R_1) = \frac{d_1^2 + d_2^2 \int_{-\infty}^{\infty} \left\{ \Phi^{k-1}(x+c) - \Phi^{k-1}(x-c) \right\} d\Phi(x)}{d_0^2}, \tag{3.3}$$

and

$$RE_{LFC}(P^*,k;R_1) = \frac{1}{kd_0^2} \Big\{ kd_1^2 + d_2^2 \Big[ \int_{-\infty}^{\infty} \left\{ \Phi^{k-1}(x+d_1+c) - \Phi^{k-1}(x+d_1-c) \right\} d\Phi(x)$$

$$+ (k-1) \int_{-\infty}^{\infty} \left\{ \Phi^{k-2}(x+c)\, \Phi(x-d_1+c) - \Phi^{k-2}(x-c)\, \Phi(x-d_1-c) \right\} d\Phi(x) \Big] \Big\}$$

$$= \frac{d_1^2 + d_2^2 G(c,d_1,d_2;k)}{d_0^2}. \tag{3.4}$$

To obtain $RE_{\underline{\mu}}(\delta^*,P^*,k;\hat{R}_1)$ we substitute in (3.2) the values of $(\hat{c},\hat{d}_1,\hat{d}_2)$, the design constants associated with $\hat{R}_1$. The values of $(\hat{c},\hat{d}_1,\hat{d}_2)$ are obtained by solving the following optimization problem:

$$\text{Minimize} \qquad \left(\frac{\delta^*}{\sigma}\right)^2 \sup_{\underline{\mu} \in \Omega} E_{\underline{\mu}}(N|R_1)$$

$$= \left(\frac{\delta^*}{\sigma}\right)^2 E_{EMC}(N|R_1)$$

$$= kd_1^2 + kd_2^2 \int_{-\infty}^{\infty} \left\{\Phi^{k-1}(x + c) - \Phi^{k-1}(x - c)\right\} d\Phi(x),$$

$$\text{Subject to} \qquad \inf_{\underline{\mu} \in \Omega(\delta^*)} P_{\underline{\mu}}(CS|R_1)$$

$$= P_{LFC}(CS|R_1)$$

$$= \Psi(c, d_1, d_2; k) \geq P^*. \tag{3.5}$$

To obtain $RE_{\underline{\mu}}(\delta^*, P^*, k; \widetilde{R}_1)$ we substitute in (3.2) the values of $(\widetilde{c}, \widetilde{d}_1, \widetilde{d}_2)$, the design constants associated with $\widetilde{R}_1$. The values of $(\widetilde{c}, \widetilde{d}_1, \widetilde{d}_2)$ are obtained by solving the following optimization problem:

$$\text{Minimize} \qquad \left(\frac{\delta^*}{\sigma}\right)^2 \sup_{\underline{\mu} \in \Omega(\delta^*)} E_{\underline{\mu}}(N|R_1)$$

$$= \left(\frac{\delta^*}{\sigma}\right)^2 E_{LFC}(N|R_1)$$

$$= kd_1^2 + d_2^2 \Big[ \int_{-\infty}^{\infty} \left\{\Phi^{k-1}(x + d_1 + c) - \Phi^{k-1}(x + d_1 - c)\right\} d\Phi(x)$$

$$+ (k - 1) \int_{-\infty}^{\infty} \left\{\Phi^{k-2}(x + c)\Phi(x - d_1 + c) - \Phi^{k-2}(x - c) \times \Phi(x - d_1 - c)\right\} d\Phi(x) \Big],$$

$$\text{Subject to} \qquad \Psi(c, d_1, d_2; k) \geq P^*. \tag{3.6}$$

It may be noted that the optimization problems (3.5) and (3.6) remain unchanged if the corresponding objective functions are replaced by (3.3) and (3.4), respectively.

The task of computing the AREs can be seen to be extremely formidable in view of the complicated nature of the nonlinear programming problems (3.5) and (3.6) associated with $\hat{R}_1$ and $\tilde{R}_1$, respectively. We have been able to simplify this task somewhat by employing the following approach: consider, say, $\tilde{R}_1$ and all the possible limiting values that $\tilde{c}$, $\tilde{d}_1$ and $\tilde{d}_2$ can take as $P^* \to 1$. It turns out that for the purpose of evaluating the $ARE_{LFC}$ we need to consider only few possibilities for the limiting values of $(\tilde{c},\tilde{d}_1,\tilde{d}_2)$. For each possible case the value of the $ARE_{LFC}$ is evaluated. Since $\tilde{R}_1$ is designed to minimize $RE_{LFC}(P^*,k;R_1)$ among all $R_1$ guaranteeing (2.1) for each $P^*$, it is clear that the actual value of the $ARE_{LFC}$ would be the minimum of those obtained for all the possible limiting values of $(\tilde{c},\tilde{d}_1,\tilde{d}_2)$. A similar consideration holds true for computing the $ARE_{EMC}$ of $R_0$ w.r.t. $\hat{R}_1$. A drawback of this approach is that it does not facilitate the computation of the other AREs of interest, namely the $ARE_{LFC}$ of $R_0$ w.r.t. $\hat{R}_1$ and $ARE_{EMC}$ of $R_0$ w.r.t. $\tilde{R}_1$. We have not been able to find any other method which would perform this task.

Now we state few preliminary lemmas which would be repeatedly used in proving the main results of the next section.

Lemma 1: For any $\underline{\mu} \in \Omega$ we have

$$\int_{-\infty}^{\infty} \left\{ \prod_{i=1}^{k-1} \Phi\left[x + \frac{(\delta_{ki} + h)\sqrt{n}_1}{\sigma}\right] + \prod_{i=1}^{k-1} \Phi\left[x + \frac{\delta_{ki}\sqrt{n}}{\sigma}\right] \right\} d\Phi(x) - 1$$

$$\leq P_{\underline{\mu}}(CS|R_1) \leq \int_{-\infty}^{\infty} \prod_{i=1}^{k-1} \Phi\left[x + \frac{(\delta_{ki} + h)\sqrt{n}_1}{\sigma}\right] d\Phi(x), \qquad (3.7)$$

where $n = n_1 + n_2$.

Proof: For a proof see [6].

Corollary: For any $R_1$ we have

$$1 - P_{LFC}(CS|R_1) = \{1 - \int_{-\infty}^{\infty} \Phi^{k-1}(x + d_1 + c)d\Phi(x)\}$$
$$+ \gamma\{1 - \int_{-\infty}^{\infty} \Phi^{k-1}(x + \sqrt{d_1^2 + d_2^2})d\Phi(x)\}, \quad (3.8)$$

where $0 \leqq \gamma \leqq 1$.

Lemma 2: For any $R_1$, as $P^* \to 1$ we have $d_1 + c \to \infty$ and $d_1^2 + d_2^2 \to \infty$.

Proof: The proof is straightforward using (3.8).

Lemma 3 (Bechhofer, Kiefer and Sobel [4]): Let $H(u) = 1 - \int_{-\infty}^{\infty} \Phi^{k-1}(x + u)d\Phi(x)$. Then as $u \to \infty$ we have $H(u) \sim \frac{(k-1)e^{-u^2/4}}{u\sqrt{\pi}}$ where $a \sim b$ means $a/b \to 1$ in the limit.

Corollary: As $P^* \to 1$ the solution in $d_0$ to equation (2.2) is given by

$$d_0^2 = 4\ln\left(\frac{k-1}{1-P^*}\right) - 2\ln\ln\left(\frac{k-1}{1-P^*}\right) - 2\ln 4\pi + o(1)$$

or

$$d_0^2 \sim -4\ln(1 - P^*)$$

Lemma 4: For any $R_1$, as $P^* \to 1$ we have

$$1 - P^* \sim \frac{(k-1)}{\sqrt{\pi}}\left\{\frac{e^{-(d_1+c)^2/4}}{(d_1 + c)} + \gamma\,\frac{e^{-(d_1^2+d_2^2)/4}}{(d_1^2 + d_2^2)^{1/2}}\right\} \quad (3.9)$$

with $0 \leqq \gamma \leqq 1$ and $\gamma = \gamma(P^*,k;R_1)$.

Proof: Use (3.8) and Lemmas 2 and 3.

We remark here that a key step in the proofs of the main theorems in the next section invloves determination of the dominating term in (3.9) as $P^* \to 1$ for the specified limiting values of $(c,d_1,d_2)$.

Lemma 5: For any $R_0$ and $R_1$ guaranteeing (2.1) we have

$$d_1^2 + d_2^2 \geqq d_0^2.$$

Proof: The proof is straightforward and is omitted.

## 4. MAIN RESULTS

We first prove a result concerning the $ARE_{EMC}$ of $R_0$ w.r.t. $\hat{R}_1$ in the following theorem.

Theorem 1: Under assumptions A1 and A2, we have,

$$ARE_{EMC}(k;\hat{R}_1) = 1 \quad \text{for all } k \geqq 2.$$

Proof: By Lemma 4, as $P^* \to 1$ we have

$$1 - P^* \sim \frac{(k-1)}{\sqrt{\pi}} \left\{ \frac{e^{-(\hat{d}_1+\hat{c})^2/4}}{(\hat{d}_1 + \hat{c})} + \gamma \frac{e^{-(\hat{d}_1^{\,2}+\hat{d}_2^{\,2})/4}}{(\hat{d}_1^{\,2} + \hat{d}_2^{\,2})^{1/2}} \right\},$$

where $0 \leqq \gamma \leqq 1$ and $\gamma = \gamma(k,P^*;\hat{R}_1)$. Now we consider the following two possibilities for the limiting values of $\hat{c}$.

Case (i): $\lim \hat{c} = \infty$ : In this case we have

$$ARE_{EMC}(k;\hat{R}_1) = \lim \left[ \frac{\hat{d}_1^{\,2} + \hat{d}_2^{\,2} \int_{-\infty}^{\infty} \{\Phi^{k-1}(x+\hat{c}) - \Phi^{k-1}(x-\hat{c})\}\, d\Phi(x)}{d_0^{\,2}} \right]$$

$$= \lim \left( \frac{\hat{d}_1^{\,2} + \hat{d}_2^{\,2}}{d_0^{\,2}} \right)$$

$$\geqq 1.$$

The last step is obtained by using Lemma 5.

Case (ii): $\lim \hat{c} < \infty$: since $\hat{d}_1 + \hat{c} \to \infty$ from Lemma 2, we have $\hat{d}_1 \to \infty$,

$(\hat{c}/\hat{d}_1) \to 0$ and

$$1 - P^* \sim \frac{(k-1)}{\sqrt{\pi}} \left\{ \frac{e^{-\hat{d}_1^2/4}}{\hat{d}_1} + \gamma \frac{e^{-(\hat{d}_1^2+\hat{d}_2^2)/4}}{(\hat{d}_1^2 + \hat{d}_2^2)^{1/2}} \right\}$$

$$\sim \frac{(k-1)A}{\sqrt{\pi}} \cdot \frac{e^{-\hat{d}_1^2/4}}{\hat{d}_1} ,$$

where $1 \leqq A \leqq 2$. Hence $\hat{d}_1^2 \sim -4\ln(1 - P^*) \sim d_0^2$ using Lemma 3 and therefore $ARE_{EMC}(k;\hat{R}_1) \geqq 1$. But $RE_{EMC}(k,P^*;\hat{R}_1) \leqq 1$ for all $P^*$ (see remark following (2.4)) implies that $ARE_{EMC}(k;\hat{R}_1) \leqq 1$. Therefore $ARE_{EMC}(k;\hat{R}_1) = 1$ and the proof is complete. $\blacksquare$

This theorem tells us that the $ARE_{EMC}$ of $R_0$ w.r.t any $R_1$ must be at least 1 since $\hat{R}_1$ minimizes $RE_{EMC}$ among all $R_1$ guaranteeing (2.1) for every $P^*$. In particular we have,

$$ARE_{EMC}(k;\tilde{R}_1) \geqq 1. \tag{4.1}$$

In the next theorem we study $ARE_{LFC}(k;\tilde{R}_1)$.

<u>Theorem 2</u>: Under assumptions A1 and A2 we have,

$$ARE_{LFC}(k;\tilde{R}_1) = \frac{1}{4} \quad \text{for all } k \geqq 2.$$

Also as $P^* \to 1$, $\tilde{d}_1 - \tilde{c} = o(\tilde{d}_1) \to \infty$ and $\lim (\tilde{d}_2/\tilde{d}_1)^2 \geqq 3$.

<u>Proof</u>: By Lemma 4, as $P^* \to 1$ we have

$$1 - P^* \sim \frac{(k-1)}{\sqrt{\pi}} \left\{ \frac{e^{-(\tilde{d}_1+\tilde{c})^2/4}}{(\tilde{d}_1 + \tilde{c})} + \gamma \frac{e^{-(\tilde{d}_1^2+\tilde{d}_2^2)/4}}{(\tilde{d}_1^2 + \tilde{d}_2^2)^{1/2}} \right\} , \tag{4.2}$$

where $0 \leqq \gamma \leqq 1$ and $\gamma = \gamma(P^*,k;\tilde{R}_1)$. The following proof proceeds by considering all the limiting values that $(\tilde{c},\tilde{d}_1,\tilde{d}_2)$ can take and evaluating $ARE_{LFC}(k;\tilde{R}_1)$ in each case. The actual limiting value of $(\tilde{c},\tilde{d}_1,\tilde{d}_2)$ would be the one for which $ARE_{LFC}(k;\tilde{R}_1)$ is minimum.

<u>Case (i)</u>: $\lim |\tilde{d}_1 - \tilde{c}| < \infty$: In this case as $P^* \to 1$ from (3.9) we have

$$1 - P^* \sim \frac{(k-1)}{\sqrt{\pi}} \left\{\frac{e^{-\tilde{d}_1^2}}{2\tilde{d}_1} + \gamma \frac{e^{-(\tilde{d}_1^2 + \tilde{d}_2^2)/4}}{(\tilde{d}_1^2 + \tilde{d}_2^2)^{1/2}}\right\}. \tag{4.3}$$

<u>Subcase (ia)</u>: $\lim (\tilde{d}_2/\tilde{d}_1)^2 < 3$: Then we have $\tilde{d}_1^2 > (\tilde{d}_1^2 + \tilde{d}_2^2)/4$ for $P^*$ arbitrarily close to 1. If $\gamma \not\to 0$ or if $\gamma \to 0$ but not at a rapid enough rate then we can write

$$1 - P^* \sim \frac{(k-1)A}{\sqrt{\pi}} \cdot \frac{e^{-(\tilde{d}_1^2+\tilde{d}_2^2)/4}}{(\tilde{d}_1^2 + \tilde{d}_2^2)^{1/2}}, \tag{4.4}$$

where $0 < A < 2$. However, if $\gamma \to 0$ at a rapid enough rate then

$$1 - P^* \sim \frac{(k-1)B}{\sqrt{\pi}} \frac{e^{-\tilde{d}_1^2}}{2\tilde{d}_1}, \tag{4.5}$$

where $1 \leqq B \leqq 2$. Suppose (4.5) holds then using Lemma 3 we have $\tilde{d}_1^2 \sim -\ln(1 - P^*) \sim d_0^2/4$ and $\lim (\tilde{d}_2/d_0)^2 < 3/4$. Therefore using (3.3) and noting that in the present case $\tilde{c} \to \infty$ we obtain,

$$ARE_{EMC}(k;\tilde{R}_1) < \frac{1}{4} + \frac{3}{4} = 1.$$

This contradicts (4.1). Hence (4.5) does not hold and therefore from (4.4) we have $(\tilde{d}_1^2 + \tilde{d}_2^2) \sim -4\ln(1 - P^*) \sim d_0^2$ using Lemma 3. Now since

$\lim\,(\tilde{d}_2/\tilde{d}_1)^2 < 3$, it follows that $\lim\,(\tilde{d}_1/d_0)^2 > 1/4$ and consequently $ARE_{LFC}(k;\tilde{R}_1) > 1/4$.

<u>Subcase (ib)</u>: $\lim\,(\tilde{d}_2/\tilde{d}_1)^2 \geq 3$: we have $\tilde{d}_1^{\,2} \leq (\tilde{d}_1^{\,2} + \tilde{d}_2^{\,2})/4$ for $P^*$ arbitrarily close to 1. Therefore from (4.3) we can write

$$1 - P^* \sim \frac{(k-1)A}{\sqrt{n}} \; \frac{e^{-\tilde{d}_1^{\,2}}}{2\tilde{d}_1}\,,$$

where $1 \leq A \leq 2$. Therefore $\tilde{d}_1^{\,2} \sim -\ln(1-P^*) \sim d_0^{\,2}/4$ using Lemma 3 and $\lim\,(\tilde{d}_2/d_0)^2 \geq 3/4$. Now $\lim\,|\tilde{d}_1 - \tilde{c}| < \infty$ implies that $\tilde{d}_1 \to \infty$, $\tilde{c} \to \infty$ and

$$0 < \lim\, G(\tilde{c},\tilde{d}_1,\tilde{d}_2;k) < 1, \tag{4.6}$$

where G is defined by (3.4). Therefore we have $ARE_{LFC}(k;\tilde{R}_1) > 1/4$.

<u>Case (ii)</u>: $\lim\,(\tilde{d}_1 - \tilde{c}) = \infty$, $\tilde{d}_1 \to \infty$, $\tilde{c} \to \infty$, $(\tilde{d}_1 - \tilde{c})/\tilde{d}_1 \to 0$:

In this case also (4.3) holds and we have subcases (a) and (b) just as in Case (i). The analysis of Subcase (iia) is the same as that of (ia) and we obtain $ARE_{LFC}(k;\tilde{R}_1) > 1/4$. The analysis of Subcase (iib) is similar to that of (ib) except now instead of (4.6) we have

$$\lim\, G(\tilde{c},\tilde{d}_1,\tilde{d}_2;k) = 0.$$

Further $\lim\,(\tilde{d}_1/d_0)^2 = 1/4$, $\lim\,(\tilde{d}_2/d_0)^2 \geq 3/4$ and $\lim\,\tilde{d}_2^{\,2}\, G(\tilde{c},\tilde{d}_1,\tilde{d}_2;k)/d_0^{\,2} = 0$. Therefore $ARE_{LFC}(k;\tilde{R}_1) = 1/4$.

<u>Case (iii)</u>: $\lim\,(\tilde{d}_1 - \tilde{c}) = \infty$ and $(\tilde{d}_1 - \tilde{c})/\tilde{d}_1 \not\to 0$: Therefore $\tilde{c}/\tilde{d}_1 \to B$ where $0 \leq B < 1$. Also denote the limiting value of $(\tilde{d}_2/\tilde{d}_1)^2$ by D. Then as $P^* \to 1$ we have,

$$1 - P^* \sim \frac{(k-1)}{\sqrt{\pi}} \left\{ \frac{e^{-\tilde{d}_1^{\,2}(1+B)^2/4}}{\tilde{d}_1(1+B)} + \gamma\, \frac{e^{-\tilde{d}_1^{\,2}(1+D)/4}}{\tilde{d}_1(1+D)^{1/2}} \right\}. \tag{4.7}$$

Subcase (iiia): $(1 + B)^2 < (1 + D)$: From (4.7) we have

$$1 - P^* \sim \frac{(k - 1)A}{\sqrt{\pi}} \cdot \frac{e^{-\tilde{d}_1^2(1+B)^2/4}}{\tilde{d}_1(1 + B)},$$

where $1 \leqq A \leqq 2$. Using Lemma 3 we have $\tilde{d}_1^2 \sim -4\ln(1 - P^*)/(1 + B)^2 \sim d_0^2/(1 + B)^2$. Therefore $ARE_{LFC}(k;\tilde{R}_1) \geqq (\tilde{d}_1/d_0)^2 = 1/(1 + B)^2 > 1/4$.

Subcase (iiib): $(1 + B)^2 \geqq (1 + D)$: In this case if $\gamma \to 0$ at a rapid enough rate then the first term in (4.7) dominates and we are back to Subcase (iiia). If $\gamma \not\to 0$ or if $\gamma \to 0$ but not at a rapid enough rate then from (4.7) we have

$$1 - P^* \sim \frac{(k - 1)A}{\sqrt{\pi}} \cdot \frac{e^{-\tilde{d}_1^2(1+D)/4}}{\tilde{d}_1(1 + D)},$$

where $0 < A \leqq 2$. Using Lemma 3 we have $\tilde{d}_1^2 \sim -4\ln(1 - P^*)/(1 + D) \sim d_0^2/(1 + D)$. But $D \leqq (1 + B)^2 - 1 < 3$. Hence $ARE_{LFC}(k;\tilde{R}_1) \geqq 1/(1 + D) > 1/4$.

Case (iv): $\lim (\tilde{d}_1 - \tilde{c}) = -\infty$: In this case $\lim G(\tilde{c},\tilde{d}_1,\tilde{d}_2;k) = 1$. Therefore using (3.4) we have $ARE_{LFC}(k;\tilde{R}_1) = \lim (\tilde{d}_1^2 + \tilde{d}_2^2)/d_0^2 \geqq 1$.

From cases (i) - (iv) we note that Subcase (iib) yields the minimum value = 1/4 for $ARE_{LFC}(k;\tilde{R}_1)$. Hence the theorem is proved. $\square$

## 5. DISCUSSION

From [4] we know that for the fully sequential procedure BKS, $ARE_{LFC}(k;BKS/R_0) = 1/4$. Thus the ratio of the expected total sample sizes in the LFC required by BKS and $\tilde{R}_1$ to guarantee (2.1) tends to 1 as $P^* \to 1$. In particular, the WSPRT to test $H_0$: $\mu_1 - \mu_2 \geqq \delta^*$ against $H_1$: $\mu_1 - \mu_2 \leqq -\delta^*$ is a special case of BKS for $k = 2$. The WSPRT is known to possess the optimum property that it simultaneously minimizes the expected total sample size at all the parameter configurations $\underline{\mu}$ for which $\mu_{[2]} - \mu_{[1]} = \delta^*$ among all tests with specified probabilities of Type I and Type II errors. If both the error probabilities are set equal to $1 - P^*$ and if $P^* \to 1$ then the ARE in the LFC of the "most economical" single-stage procedure which is known to be $R_0$ (see [5]) w.r.t. the WSPRT is 1/4. Thus for $k = 2$, we find that as $P^* \to 1$, the two-stage procedure $\tilde{R}_1$ performs as well as the WSPRT which is the optimum procedure for the given testing problem. This is a somewhat surprising, but a very important result.

The result of Theorem 1 indicates that as $P^* \to 1$, no two-stage procedure of the form of $R_1$ can have the expected total sample size in the EMC less than the total sample size needed by $R_0$. In particular, (4.1) holds. Our efforts to arrive at the exact value of $ARE_{EMC}(k;\tilde{R}_1)$ were fruitless. However the following table of values of the REs at the LFC and the EMC for $\hat{R}_1$ and $\tilde{R}_1$ throws some light on the behavior of these quantities.

Table I

Relative Efficiencies of $\hat{R}_1$ and $\tilde{R}_1$ (k = 2)

| $P^*$ | $RE_{LFC}(P^*,k;\hat{R}_1)$ | $RE_{EMC}(P^*,k;\hat{R}_1)$ | $RE_{LFC}(P^*,k;\tilde{R}_1)$ | $RE_{EMC}(P^*,k;\tilde{R}_1)$ |
|---|---|---|---|---|
| .9999 | .74661 | .90047 | .53237 | 1.01811 |
| .9995 | .73470 | .89442 | .56079 | .99273 |
| .999 | .72786 | .89137 | .57593 | .97828 |
| .99 | .72195 | .87850 | .64581 | .92509 |
| .95 | .74574 | .86540 | .72127 | .88560 |
| .90 | .77158 | .85799 | .76175 | .86724 |
| .85 | .79026 | .85319 | .78641 | .85695 |
| .80 | .80587 | .84982 | .80408 | .85162 |
| .75 | .81764 | .84719 | .81689 | .84795 |

From this table it appears that the rates of approach of $RE_{EMC}(P^*,;\hat{R}_1)$ to 1 and that of $RE_{LFC}(P^*,k;\tilde{R}_1)$ to 1/4 as $P^* \to 1$ are monotonic but are fairly slow. As $P^*$ increases $RE_{LFC}(P^*,k;\hat{R}_1)$ first decreases and then starts increasing. We also notice that $RE_{EMC}(P^*,k;\tilde{R}_1)$ increases with $P^*$ and is greater than 1 for $P^* = .9999$. At present we do not know the exact value of $ARE_{EMC}(k;\tilde{R}_1)$. We only know that $ARE_{EMC}(k;\tilde{R}_1) = \lim (\tilde{d}_1^{\,2} + \tilde{d}_2^{\,2})/d_0^{\,2} = 1/4 + \lim (\tilde{d}_2/d_0)^2 \geq 1$. But we conjecture that $1 < ARE_{EMC}(k;\tilde{R}_1) < \infty$.

In [3] Bechhofer showed that for the testing problem described above the ratio of the expected total sample size required by the WSPRT in the EMC $(\mu_1 = \mu_2)$ to the total sample size required by $R_0$ tends to $\infty$ as $P^* \to 1$. Thus our conjecture is that $\tilde{R}_1$ does not possess this extremely undesirable property possessed by the WSPRT.

## 6. SUGGESTIONS FOR FUTURE RESEARCH

Clearly it would be very useful to develop a general method to evaluate the AREs for $\hat{R}_1$ and $\tilde{R}_1$ at any $\underline{\mu}$, and in particular at the LFC and the EMC. Also because of the screening aspect of the two-stage procedure $R_1$, it is anticipated that the gains achieved by $\hat{R}_1$ and $\tilde{R}_1$ over that of $R_0$ in terms of the expected total sample sizes would increase substantially as k increases. Therefore it would be of some interest to study the limiting behavior of the RE for $\hat{R}_1$ and $\tilde{R}_1$ as $k \to \infty$, $\delta^*$ and $P^*$ being kept fixed.

## ACKNOWLEDGEMENTS

The writer wishes to thank Professors Robert Bechhofer and Thomas Santner for their suggestions and comments.

# BIBLIOGRAPHY


1. Alam, K. [1970]: "A two-sample decision procedure for selecting the population with the largest mean from k normal populations," Ann. Inst. Statist. Math (Tokyo), 22, 127-36.

2. Bechhofer, R. E. [1954]: "A single-sample multiple decision procedure for ranking means of normal populations with known variances," Ann. Math. Statist., 25, 16-39.

3. Bechhofer, R. E. [1960]: "A note on the limiting relative efficiency of the Wald sequential probability ratio test," J. Amer. Statist. Assoc., 55, 660-3.

4. Bechhofer, R. E., Kiefer, J., and Sobel, M. [1968]: Sequential Identification and Ranking Procedures, Statistical Research Monographs, vol. III, The University of Chicago Press, Chicago, Illinois.

5. Hall, W. J. [1959]: "The most economical character of some Bechhofer and Sobel decision rules," Ann. Math. Statist., 31, 296-304.

6. Tamhane, A. C., and Bechhofer, R. E. [1975]: "A minimax two-stage permanent elimination type procedure for selecting the largest normal mean (common known variance)," A paper under preparation.

7. Wald, A., and Wolfowitz, J. [1948]: "Optimum character of the sequential probability ratio test," Ann. Math. Statist., 19, 326-39.