ON BAYES AND EMPIRICAL BAYES TWO-STAGE ALLOCATION PROCEDURES FOR SELECTION PROBLEMS*

by

Shanti S. Gupta
Purdue University

TaChen Liang
Wayne State University

Technical Report#87-53

Department of Statistics
Purdue University

December 1987
Revised March 1989

* This research was supported in part by the Office of Naval Research Contract N00014-84-C-0167 at Purdue University and NSF Grant DMS-8606964. Reproduction in whole or in part is permitted for any purpose of the United States Government.

o

ON BAYES AND EMPIRICAL BAYES TWO-STAGE ALLOCATION PROCEDURES FOR SELECTION PROBLEMS*

Shanti S. GUPTA and TaChen LIANG

Department of Statistics, Purdue University, West Lafayette, IN 47907, USA
Department of Mathematics, Wayne State University, Detroit, MI 48202, USA

ABSTRACT: We investigate the problem of deriving two-stage allocation procedures for selecting the best normal population. If the prior distribution is assumed to be known, an exact Bayes two-stage allocation procedure is obtained. If the prior distribution depends on some unknown parameter, an adaptive two-stage allocation procedure is proposed. Using the empirical Bayes formulation, we prove that the proposed adaptive two-stage allocation procedure has some asymptotic optimality property.

1. INTRODUCTION

Suppose that an experimenter (a customer) wishes to purchase M items of some product. We assume that these items are supplied by k different manufacturers (suppliers), say, $\pi_1, \dots, \pi_k$. At first, the experimenter carries out an inspection on m items of the product from each of the k suppliers in order to obtain data for determining their quality. Then, based on the resulting information, he allocates the remaining $M - km$ items to the k suppliers, say, $N_1, \dots, N_k$, respectively, where N_i, $i = 1, \dots, k$, are nonnegative integers such that $\sum_{i=1}^{k} N_i = M - km$. Let θ_i denote a measure of the quality of the product from the ith manufacturer π_i. Let $\theta_{[1]} \leq \dots \leq \theta_{[k]}$ be the ordered values of the parameters $\theta_1, \dots, \theta_k$. It is assumed that the exact pairing between the ordered and the unordered parameters is unknown. The supplier π_i with $\theta_i = \theta_{[k]}$ is called the *best*. Of course, the experimenter would ideally like to allocate (purchase) the remaining $M - km$ items from the best supplier. Thus, the experimenter is faced with the so-called two-stage allocation and selection problem.

For the two-stage allocation problem described above, we define the loss function:

$$L(\underset{\sim}{\theta}; m, N_1, \dots N_k) = m \sum_{i=1}^{k} (\theta_{[k]} - \theta_i) + \sum_{i=1}^{k} N_i (\theta_{[k]} - \theta_i), \tag{1.1}$$

where $\underset{\sim}{\theta} = (\theta_1, \dots, \theta_k)$, $0 \leq m \leq [\frac{M}{k}]$, $0 \leq N_i \leq M - km$, $i = 1, \dots, k$, $\sum_{i=1}^{k} N_i = M - km$,

* This research was supported in part by the Office of Naval Research Contract N00014-88-K-0170 and NSF Grants DMS–8606964, DMS-8702620 at Purdue University.

and $[y]$ denotes the largest integer not greater than y. Note that the first summation in (1.1) is the loss due to the choice of the common initial number of items to be supplied by each of the k manufacturers, and the second summation in (1.1) is the loss due to the allocation made at the second stage. Our goal here is to derive optimal two-stage allocation procedures with respect to the loss function (1.1). We study the problem for normal populations, say $\pi_1, \ldots, \pi_k$, with unknown means $\theta_1, \ldots, \theta_k$, and a common known variance σ^2. The unknown means $\theta_1, \ldots, \theta_k$ are assumed to be independent and identically distributed (iid) with a normal prior distribution $N(\theta_0, \tau^2)$, where the parameter τ^2 may be either known or unknown.

We note that Somerville [6, 7] studied a two-stage minimax allocation procedure for the normal distribution model with a different loss function. However, since the loss function considered by Somerville [6] is not bounded, the minimax solution does not exist (see Ofosu [3] for a comment). Ofosu [4] also studied a two-stage allocation procedure via a Bayesian approach (see Gupta and Panchapakesan [2]).

2. NORMAL MODEL

Let $\pi_1, \ldots, \pi_k$ be normal populations with unknown means $\theta_1, \ldots, \theta_k$, and a common known variance σ^2. The unknown means $\theta_1, \ldots, \theta_k$ are assumed to be iid with a normal prior distribution $N(\theta_0,\ \tau^2)$. In this section, we assume that τ^2 is known. Also, for simplicity, we assume that $M = kN$ for some positive integer N.

2.1. Bayes Allocation Procedure for a Fixed m

First, we take m, $0 \leq m \leq N$, random observations from each of the k populations. Let $\overline{X}_i$ denote the sample mean of the m random observations taken from population π_i and let $\overline{x}_i$ denote the associated observed value, $i = 1, \ldots, k$. At the second stage, based on the observed values $\underset{\sim}{\overline{x}} = (\overline{x}_1, \ldots, \overline{x}_k)$, allocate $N_i(\underset{\sim}{\overline{x}})$ random observations to population π_i, $i = 1, \ldots, k$, where $N_1(\underset{\sim}{\overline{x}}), \ldots, N_k(\underset{\sim}{\overline{x}})$ are nonnegative integers such that $\sum_{i=1}^{k} N_i(\underset{\sim}{\overline{x}}) = k(N-m)$. Let $\overline{Y}_i$ denote the sample mean of the $N_i(\underset{\sim}{\overline{x}})$ random observations taken from the population π_i at the second stage, and let $\overline{y}_i$ be the associated observed value, $i = 1, \ldots, k$. Also, let $\underset{\sim}{\overline{y}} = (\overline{y}_1, \ldots, \overline{y}_k)$. Note that when either $m = 0$ or $m = N$, the above allocation procedure is reduced to a one-stage allocation procedure.

At stage two, given $\underset{\sim}{\overline{X}} = \underset{\sim}{\overline{x}}$ and $\underset{\sim}{\overline{Y}} = \underset{\sim}{\overline{y}}$, respectively, the posterior expected loss is:

$$\begin{aligned} r_m(\underset{\sim}{\overline{x}}, \underset{\sim}{\overline{y}}) &= E[L(\underset{\sim}{\theta}; m,\ N_1(\underset{\sim}{\overline{x}}), \ldots, N_k(\underset{\sim}{\overline{x}}))|\underset{\sim}{\overline{X}} = \underset{\sim}{\overline{x}},\ \underset{\sim}{\overline{Y}} = \underset{\sim}{\overline{y}}] \\ &= kNE[\theta_{[k]}|\underset{\sim}{\overline{X}} = \underset{\sim}{\overline{x}},\ \underset{\sim}{\overline{Y}} = \underset{\sim}{\overline{y}}] - \sum_{j=1}^{k}(m + N_j(\underset{\sim}{\overline{x}}))E[\theta_j|\underset{\sim}{\overline{X}} = \underset{\sim}{\overline{x}},\ \underset{\sim}{\overline{Y}} = \underset{\sim}{\overline{y}}]. \end{aligned} \tag{2.1}$$

Therefore, at stage one, given $\underset{\sim}{\overline{X}} = \underset{\sim}{\overline{x}}$, the posterior expected loss is given by

$$
\begin{aligned}
r_m(\underset{\sim}{\overline{x}}) &= E[r_m(\underset{\sim}{\overline{X}},\underset{\sim}{\overline{Y}})|\underset{\sim}{\overline{X}} = \underset{\sim}{\overline{x}}] \\
&= kNE[\theta_{[k]}|\underset{\sim}{\overline{X}} = \underset{\sim}{\overline{x}}] - \sum_{j=1}^{k}(m + N_j(\underset{\sim}{\overline{x}}))E[\theta_j|\underset{\sim}{\overline{X}} = \underset{\sim}{\overline{x}}] \\
&= kNE[\theta_{[k]}|\underset{\sim}{\overline{X}} = \underset{\sim}{\overline{x}}] - \sum_{j=1}^{k}(m + N_j(\underset{\sim}{\overline{x}}))\frac{\theta_0\sigma^2 + m\tau^2\overline{x}_j}{\sigma^2 + m\tau^2} \qquad (2.2) \\
&= kNE[\theta_{[k]}|\underset{\sim}{\overline{X}} = \underset{\sim}{\overline{x}}] - m\sum_{j=1}^{k}\frac{\theta_0\sigma^2 + m\tau^2\overline{x}_j}{\sigma^2 + m\tau^2} - \sum_{j=1}^{k}N_j(\underset{\sim}{\overline{x}})\frac{\theta_0\sigma^2 + m\tau^2\overline{x}_j}{\sigma^2 + m\tau^2}.
\end{aligned}
$$

For each observed $\underset{\sim}{\overline{X}} = \underset{\sim}{\overline{x}}$, let $A(\underset{\sim}{\overline{x}}) = \{i|\overline{x}_i = \max_{1\le j\le k} \overline{x}_j\}$. Then, for a fixed m, the Bayes allocation at the second stage is to choose the nonnegative integers $N_1(\underset{\sim}{\overline{x}}), \ldots N_k(\underset{\sim}{\overline{x}})$ such that $\sum_{i\in A(\underset{\sim}{\overline{x}})} N_i(\underset{\sim}{\overline{x}}) = k(N-m)$. Then, conditional on m and the observed value $\underset{\sim}{\overline{X}} = \underset{\sim}{\overline{x}}$, the minimum posterior expected loss is:

$$
\begin{aligned}
r_m^B(\underset{\sim}{\overline{x}}) = kNE[\theta_{[k]}|\underset{\sim}{\overline{X}} = \underset{\sim}{\overline{x}}] - m\sum_{j=1}^{k}\frac{\theta_0\sigma^2 + m\tau^2\overline{x}_j}{\sigma^2 + m\tau^2} \\
- \frac{k(N-m)[\theta_0\sigma^2 + m\tau^2 \max_{1\le i\le k}\overline{x}_i]}{\sigma^2 + m\tau^2},
\end{aligned} \qquad (2.3)
$$

and the minimum Bayes risk for a fixed m is:

$$
\begin{aligned}
r_m^B &= E[r_m^B(\underset{\sim}{\overline{X}})] \\
&= kNE[\theta_{[k]}] - m\sum_{j=1}^{k}\frac{\theta_0\sigma^2 + m\tau^2E[\overline{X}_j]}{\sigma^2 + m\tau^2} \qquad (2.4) \\
&\quad - \frac{k(N-m)\{\theta_0\sigma^2 + m\tau^2E[\max_{1\le j\le k}\overline{X}_j]\}}{\sigma^2 + m\tau^2}.
\end{aligned}
$$

Note that under the statistical model, $\overline{X}_1, \ldots, \overline{X}_k$ are iid and have a marginal normal distribution with mean θ_0 and variance $\frac{\sigma^2}{m}+\tau^2$. Thus, $E[\max_{1\le i\le k}\overline{X}_i] = \theta_0+\sqrt{\frac{\sigma^2}{m}+\tau^2}E[\max_{1\le j\le k}Z_j]$ $= \theta_0 + \sqrt{\frac{\sigma^2}{m}+\tau^2}\ \alpha$, where $Z_1, \ldots, Z_k$ are iid $N(0,1)$ and $\alpha = E[\max_{1\le j\le k}Z_j]$. Also, $E[\theta_{[k]}] = \theta_0 + \tau\alpha$. Hence, we have

$$
r_m^B = k\tau\alpha\{N - \frac{(N-m)\sqrt{m}\ \tau}{\sqrt{\sigma^2 + m\tau^2}}\}. \qquad (2.5)
$$

Note that the minimum Bayes risk r_m^B does not depend on the parameter θ_0.

2.2. Optimal Initial Sample Size

Next, we want to find an integer, say m_B, $0 \le m_B \le N$ such that $r_{m_B}^B \le r_m^B$ for all integers m in $[0, N]$. We call such an integer m_B as an optimal initial sample size. When m_B is determined, a Bayes two-stage allocation procedure, say P_B, is given as follows:

First, take m_B random observations from each of the k populations. Compute the observed sample mean $\overline{x}_i$, $i = 1, \ldots, k$. Then, take $k(N - m_B)$ random observations from the population which yields the largest sample mean value.

Note that finding an integer m in $[0, N]$ to minimize the Bayes risk r_m^B is equivalent to finding an integer m in $[0, N]$ to maximize $(N-m)\sqrt{m}/\sqrt{\sigma^2 + m\tau^2}$ [see (2.5)]. In general, we assume m to be a variable and for each fixed $\tau^2 > 0$, let

$$H_{\tau^2}(m) = \frac{(N-m)^2 m}{\sigma^2 + m\tau^2} \tag{2.6}$$

be a function defined on the interval $[0, N]$. Then, the first derivative of the function $H_{\tau^2}(m)$ with respect to m is

$$H'_{\tau^2}(m) = \frac{(m-N)[(3m-N)\sigma^2 + 2m^2\tau^2]}{(\sigma^2 + m\tau^2)^2},$$

which is nonpositive if $\frac{N}{3} \leq m \leq N$. That is, the function $H_{\tau^2}(m)$ is nonincreasing in m for m in the interval $[\frac{N}{3},\ N]$. Thus, to find a number m in the interval $[0, N]$ to maximize the function $H_{\tau^2}(m)$, it suffices to consider those m in the subinterval $[0,\ \frac{N}{3}]$. Let

$$G(m) = (m-N)[(3m-N)\sigma^2 + 2m^2\tau^2],\ m \in [0,\ \frac{N}{3}].$$

Then,

$$G'(m) = (3m - 2N)(2\sigma^2 + 2m\tau^2) < 0, \quad \text{for all } m \in [0,\ \frac{N}{3}].$$

In other words, $G(m)$ is a decreasing function of m for $m \in [0,\ \frac{N}{3}]$. Also, note that $G(0) > 0$, $G(\frac{N}{3}) < 0$. Thus, there exists a unique number in $(0,\ \frac{N}{3})$, say m^*, such that $G(m^*) = 0$. Hence, $H'_{\tau^2}(m) > 0$ for all $m \in [0, m^*)$; $H'_{\tau^2}(m) < 0$ for all $m \in (m^*,\ \frac{N}{3})$, and $H'_{\tau^2}(m^*) = 0$. This implies that the function $H_{\tau^2}(m)$ achieves its maximum at $m = m^*$. Note that m^* is the positive solution of the equation $(3m - N)\sigma^2 + 2m^2\tau^2 = 0$. That is,

$$\begin{aligned} m^* &= (-3\sigma^2 + \sqrt{8N\tau^2\sigma^2 + 9\sigma^4})/(4\tau^2) \\ &= 2N\sigma/[\sqrt{8N\tau^2 + 9\sigma^2} + 3\sigma]. \end{aligned} \tag{2.7}$$

Let

$$m_B = \begin{cases} [m^*] & \text{if } H_{\tau^2}([m^*]) \geq H_{\tau^2}([m^*]+1), \\ [m^*]+1 & \text{if } H_{\tau^2}([m^*]) < H_{\tau^2}([m^*]+1). \end{cases} \tag{2.8}$$

Therefore, the minimum Bayes risk, denoted by r^B, of the Bayes two-stage allocation procedure is:

$$r^B = k\tau\alpha\{N - \frac{(N - m_B)\sqrt{m_B}\ \tau}{\sqrt{\sigma^2 + m_B\tau^2}}\}. \tag{2.9}$$

Remarks 2.1

a) For fixed N and σ^2, the optimal initial sample size m_B can be viewed as a function of the parameter τ^2, and hence is denoted by $m_B(\tau^2)$. From (2.6), (2.7), (2.8), one can see that

$$1 \leq m_B(\tau^2) \leq [\frac{N}{3}] + 1 \text{ for any } \tau^2 > 0.$$

Furthermore, we have the following results:

$$\lim_{\tau^2 \to \infty} m_B(\tau^2) = 1 \text{ and}$$

$$\lim_{\tau^2 \to 0} m_B(\tau^2) = \begin{cases} [\frac{N}{3}] & \text{if } N \equiv 0 \text{ or } 1 \text{ (mod 3)}, \\ [\frac{N}{3}] + 1 & \text{if } N \equiv 2 \text{ (mod 3)}. \end{cases}$$

b) From (2.7), m^* is a decreasing function of the parameter τ^2. Thus, from (2.8), one may expect that $m_B(\tau^2)$ is nonincreasing in τ^2. Actually, we have the following results:

$$\begin{cases} \text{If} & \tau_1^2 > \tau_2^2, \text{ then } m_B(\tau_1^2) \leq m_B(\tau_2^2) \\ \text{If} & m_B(\tau_1^2) < m_B(\tau_2^2), \text{ then } \tau_1^2 > \tau_2^2, \end{cases} \tag{2.10}$$

which can be obtained directly from the following lemma.

Lemma 2.1. Let $H_{\tau^2}(m) = \frac{(N-m)^2 m}{\sigma^2 + m\tau^2}$, $1 \leq m \leq [\frac{N}{3}] + 1$ and $\tau^2 > 0$. If $H_{\tau_2^2}(m) \geq H_{\tau_2^2}(m+1)$, then $H_{\tau_1^2}(m) > H_{\tau_1^2}(m+1)$ for all $\tau_1^2 > \tau_2^2$.

Proof: By the given condition,

$$\begin{aligned} 0 &\leq H_{\tau_2^2}(m) - H_{\tau_2^2}(m+1) \\ &= \frac{(N-m)^2 m[\sigma^2 + (m+1)\tau_2^2] - (\sigma^2 + m\tau_2^2)(N-m-1)^2(m+1)}{(\sigma^2 + m\tau_2^2)[\sigma^2 + (m+1)\tau_2^2]}. \end{aligned}$$

Let

$$h(\tau^2) = (N-m)^2 m[\sigma^2 + (m+1)\tau^2] - (\sigma^2 + m\tau^2)(N-m-1)^2(m+1).$$

Hence, $h(\tau_2^2) \geq 0$. Also, the first derivative of $h(\tau^2)$ with respect to τ^2 is

$$h'(\tau^2) = \frac{d\, h(\tau^2)}{d\tau^2} = m(m+1)[2(N-m) - 1] > 0 \text{ for all } 1 \leq m \leq [\frac{N}{3}] + 1,$$

which implies $h(\tau^2)$ is an increasing function of τ^2. Thus, $h(\tau_1^2) > h(\tau_2^2) \geq 0$ since $\tau_1^2 > \tau_2^2$. Therefore, we have $H_{\tau_1^2}(m) > H_{\tau_1^2}(m+1)$.

3. An Adaptive Two-Stage Allocation Procedure

In this section, we still assume the normal model except that the value of the parameter τ^2 is unknown. Thus, the Bayes two-stage allocation procedure derived in Section 2 can not be applied in this situation. To overcome this difficulty, we propose an adaptive two-stage allocation procedure via the empirical Bayes approach.

We now consider the following situation. Suppose that one is confronted repeatedly and independently with a sequence of the allocation problems as described in Section 1. We can then use the past observations at hand to construct an estimator for the unknown parameter τ^2. This estimator is then applied to form an adaptive two-stage allocation procedure for the next allocation problem. Suppose now, we are at time $t = n + 1$. We have already had n past observations at hand. We let m_j denote the adaptive optimal initial sample size taken at stage one at time $t = j$, $j = 1, \ldots, n$. The determination of m_j will be described later. From Remark 2.1 a), $1 \leq m_j \leq [\frac{N}{3}] + 1$. That is, we take at least one observation from each of the k populations at each time $j = 1, \ldots, n$. We let X_{ij} denote an observation taken from population π_i at time j, $j = 1, \ldots, n$. Then, under the normal model, X_{ij} has a marginal normal distribution with mean θ_0 and variance $\sigma^2 + \tau^2$. Also, following the usual empirical Bayes formulation (for example, see Robbins [5] or Gupta and Liang [1]), we can assume that X_{ij}, $j = 1, \ldots, n$; $i = 1, \ldots, k$, are independently distributed. In the following, we only consider the case when the parameter θ_0 is unknown. Let

$$\begin{cases} \overline{X}(n) = \frac{1}{kn} \sum_{i=1}^{k} \sum_{j=1}^{n} X_{ij}, \\ S^2(n) = \frac{1}{kn-1} \sum_{i=1}^{k} \sum_{j=1}^{n} (X_{ij} - \overline{X}(n))^2. \end{cases} \tag{3.1}$$

Then, $(kn-1)S^2(n)/(\sigma^2 + \tau^2)$ has a χ^2-distribution with $kn - 1$ degrees of freedom. Since τ^2 is positive, we suggest using

$$\tau_{n+1}^2 = (S^2(n) - \sigma^2)^+ \tag{3.2}$$

to estimate the unknown parameter τ^2, where $y^+ = \max(0, y)$. When $\tau_{n+1}^2 > 0$, we define m_{n+1}, the adaptive optimal initial sample size at time $t = n + 1$, to be an integer in the interval $[0, N]$ which maximizes the function $H_{\tau_{n+1}^2}(m) = \frac{(N-m)^2 m}{\sigma^2 + m\tau_{n+1}^2}$ among all the integers in the interval $[0, N]$. From Remark 2.1 a), $1 \leq m_{n+1} \leq [\frac{N}{3}] + 1$. When $\tau_{n+1}^2 = 0$, we let $m_{n+1} = [\frac{N}{3}]$ (or $[\frac{N}{3}] + 1$) if $H_{\tau_{n+1}^2}([\frac{N}{3}]) \geq (<) H_{\tau_{n+1}^2}([\frac{N}{3}] + 1)$. Note that when $n = 0$, i.e. there is no past observation available, we arbitrarily choose an integer m_1 in the interval $[1, [\frac{N}{3}] + 1]$ as the initial sample size.

We then propose an adaptive two-stage allocation procedure, say P_{n+1}, at $t = n + 1$ as follows:

At time $t = n+1$, first take m_{n+1} observations from each of the k populations. Compute the observed sample mean $\overline{x}_i$ based on the m_{n+1} observations taken from population π_i, $i = 1, \ldots k$. Then, take $k(N - m_{n+1})$ random observations from the population which yields the largest sample mean value.

We denote the conditional Bayes risk given m_{n+1} and the Bayes risk of the adaptive two-stage allocation procedure P_{n+1} by $r_{n+1}(m_{n+1})$ and r_{n+1}, respectively. That is,

$$\begin{cases} r_{n+1}(m_{n+1}) = k\tau\alpha\{N - (N - m_{n+1})\sqrt{m_{n+1}}\ \tau/\sqrt{\sigma^2 + m_{n+1}\tau^2}\}, \\ r_{n+1} = E[r_{n+1}(m_{n+1})]; \end{cases} \tag{3.3}$$

where the expectation E is taken with respecpt to m_{n+1} or the probability space generated by $(X_{ij},\ j = 1, \ldots, n,\ i = 1, \ldots, k)$.

Note that $r_{n+1}(m_{n+1}) - r^B \geq 0$ since r^B is the minimum Bayes risk, and therefore $r_{n+1} - r^B \geq 0$. The two differences $r_{n+1}(m_{n+1}) - r^B$ and $r_{n+1} - r^B$ are always used as measures of the performance of the proposed two-stage allocation procedure P_{n+1}.

Definition 3.1

a) The sequence of adaptive two-stage allocation procedures $\{P_{n+1}\}$ is said to be asymptotically optimal in probability of order $\{\alpha_n\}$ if for any $\varepsilon > 0$, $P\{r_{n+1}(m_{n+1}) - r^B \geq \varepsilon\} \leq 0(\alpha_n)$ as $n \to \infty$ where $\{\alpha_n\}$ is a sequence of positive numbers such that $\lim_{n\to\infty} \alpha_n = 0$.

b) The sequence of adaptive two-stage allocation procedures $\{P_{n+1}\}$ is said to be asymptotically optimal of order $\{\beta_n\}$ if $r_{n+1} - r^B \leq 0(\beta_n)$ as $n \to \infty$ where $\{\beta_n\}$ is a sequence of positive numbers such that $\lim_{n\to\infty} \beta_n = 0$.

In the following, we will investigate some asymptotically optimal properties of the proposed adaptive two-stage allocation procedures $\{P_{n+1}\}$.

Let $I = \{m | m$ is an integer in $[1, [\frac{N}{3}] + 1]$ such that $H_{\tau^2}(m_B) - H_{\tau^2}(m) \neq 0\}$, and let $c = \min\{H_{\tau^2}(m_B) - H_{\tau^2}(m) | m \in I\}$. Then, by the definitions of m_B and the set I, $c > 0$.

Lemma 3.1.

a) Suppose that $m_{n+1} \in I$ and $m_{n+1} < m_B$. Then

$$c \leq H_{\tau^2}(m_B) - H_{\tau^2}(m_{n+1}) \leq d^{-1}(\tau^2_{n+1} - \tau^2)$$

where $d^{-1} = N^4/(16\sigma^4)$.

b) Suppose that $m_{n+1} \in I$ and $m_{n+1} > m_B$. Then,

$$c \leq H_{\tau^2}(m_B) - H_{\tau^2}(m_{n+1}) \leq d^{-1}(\tau^2 - \tau^2_{n+1}).$$

Proof:

a) By Lemma 2.1, as $m_{n+1} \in I$ and $m_B > m_{n+1}$, we have $\tau^2 < \tau^2_{n+1}$. Thus, on the event that $m_{n+1} \in I$ and $\tau^2 < \tau^2_{n+1}$, we have

$$\begin{aligned} c &\leq H_{\tau^2}(m_B) - H_{\tau^2}(m_{n+1}) \\ &= \frac{(N-m_B)^2 m_B}{\sigma^2 + m_B\tau^2} - \frac{(N-m_{n+1})^2 m_{n+1}}{\sigma^2 + m_{n+1}\tau^2} \\ &= [\frac{(N-m_B)^2 m_B}{\sigma^2 + m_B\tau^2} - \frac{(N-m_B)^2 m_B}{\sigma^2 + m_B\tau^2_{n+1}}] \\ &\quad + [\frac{(N-m_B)^2 m_B}{\sigma^2 + m_B\tau^2_{n+1}} - \frac{(N-m_{n+1})^2 m_{n+1}}{\sigma^2 + m_{n+1}\tau^2_{n+1}}] \\ &\quad + [\frac{(N-m_{n+1})^2 m_{n+1}}{\sigma^2 + m_{n+1}\tau^2_{n+1}} - \frac{(N-m_{n+1})^2 m_{n+1}}{\sigma^2 + m_{n+1}\tau^2}]. \end{aligned} \tag{3.4}$$

In (3.4), $(N-m_B)^2 m_B/(\sigma^2 + m_B\tau^2_{n+1}) - (N-m_{n+1})^2 m_{n+1}/(\sigma^2 + m_{n+1}\tau^2_{n+1}) \leq 0$ which is obtained by the definition of m_{n+1}, and $(N-m_{n+1})^2 m_{n+1}/(\sigma^2 + m_{n+1}\tau^2_{n+1}) - (N - m_{n+1})^2 m_{n+1}/(\sigma^2 + m_{n+1}\tau^2) < 0$ by noting that $\tau^2 < \tau^2_{n+1}$. Thus, we obtain

$$\begin{aligned} c &\leq H_{\tau^2}(m_B) - H_{\tau^2}(m_{n+1}) \\ &\leq \frac{(N-m_B)^2 m_B}{\sigma^2 + m_B\tau^2} - \frac{(N-m_B)^2 m_B}{\sigma^2 + m_B\tau^2_{n+1}} \\ &= \frac{(N-m_B)^2 m_B^2(\tau^2_{n+1} - \tau^2)}{(\sigma^2 + m_B\tau^2)(\sigma^2 + m_B\tau^2_{n+1})} \\ &\leq \frac{N^4}{16\sigma^4}(\tau^2_{n+1} - \tau^2) \\ &= d^{-1}(\tau^2_{n+1} - \tau^2) \end{aligned} \tag{3.5}$$

which completes the proof of part a).

b) By Lemma 2.1 again, as $m_{n+1} \in I$ and $m_B < m_{n+1}$, we have $\tau^2 > \tau_{n+1}^2$. Thus, under the event that $m_{n+1} \in I$ and $\tau^2 > \tau_{n+1}^2$, we have

$$
\begin{aligned}
c &\le H_{\tau^2}(m_B) - H_{\tau^2}(m_{n+1}) \\
&= \frac{(N-m_B)^2 m_B}{\sigma^2 + m_B\tau^2} - \frac{(N-m_{n+1})^2 m_{n+1}}{\sigma^2 + m_{n+1}\tau^2} \\
&= \Big[\frac{(N-m_B)^2 m_B}{\sigma^2 + m_B\tau^2} - \frac{(N-m_B)^2 m_B}{\sigma^2 + m_B\tau_{n+1}^2}\Big] + \Big[\frac{(N-m_B)^2 m_B}{\sigma^2 + m_B\tau_{n+1}^2} - \frac{(N-m_{n+1})^2 m_{n+1}}{\sigma^2 + m_{n+1}\tau_{n+1}^2}\Big] \\
&\quad + \Big[\frac{(N-m_{n+1})^2 m_{n+1}}{\sigma^2 + m_{n+1}\tau_{n+1}^2} - \frac{(N-m_{n+1})^2 m_{n+1}}{\sigma^2 + m_{n+1}\tau^2}\Big],
\end{aligned} \tag{3.6}
$$

where $(N-m_B)^2 m_B/(\sigma^2+m_B\tau^2) - (N-m_B)^2 m_B/(\sigma^2+m_B\tau_{n+1}^2) < 0$ since $\tau^2 > \tau_{n+1}^2$ and $(N-m_B)^2 m_B/(\sigma^2+m_B\tau_{n+1}^2) - (N-m_{n+1})^2 m_{n+1}/(\sigma^2+m_{n+1}\tau_{n+1}^2) \le 0$, by the definition of m_{n+1}. Therefore,

$$
\begin{aligned}
c &\le H_{\tau^2}(m_B) - H_{\tau^2}(m_{n+1}) \\
&\le \frac{(N-m_{n+1})^2 m_{n+1}}{\sigma^2 + m_{n+1}\tau_{n+1}^2} - \frac{(N-m_{n+1})^2 m_{n+1}}{\sigma^2 + m_{n+1}\tau^2} \\
&= \frac{(N-m_{n+1})^2 m_{n+1}^2(\tau^2 - \tau_{n+1}^2)}{(\sigma^2 + m_{n+1}\tau_{n+1}^2)(\sigma^2 + m_{n+1}\tau^2)} \\
&\le d^{-1}(\tau^2 - \tau_{n+1}^2).
\end{aligned} \tag{3.7}
$$

Lemma 3.2.

a) $P\{r_{n+1}(m_{n+1}) > r_B\} \le P\{|\tau_{n+1}^2 - \tau^2| \ge dc\}$.

b) $r_{n+1} - r_B \le k\alpha\tau^2[H_{\tau^2}(m_B)]^{\frac{1}{2}} P\{|\tau_{n+1}^2 - \tau^2| \ge dc\}$.

Proof:

a)

$$
\begin{aligned}
& P\{r_{n+1}(m_{n+1}) > r_B\} \\
&= P\{H_{\tau^2}(m_B) - H_{\tau^2}(m_{n+1}) > 0,\ m_{n+1} \in I\} \\
&= P\{H_{\tau^2}(m_B) - H_{\tau^2}(m_{n+1}) \ge c,\ m_{n+1} \in I\} \\
&\qquad \text{(by the definition of the set } I) \\
&= P\{H_{\tau^2}(m_B) - H_{\tau^2}(m_{n+1}) \ge c,\ m_{n+1} \in I,\ m_B < m_{n+1}\} \\
&\qquad + P\{H_{\tau^2}(m_B) - H_{\tau^2}(m_{n+1}) \ge c,\ m_{n+1} \in I,\ m_B > m_{n+1}\} \\
&\le P\{\tau^2 - \tau_{n+1}^2 \ge dc\} + P\{\tau_{n+1}^2 - \tau^2 \ge dc\} \\
&\qquad \text{(by Lemma 3.1)} \\
&= P\{|\tau_{n+1}^2 - \tau^2| \ge dc\}.
\end{aligned}
$$

b)

$$\begin{aligned}
& r_{n+1} - r_B \\
&= E[r_{n+1}(m_{n+1}) - r_B] \\
&= E[k\alpha\tau^2[(H_{\tau^2}(m_B))^{\frac{1}{2}} - (H_{\tau^2}(m_{n+1}))^{\frac{1}{2}}]] \\
&\leq k\alpha\tau^2[H_{\tau^2}(m_B)]^{\frac{1}{2}} P\{H_{\tau^2}(m_B) - H_{\tau^2}(m_{n+1}) > 0\} \\
&= k\alpha\tau^2[H_{\tau^2}(m_B)]^{\frac{1}{2}} P\{H_{\tau^2}(m_B) - H_{\tau^2}(m_{n+1}) \geq c\} \\
&\leq k\alpha\tau^2[H_{\tau^2}(m_B)]^{\frac{1}{2}} P\{|\tau_{n+1}^2 - \tau^2| \geq dc\}
\end{aligned}$$

where the last equality is obtained from the definition of the constant c, and the last inequality is obtained from the proof of part a) of this lemma.

From Lemma 3.2, in order to investigate the asymptotic behavior of $P\{r_{n+1}(m_{n+1}) > r_B\}$ and $r_{n+1} - r_B$, it suffices to study the asymptotic behavior of the probability $P\{|\tau_{n+1}^2 - \tau^2| \geq dc\}$.

Lemma 3.3. Let $\{\tau_{n+1}^2\}_{n=1}^{\infty}$ be a sequence of estimators defined in (3.2). Then, τ_{n+1}^2 converges to τ^2 in probability. Furthermore, for any $\varepsilon > 0$, we have $P\{|\tau_{n+1}^2 - \tau^2| \geq \varepsilon\} \leq 0(n^{-1})$ as $n \to \infty$.

Proof: First note that $Y \equiv \frac{(kn-1)S^2(n)}{\sigma^2+\tau^2}$ follows a χ^2-distribution with $(kn-1)$ degrees of freedom. By the definition of τ_{n+1}^2 given in (3.2), letting $\varepsilon_1 = \varepsilon/(\sigma^2 + \tau^2)$, we have

$$\begin{aligned}
& P\{|\tau_{n+1}^2 - \tau^2| \geq \varepsilon\} \\
&= P\{\tau_{n+1}^2 \geq \tau^2 + \varepsilon\} + P\{\tau_{n+1}^2 \leq \tau^2 - \varepsilon\} \\
&\leq P\{S^2(n) \geq \tau^2 + \sigma^2 + \varepsilon\} + P\{S^2(n) \leq \tau^2 + \sigma^2 - \varepsilon\} \\
&= P\{Y \geq (kn-1)(1+\varepsilon_1)\} + P\{Y \leq (kn-1)(1-\varepsilon_1)\} \\
&= P\Big\{\Big|\frac{Y-(kn-1)}{\sqrt{2(kn-1)}}\Big| \geq \sqrt{\frac{kn-1}{2}}\varepsilon_1\Big\} \\
&\leq \frac{2}{(kn-1)\varepsilon_1^2}
\end{aligned}$$

which can be obtained by Chebyschev's inequality. Hence we obtain that

$$P\{|\tau_{n+1}^2 - \tau^2| \geq \varepsilon\} \leq 0(n^{-1}) \text{ as } n \to \infty.$$

From Lemmas 3.2 and 3.3, we conclude the following theorem.

Theorem 3.1. The sequence of adaptive two-stage allocation procedures $\{P_{n+1}\}$ is asymptotically optimal in probability of order $\{n^{-1}\}$ and asymptotically optimal of order $\{n^{-1}\}$. That is,

$$P\{r_{n+1}(m_{n+1}) - r^B \geq \varepsilon\} \leq 0(n^{-1}) \text{ as } n \to \infty \text{ for any } \varepsilon > 0,$$

and

$$r_{n+1} - r^B \leq 0(n^{-1}) \text{ as } n \to \infty.$$

REFERENCES

[1] Gupta, S. S. and Liang, T. (1987). On some Bayes and empirical Bayes selection procedures. In Probability and Bayesian Statistics (Ed. R. Viertl), Plenum Publishing Corporation, 233–246.

[2] Gupta, S. S. and Panchapakesan, S. (1979). Multiple Decision Procedures. New York, John Wiley.

[3] Ofosu, J. B. (1974). Optimum sample size for a problem in choosing the population with the largest mean: some comments on Somerville's paper. Journal of American Statistical Association **69**, 270.

[4] Ofosu, J. B. (1975). A Bayesian procedure for selecting the largest of k normal means. Calcutta Statistical Association Bulletin **24**, 43–51.

[5] Robbins, H. (1983). Some thoughts on empirical Bayes estimation. Annals of Statistics **11**, 713–723.

[6] Somerville, P. N. (1970). Optimum sample size for a problem in choosing the population with the largest mean. Journal of American Statistical Association **65**, 763–775.

[7] Somerville, P. N. (1974). Rejoinder to Ofosu's comments. Journal of American Statistical Association **69**, 270–271.

REPORT DOCUMENTATION PAGE

1a. REPORT SECURITY CLASSIFICATION

1b. RESTRICTIVE MARKINGS

2a. SECURITY CLASSIFICATION AUTHORITY
Unclassified

2b. DECLASSIFICATION / DOWNGRADING SCHEDULE

3. DISTRIBUTION / AVAILABILITY OF REPORT

4. PERFORMING ORGANIZATION REPORT NUMBER(S)
Technical Report #87-53

5. MONITORING ORGANIZATION REPORT NUMBER(S)

6a. NAME OF PERFORMING ORGANIZATION
Purdue University

6b. OFFICE SYMBOL *(If applicable)*

7a. NAME OF MONITORING ORGANIZATION

6c. ADDRESS *(City, State, and ZIP Code)*
Department of Statistics
West Lafayette, IN 47907

7b. ADDRESS *(City, State, and ZIP Code)*

8a. NAME OF FUNDING / SPONSORING ORGANIZATION
Office of Naval Research

8b. OFFICE SYMBOL *(If applicable)*

9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER
N00014-84-C-0167 and NSF Grant DMS-8606964

8c. ADDRESS *(City, State, and ZIP Code)*
Arlington, VA 22217-5000

10. SOURCE OF FUNDING NUMBERS

PROGRAM ELEMENT NO.	PROJECT NO.	TASK NO.	WORK UNIT ACCESSION NO.

11. TITLE *(Include Security Classification)*
ON BAYES AND EMPIRICAL BAYES TWO-STAGE ALLOCATION PROCEDURES FOR SELECTION PROBLEMS

12. PERSONAL AUTHOR(S)
Shanti S. Gupta and TaChen Liang

13a. TYPE OF REPORT
Technical

13b. TIME COVERED
FROM ________ TO ________

14. DATE OF REPORT *(Year, Month, Day)*
December 8, 1987

15. PAGE COUNT
10

16. SUPPLEMENTARY NOTATION

17. COSATI CODES

FIELD	GROUP	SUB-GROUP

18. SUBJECT TERMS *(Continue on reverse if necessary and identify by block number)*
Allocation, selection, asymptotically optimal, Bayes, empircal Bayes, initial sample size, two-stage procedure.

19. ABSTRACT *(Continue on reverse if necessary and identify by block number)*
This paper deals with the problem of deriving two-stage allocation procedures for selecting the best normal population. If the prior distribution is assumed to be known, an exact Bayes two-stage allocation procedure is obtained. If the prior distribution depends on some unknown parameter, an adaptive two-stage allocation procedure is proposed. Using the empirical Bayes formulation, we prove that the proposed adaptive two-stage allocation procedure has some asymptotic optimality property.

20. DISTRIBUTION / AVAILABILITY OF ABSTRACT
☒ UNCLASSIFIED/UNLIMITED ☐ SAME AS RPT. ☐ DTIC USERS

21. ABSTRACT SECURITY CLASSIFICATION
Unclassified

22a. NAME OF RESPONSIBLE INDIVIDUAL
Shanti S. Gupta

22b. TELEPHONE *(Include Area Code)*
317-494-6031

22c. OFFICE SYMBOL