SIMULTANEOUSLY SELECTING NORMAL POPULATIONS CLOSE TO A CONTROL*

by

TaChen Liang
Wayne State University

Technical Report # 95-26C

Department of Statistics
Purdue University

June 1995
Revised June 1996

*

SIMULTANEOUSLY SELECTING NORMAL POPULATIONS CLOSE TO A CONTROL*

TaChen Liang

Department of Mathematics
Wayne State University
Detroit, MI 48202

Abstract

We study the problem of selecting populations close to a control from among k normal populations using the parametric empirical Bayes approach. A Bayes selection rule is derived, which depends on certain parameters. When those parameters are unknown, using the empirical Bayes idea, we first present estimators, based on information collected from the k populations for the unknown parameters. Then, mimicking the behavior of the Bayes selection rule, an empirical Bayes selection rule is constructed. The relative regret Bayes risk is used as a measure of performance of the empirical Bayes selection rule. It is shown that the relative regret Bayes risk of the proposed empirical Bayes selection rule converges to zero at a rate of order $O(k^{-1})$. A simulation study is also carried out to investigate the performance of the proposed empirical Bayes selection rule for small to moderate values of k.

AMS 1991 Subject Classification: Primary 62F07; Secondary 62C12.

Keywords and phrases: Asymptotically optimal; empirical Bayes, rate of convergence; relative regret Bayes risk; simultaneous selection.

*This research was supported in part by US Army Research Office, Grant DAAH04–95–1-0165 under the direction of Professor S. S. Gupta

1. Introduction

Problems of selecting populations close to a control arise frequently in many applications. Assume that there are k populations and a target value used as a control. Our goal is to select those populations which are sufficiently close to the control. To motivate such a study, for example, consider the matching parts problem in industrial production described in Burr (1976). A diesel engine plant has to make plunger rods for forcing fuel through small holes. The diameter of the plunger rods should meet certain specification limits. Suppose there are several plunger rods. Then, we may wish to select one or all those rods which meet the specification limits. Also, as described by Wellek and Michaelis (1991), such selection problem arises in clinical trials and bioavailability trials, for instance, to identify the equivalence of a newly developed formulation of a drug with different administration methods to a standard formulation.

In the literature, Gupta and Singh (1979) and Gupta and Hsiao (1981) have derived Bayes, Γ-minimax and minimax procedures for selecting populations close to a control. Mee, Shah and Lefante (1987) have developed multiple testing procedures to compare the means of k normal populations with respect to a control. Giani and Straβburger (1994) have studied testing and selection procedures for equivalence of k populations with respect to a control. It should be noted that comparing populations with a control under different types of formulation has been investigated in the literature. To mention a few, for example, Bechhofer and Turnbull (1978), Dunnett (1984), Wilcox (1984) and Gupta, Liang and Rau (1994) have discussed problems of selecting the best population provided that the best is better than a control. Paulson (1952) and Gupta and Sobel (1958) have studied problems of selecting a subset containing all populations better than a control. Randles and Hollander (1971) and Miescke (1981) have derived optimal selection rules via the Γ-minimax and minimax approaches for selecting good populations. Huang (1975) has derived Bayes selection rules to partition normal populations.

In this paper, we study the problem of selecting populations close to a control from among k normal populations according to the Kullback-Leibler discrimination information. We will derive empirical Bayes simultaneous selection rules for this selection problem and investigate the corresponding optimality of the empirical Bayes selection rule. The paper is organized as follows. The framework of the selection problem is given in Section 2. An empirical Bayes selection rule is proposed is Section 3. The asymptotic optimality of the empirical Bayes selection rule is investigated in Section 4. The relative regret Bayes risk is used as a measure of performance of the empirical Bayes selection rule. It is shown that the relative regret Bayes risk of the concerned empirical Bayes selection rule converges to zero with a rate of order $O(k^{-1})$. A simulation study is carried out to investigate the performance of the empirical Bayes selection rule for small to moderate values of k.

2. Formulation of the Selection Problem and A Bayes Selection Rule

Consider k independent normal populations $\pi_1, \ldots, \pi_k$ with unknown means $\theta_1, \ldots, \theta_k$, respectively, and a common variance σ^2. Let θ_0 be a known control value. Also, let X_i denote a random variable arising from population π_i and $f(x|\theta_i, \sigma^2)$ denote the density of a $N(\theta_i, \sigma^2)$ distribution. Then, the distance between population π_i and the control θ_0 is defined as:

$$\delta_i = E_{\theta_i}\left[\ell n \frac{f(X_i|\theta_i, \sigma^2)}{f(X_i|\theta_0, \sigma^2)}\right] = \frac{(\theta_i - \theta_0)^2}{2\sigma^2} \tag{2.1}$$

the Kullback-Leibler discrimination information between two normal distributions $N(\theta_i, \sigma^2)$ and $N(\theta_0, \sigma^2)$. Note that δ_i is increasing in $|\theta_i - \theta_0|$ and $\delta_i = 0$ as $\theta_i = \theta_0$.

For a given constant $c > 0$, population π_i is said to be good if $\delta_i \leq c$, and bad otherwise. Our selection goal is to select all good populations (or all populations with at most distance c from the control θ_0) and to exclude all bad populations.

Let $\Omega = \{\underset{\sim}{\theta} = (\theta_1, \ldots, \theta_k, \sigma^2) \; |\theta_i \epsilon R, i = 1, \ldots, k; \; \sigma^2 > 0\}$ be the parameter space. Let $\underset{\sim}{a} = (a_1, \ldots, a_k)$ denote an action, where $a_i = 0, 1, i = 1, \ldots, k$. Whenever action $\underset{\sim}{a}$ is taken, it means that population π_i is selected as good if $a_i = 1$ and excluded as bad if $a_i = 0$. The following loss function is adopted:

$$L(\underset{\sim}{\theta}, \underset{\sim}{a}) = \sum_{i=1}^{k} L_i(\underset{\sim}{\theta}, a_i) \tag{2.2}$$

where for each $i = 1, \ldots, k$,

$$L_i(\underset{\sim}{\theta}, a_i) = a_i(\delta_i - c)I_{(c,\infty)}(\delta_i) + (1 - a_i)(c - \delta_i)I_{[0,c]}(\delta_i), \tag{2.3}$$

where I_s denotes the indicator function of the set S.

In (2.3), the first term is the loss of selecting population π_i as good while π_i is at least c distance away from the control θ_0, and the second term is the loss of wrongly excluding π_i as bad one while π_i is within c distance from the control θ_0.

For each $i = 1, \ldots, k$, let $Y_{i1} \ldots, Y_{im}$ be a sample of size m taken from population π_i. It is assumed that θ_i is a realization of a random variable Θ_i which has a $N(\theta_0, \tau^2)$ prior distribution with unknown common variance τ^2. The random variables $\Theta_1, \ldots, \Theta_k$ are assumed to be mutually independent.

Let $\underset{\sim}{Y}_i = (Y_{i1}, \ldots, Y_{im})$, $i = 1, \ldots, k$, and $\underset{\sim}{Y} = (\underset{\sim}{Y}_1, \ldots, \underset{\sim}{Y}_k)$ and let $\mathcal{Y}$ denote the sample space of $\underset{\sim}{Y}$. A selection rule $\underset{\sim}{d} = (d_1, \ldots, d_k)$ is a mapping defined on the sample space $\mathcal{Y}$ into $[0,1]^k$, such that for each $y \in \mathcal{Y}$, $\underset{\sim}{d}(\underset{\sim}{y}) = (d_1(\underset{\sim}{y}), \ldots, d_k(\underset{\sim}{y}))$ where $d_i(\underset{\sim}{y})$ is the probability of selecting population π_i as good.

Under the preceding statistical model, the Bayes risk of a selection rule $\underset{\sim}{d}$ is:

$$R_k(\underset{\sim}{d}) = \sum_{i=1}^{k} R_{ki}(d_i) \tag{2.4}$$

where

$$R_{ki}(d_i) = \int_{\mathcal{Y}} d_i(\underset{\sim}{y})[\varphi_i(\underset{\sim}{y}_i) - c] \prod_{j=1}^{k} f_j(\underset{\sim}{y}_j) d\underset{\sim}{y} + C_i \tag{2.5}$$

and

$$\begin{cases} C_i = E[(c - \frac{(\Theta_i - \theta_0)^2}{2\sigma^2}) I_{[0,c]}(\frac{(\Theta_i - \theta_0)^2}{2\sigma^2})], \\ f_j(\underset{\sim}{y}_j) \text{ is the marginal probability density of } \underset{\sim}{Y}_j, \\ \varphi_i(\underset{\sim}{y}_i) = E[\frac{(\Theta_i - \theta_0)^2}{2\sigma^2} | \underset{\sim}{Y}_i = \underset{\sim}{y}_i]. \end{cases}$$

Since given $\underset{\sim}{Y}_i = \underset{\sim}{y}_i$, Θ_i has a posterior normal distribution with mean $B\theta_0 + (1-B)\overline{y}_i$ and variance $B\tau^2$, where $\overline{y}_i = \frac{1}{m} \sum_{j=1}^{m} y_{ij}$ and $B = \frac{\sigma^2}{m} / (\frac{\sigma^2}{m} + \tau^2)$, it follows that

$$\begin{aligned} \varphi_i(\underset{\sim}{y}_i) &= \frac{\mathrm{Var}(\Theta_i | \underset{\sim}{Y}_i = \underset{\sim}{y}_i)}{2\sigma^2} + \frac{\{E[\Theta_i | \underset{\sim}{Y}_i = \underset{\sim}{y}_i] - \theta_0\}^2}{2\sigma^2} \\ &= \frac{\frac{\sigma^2}{m}\tau^2}{2\sigma^2(\frac{\sigma^2}{m} + \tau^2)} + \frac{(\overline{y}_i - \theta_0)^2}{2\sigma^2} \times \frac{\tau^4}{(\frac{\sigma^2}{m} + \tau^2)^2} \\ &= \frac{1}{2m}[1 - B] + \frac{(\overline{y}_i - \theta_0)^2}{2\sigma^2}(1 - B)^2 \equiv \psi_i(\overline{y}_i). \end{aligned} \tag{2.6}$$

Hence, a Bayes selection rule $\underset{\sim}{d}_B = (d_{B1}, \ldots, d_{Bk})$, which minimizes the Bayes risks $R_k(\underset{\sim}{d})$ among all selection rules, is given as follows:

For each $\underset{\sim}{y} \epsilon \mathcal{Y}$, and $i = 1, \ldots, k$,

$$d_{Bi}(\underset{\sim}{y}) = \begin{cases} 1 & \text{if } \psi_i(\overline{y}_i) \leq c, \\ 0 & \text{otherwise.} \end{cases} \tag{2.7}$$

From (2.6) and (2.7), we see that for each component i, the Bayes selection rule d_{Bi} is independent of $\underset{\sim}{y}_j$, for all $j \neq i$, and depends on $\underset{\sim}{y}_i$ only through the sample mean value $\overline{y}_i$. Therefore, it can be written as $d_{Bi}(\overline{y})$. That is, $d_{Bi}(\overline{y}_i) = d_{Bi}(\underset{\sim}{y}_i)$. The minimum Bayes risk is:

$$R_k(\underset{\sim}{d}_B) = \sum_{i=1}^{k} R_{ki}(d_{Bi}) \tag{2.8}$$

and

$$R_{ki}(d_{Bi}) = \int_{-\infty}^{\infty} d_{Bi}(\overline{y}_i)[\psi_i(\overline{y}_i) - c]g_i(\overline{y}_i)d\overline{y}_i + C_i \tag{2.9}$$

where $g_i(\overline{y}_i)$ is the marginal pdf of the sample mean $\overline{Y}_i = \frac{1}{m}\sum_{j=1}^{m} Y_{ij}$. According to the statistical model described previously, it is known that $\overline{Y}_1, \ldots, \overline{Y}_k$ are iid, with a normal distribution $N(\theta_0, \frac{\sigma^2}{m} + \tau^2)$.

Let $c^* = 2\sigma^2[c - \frac{1-B}{2m}]/(1-B)^2$. Then $\psi_i(\overline{y}_i) - c \leq 0$ if and only if $(\overline{y}_i - \theta_0)^2 \leq c^*$. If $c^* < 0$, then $\psi_i(\overline{y}_i) - c > 0$ for all $\overline{y}_i$, $i = 1, \ldots, k$. Hence $d_{Bi}(\overline{y}_i) = 0$ for all $\overline{y}_i, i = 1, \ldots, k$. When $c^* > 0$, the Bayes selection rule $\underset{\sim}{d}_B$ can be rewritten as follows: For each $i = 1, \ldots, k$,

$$d_{Bi}(\overline{y}_i) = \begin{cases} 1 & \text{if } \overline{y}_i \;\epsilon\; I, \\ 0 & \text{otherwise,} \end{cases} \tag{2.10}$$

where $I = [\theta_0 - \sqrt{c^*}, \theta_0 + \sqrt{c^*}]$.

Finally, we note that for each i, $\psi_i(\overline{y}_i)$ is increasing in $|\overline{y}_i - \theta_0|$ and $d_{Bi}(\overline{y}_i)$ is nonincreasing in $|\overline{y}_i - \theta_0|$. Also, since when $c^* < 0$, $d_{Bi}(\overline{y}_i) = 0$ for all $\overline{y}_i, i = 1, \ldots, k$, and this may be an extreme case. Hence, in the following analysis, it is assumed that $c > \frac{1}{2m}$ and therefore $c^* > 0$.

3. An Empirical Bayes Selection Rule

It should be noted that the Bayes selection rule $\underset{\sim}{d}_B$ depends on $\psi_i(\overline{y}_i), i = 1, \ldots, k$, which are dependent on σ^2 and τ^2. When the parameters are unknown, the Bayes selection rule $\underset{\sim}{d}_B$ cannot be implemented for the selection problem at hand. The unknown parameters should be estimated. In the following, the parametric empirical Bayes approach is employed for estimating the unknown parameters and deriving a selection rule.

For each $i = 1, \ldots, k$, let $S_i = \sum_{j=1}^{m}(Y_{ij} - \overline{Y}_i)^2$ and $S = \sum_{i=1}^{k} S_i$. It is known that $S_i/\sigma^2 \sim \chi^2_{m-1}, i = 1, \ldots, k$; $S_1, \ldots, S_k$ are mutually independent and hence $S/\sigma^2 \sim \chi^2_{k(m-1)}$. Let $W = \sum_{i=1}^{k}(\overline{Y}_i - \theta_0)^2$. Since $\overline{Y}_1, \ldots, \overline{Y}_k$ are iid, having a $N(\theta_0, \frac{\sigma^2}{m} + \tau^2)$ distribution, $W/(\frac{\sigma^2}{m} + \tau^2) \sim \chi^2_k$. Note that $E[\frac{S}{k(m-1)}] = \sigma^2$, $E[\frac{W}{k}] = \frac{\sigma^2}{m} + \tau^2$. Hence we may use $\hat{B} = \left(\frac{S/[k(m-1)m]}{W/k} \wedge 1\right)$ to estimate B by noting that $B = \frac{\sigma^2}{m}/(\frac{\sigma^2}{m} + \tau^2) < 1$ where $a \wedge b = \min(a, b)$. Also, we use $\hat{\sigma}^2 = \frac{S}{k(m-1)}$ to estimate σ^2.

Define

$$\psi_i^*(\overline{y}_i) = \frac{1}{2m}[1 - \hat{B}] + \frac{(\overline{y}_i - \theta_0)^2}{2\hat{\sigma}^2}[1 - \hat{B}]^2, i = 1, \ldots, k. \tag{3.1}$$

$\psi_i^*(\overline{y}_i)$ is a mimicry of $\psi_i(\overline{y}_i)$ with the unknown parameters B and σ^2 being replaced by the corresponding estimators $\hat{B}$ and $\hat{\sigma}^2$, respectively. Now, an empirical Bayes simultaneous selection rule $\underset{\sim}{d}_k^* = (d_{k1}^*, \ldots, d_{kk}^*)$ is proposed as follows.

For each $i = 1, \ldots, k$; and each $\underset{\sim}{y} \in \mathcal{Y}$, define

$$d_{ki}^*(\overline{y}_i) = d_{ki}^*(\overline{y}_i|\hat{B}, \hat{\sigma}^2) = \begin{cases} 1 & \text{if } \psi_i^*(\overline{y}_i) \leq c, \\ 0 & \text{otherwise.} \end{cases} \tag{3.2}$$

Note that for each $i = 1, \ldots, k$, d_{ki}^* depends on y_i as well as $y_j, j \neq i$, through $\overline{y}_i, \hat{B}$ and $\hat{\sigma}^2$. Also, it can be seen that $\psi_i^*(\overline{y}_i)$ is increasing in $|\overline{y}_i - \theta_0|$ and therefore, $d_{ki}^*(\overline{y}_i|\hat{B}, \hat{\sigma}^2)$ is nonincreasing in $|\overline{y}_i - \theta_0|$.

The Bayes risk of the empirical Bayes selection rule $\underset{\sim}{d}_k^*$ is

$$R_k(\underset{\sim}{d}_k^*) = \sum_{i=1}^{k} R_{ki}\ (d_{ki}^*). \tag{3.3}$$

where

$$
\begin{aligned}
R_{ki}(d_{ki}^*) &= \int E_i[d_{ki}^*(\overline{Y}_i|\hat{B},\hat{\sigma}^2)|\overline{Y}_i=\overline{y}_i][\psi_i(\overline{y}_i)-c]g_i(\overline{y}_i)d\overline{y}_i + C_i \\
&= \int P_i\{d_{ki}^*(\overline{y}_i|\hat{B},\hat{\sigma}^2)=1|\overline{Y}_i=\overline{y}_i\}[\psi_i(\overline{y}_i)-c]g_i(\overline{y}_i)d\overline{y}_i + C_i.
\end{aligned}
\tag{3.4}
$$

In (3.4), P_i is the conditional probability measure generated by $\hat{B}$ and $\hat{\sigma}^2$ conditioning on $\overline{Y}_i = \overline{y}_i$ and E_i is the expectation taken with respect to the conditional probability measure P_i given $\overline{Y}_i = \overline{y}_i$.

Example. The example of Romano (1977, page 248) is used to illustrate the application of the empirical Bayes selection rule $\underset{\sim}{d}_k^*$. Four product lines in an industrial corporation are set to manufacture a specific type of ball bearing with a diameter of 1mm. An experimenter is interested in finding out all those product lines for which the associated Kullback Leibler discrimination information from the control value $\theta_0 = 1mm$ is at most 0.1. For this purpose, at the end of a day's production, ten ball bearings are randomly and independently selected from each of the four lots manufactured by the product lines. The data is given below.

i	1	2	3	4
$\overline{Y}_i$	1.194	1.406	1.129	1.226 (mm)
S_i	0.7552	1.6510	1.5319	0.5328

Note that $\theta_0 = 1mm$, $c = 0.1$, $k = 4$ and $m = 10$. Then, $S = 4.4709$, $W = 0.270189, \hat{\sigma}^2 = 0.124192$ and $\hat{B} = 0.183859$. Hence,

i	1	2	3	4
$\psi_i^*(\overline{y}_i)$	0.1417	0.4828	0.0854	0.1778
$d_{ki}^*(\overline{y}_i)$	0	0	1	0

That is, the empirical Bayes selection rule $\underset{\sim}{d}_k^*$ selects product line 3 only and excludes the other three product lines.

4. Asymptotic Optimality

For a selection rule $\underset{\sim}{d} = (d_1, \ldots, d_k)$ let $R_k(\underset{\sim}{d})$ denote the Bayes risk of $\underset{\sim}{d}$. Since $\underset{\sim}{d}_B$ is the Bayes selection rule. $D_{ki}(d_i) = R_{ki}(d_i) - R_{ki}(d_{Bi}) \geq 0$ for each $i = 1, \ldots, k$. Hence,

$D_k(\underset{\sim}{d}) = R_k(\underset{\sim}{d}) - R_k(\underset{\sim}{d}_B) = \sum_{i=1}^{k} D_{ki}(d_i) \geq 0$. $D_k(\underset{\sim}{d})$ and $\rho_k(\underset{\sim}{d}) = D_k(\underset{\sim}{d})/R_k(\underset{\sim}{d}_B)$ are called regret Bayes risk and relative regret Bayes risk, respectively, of the selection rule $\underset{\sim}{d}$. In the following, the relative regret Bayes risk $\rho_k(\underset{\sim}{d})$ is used as a measure of performance of the selection rule $\underset{\sim}{d}$.

A selection rule $\underset{\sim}{d}$ is said to be asymptotically optimal of order $\{\beta_k\}$ if $\rho_k(\underset{\sim}{d}) = O(\beta_k)$ where $\{\beta_k\}$ is a sequence of positive numbers such that $\lim_{k\to\infty} \beta_k = 0$.

In the following, we will investigate the asymptotic optimality of the empirical Bayes selection rule $\underset{\sim}{d}_k^*$. For doing so, note that under the previously described statistical model, for the Bayes and empirical Bayes selection rules $\underset{\sim}{d}_B$ and $\underset{\sim}{d}_k^*$, we have: $R_{k1}(d_{B1}) = \cdots = R_{kk}(d_{Bk})$ and $R_{k1}(d_{k1}^*) = \cdots = R_{kk}(d_{kk}^*)$. Hence, $\rho_k(\underset{\sim}{d}_k^*) = [R_{k1}(d_{k1}^*) - R_{k1}(d_{B1})]/R_{k1}(d_{B1})$. Since $R_{k1}(d_{B1})$ is fixed for all k, therefore, it suffices to investigate the asymptotic behavior of the regret Bayes risk $R_{k1}(d_{k1}^*) - R_{k1}(d_{B1})$.

Let $c_1(\overline{y}_1) = \psi_1(\overline{y}_1) - c$. Also, let $J = R - I$, the complement of the interval $I = [\theta_0 - \sqrt{c^*}, \theta_0 + \sqrt{c^*}]$. Note that $d_{B1}(\overline{y}_1) = 1, c_1(\overline{y}_1) \leq 0$ if $\overline{y}_1 \epsilon I$; and $d_{B1}(\overline{y}_1) = 0, c_1(\overline{y}_1) > 0$ for $\overline{y}_1 \epsilon J$. From (2.9), (3.4) and the fact that $\underset{\sim}{d}_B$ is a Bayes selection rule, we have that

$$\begin{aligned}
0 \leq & R_{k1}(d_{k1}^*) - R_{k1}(d_{B1}) \\
= & \int E_1[d_{k1}^*(\overline{y}_1|\hat{B}, \hat{\sigma}^2) - d_{B1}(\overline{y}_1)|\overline{Y}_1 = \overline{y}_1] c_1(\overline{y}_1) g_1(\overline{y}_1) d\overline{y}_1 \\
= & \int_I P_1\{d_{k1}^*(\overline{y}_1|\hat{B}, \hat{\sigma}^2) = 0|\overline{Y}_1 = \overline{y}_1\}[-c_1(\overline{y}_1)] g_1(\overline{y}_1) d\overline{y}_1 \\
& + \int_J P_1\{d_{k1}^*(\overline{y}_1|\hat{B}, \hat{\sigma}^2) = 1|\overline{Y}_1 = \overline{y}_1\} c_1(\overline{y}_1) g_1(\overline{y}_1) d\overline{y}_1 .
\end{aligned} \tag{4.1}$$

By the definitions of d_{k1}^*, ψ_1^* and by an application of Bonferroni inequality, for each $\overline{y}_1 \in I$,

$$
\begin{aligned}
&P_1\{d_{k1}^*(\overline{y}_1|\hat{B},\hat{\sigma}^2) = 0|\overline{Y}_1 = \overline{y}_1\}\\
=&P_1\{\psi_1^*(\overline{y}_1) > c|\overline{Y}_1 = \overline{y}_1\}\\
=&P_1\Big\{\Big[\frac{1}{2m}(1-\hat{B}) + \frac{(\overline{y}_1-\theta_0)^2}{2\hat{\sigma}^2}(1-\hat{B})^2\Big] - [\frac{1}{2m}(1-B) + \frac{(\overline{y}_1-\theta_0)^2}{2\sigma^2}(1-B)^2] > -c_1(\overline{y}_1)|\overline{Y}_1 = \overline{y}_1\}\\
\le&P_1\Big\{\frac{B-\hat{B}}{2m} > -\frac{c_1(\overline{y}_1)}{3}|\overline{Y}_1 = \overline{y}_1\Big\}\\
&+P_1\{\frac{(\overline{y}_1-\theta_0)^2}{2\sigma^2}[2-\hat{B}-B][B-\hat{B}] > -\frac{c_1(\overline{y}_1)}{3}|\overline{Y}_1 = \overline{y}_1\}\\
&+P_1\{\frac{(\overline{y}_1-\theta_0)^2(1-\hat{B})^2}{2}[\frac{1}{\hat{\sigma}^2} - \frac{1}{\sigma^2}] > -\frac{c_1(\overline{y}_1)}{3}|\overline{Y}_1 = \overline{y}_1\}\\
\le&P_1\{\hat{B}-B < \frac{2mc_1(\overline{y}_1)}{3}|\overline{Y}_1 = \overline{y}_1\}\\
&+P_1\{\hat{B}-B < \frac{2\sigma^2 c_1(\overline{y}_1)}{3(\overline{y}_1-\theta_0)^2}|\overline{Y}_1 = \overline{y}_1\}\\
&+P_1\{\hat{\sigma}^2-\sigma^2 < \frac{\sigma^4 c_1(\overline{y}_1)}{2\sigma^2[-c_1(\overline{y}_1)] + 3(\overline{y}_1-\theta_0)^2}|\overline{Y}_1 = \overline{y}_1\}
\end{aligned}
\tag{4.2}
$$

Similarly, for each $\overline{y}_1 \epsilon J$,

$$
\begin{aligned}
&P_1\{d_{k1}^*(\overline{y}_1|\hat{B},\hat{\sigma}^2) = 1|\overline{Y}_1 = \overline{y}_1\}\\
=&P_1\{\psi_1^*(\overline{y}_1) \le c|\overline{Y}_1 = \overline{y}_1\}\\
=&P_1\{\frac{B-\hat{B}}{2m} + \frac{(\overline{y}_1-\theta_0)^2}{2\hat{\sigma}^2}(1-\hat{B})^2 - \frac{(\overline{y}_1-\theta_0)^2}{2\sigma^2}(1-B)^2 < -c_1(\overline{y}_1)|\overline{Y}_1 = \overline{y}_1\}\\
\le&P_1\{\hat{B}-B > \frac{2mc_1(\overline{y}_1)}{3}|\overline{Y}_1 = \overline{y}_1\}\\
&+P_1\{\hat{B}-B > \frac{2\sigma^2 c_1(\overline{y}_1)}{3(\overline{y}_1-\theta_0)^2}|\overline{Y}_1 = \overline{y}_1\}\\
&+P_1\{\hat{\sigma}^2-\sigma^2 > \frac{2\sigma^4 c_1(\overline{y}_1)}{2\sigma^2[-c_1(\overline{y}_1)] + 3(\overline{y}_1-\theta_0)^2}|\overline{Y}_1 = \overline{y}_1\}.
\end{aligned}
\tag{4.3}
$$

Combining (4.1)-(4.3) together yields that

$$
0 \le R_{k1}(d_{k1}^*) - R_{k1}(d_{B1}) \le A_1 + A_2 + A_3 + B_1 + B_2 + B_3 \tag{4.4}
$$

where

$$A_1 = \int_I P_1\{\hat{B} - B < \frac{2mc_1(\overline{y}_1)}{3}|\overline{Y}_1 = \overline{y}_1\}[-c_1(\overline{y}_1)]g_1(\overline{y}_1)d\overline{y}_1,$$

$$A_2 = \int_I P_1\{\hat{B} - B < \frac{2\sigma^2 c_1(\overline{y}_1)}{3(\overline{y}_1 - \theta_0)^2}|\overline{Y}_1 = \overline{y}_1\}[-c_1(\overline{y}_1)]g_1(\overline{y}_1)d\overline{y}_1,$$

$$A_3 = \int_I P_1\{\hat{\sigma}^2 - \sigma^2 < \frac{\sigma^4 c_1(\overline{y}_1)}{2\sigma^2[-c_1(\overline{y}_1)] + 3(\overline{y}_1 - \theta_0)^2}|\overline{Y}_1 = \overline{y}_1\}[-c_1(\overline{y}_1)]g_1(\overline{y}_1)d\overline{y}_1,$$

$$B_1 = \int_J P_1\{\hat{B} - B > \frac{2mc_1(\overline{y}_1)}{3}|\overline{Y}_1 = \overline{y}_1\}c_1(\overline{y}_1)g_1(\overline{y}_1)d\overline{y}_1,$$

$$B_2 = \int_J P_1\{\hat{B} - B > \frac{2\sigma^2 c_1(\overline{y}_1)}{3(\overline{y}_1 - \theta_0)^2}|\overline{Y}_1 = \overline{y}_1\}c_1(\overline{y}_1)g_1(\overline{y}_1)d\overline{y}_1,$$

and

$$B_3 = \int_J P_1\{\hat{\sigma}^2 - \sigma^2 > \frac{2\sigma^4 c_1(\overline{y}_1)}{2\sigma^2[-c_1(\overline{y}_1)] + 3(\overline{y}_1 - \theta_0)^2}|\overline{Y}_1 = \overline{y}_1\}c_1(\overline{y}_1)g_1(\overline{y}_1)d\overline{y}_1.$$

Therefore, it suffices to investigate the asymptotic behavior for each of the six terms in (4.4). For this purpose, certain useful lemmas are introduced as follows.

Lemma 4.1 For a random variable $S \sim \chi^2_n$.

(a) $P\{\frac{S}{n} - 1 \leq C\} \leq \exp\{-\frac{n}{2}[C - ln(1 + C)]\}$ for $-1 < C < 0$;

(b) $P\{\frac{S}{n} - 1 \geq C\} \leq \exp\ \{-\frac{n}{2}[C - ln(1 + C)]\}$ for $C > 0$.

Note: Lemma 4.1 is from Corollary 4.1 of Gupta, Liang and Rau (1994).

For each real value b and y, define $\alpha_1(b) = \frac{(k-1)b}{4kB}$, $\alpha_2(b) = \frac{-b}{4(B+b)}$, $c_2(y, b) = \frac{(k-1)b}{kB} + \frac{m(B+b)(y-\theta_0)^2}{k\sigma^2} - \frac{1}{k}$, and $c_3(y, b) = \frac{(k-1)b}{kB} + \frac{m(B+b)(y-\theta_0)^2}{k\sigma^2}$. Also, let $W_1 = \sum_{j=2}^{k}(\overline{Y}_j - \theta_0)^2$. Note that $W_1/(\frac{\sigma^2}{m} + \tau^2) \sim \chi^2_{k-1}$. Finally, set $h(c) = c - ln(1 + c)$.

Lemma 4.2 For $\overline{y}_1 \in J$ and $b > 0$ such that $(k-1)b > 2B$, we have $P_1\{\hat{B} - B > b|\overline{Y}_1 = \overline{y}_1\} \leq \exp\{-\frac{k(m-1)}{2}\ h(\alpha_1(b))\} + \exp\{-\frac{k-1}{2}h(\alpha_2(b))\}$.

Proof: First note that $c_2(y, b) \geq \frac{(k-1)b}{kB} - \frac{1}{k} \geq \frac{(k-1)b}{2kB} = 2\alpha_1(b) > 0$, since $(k-1)b > 2B$.

Then, by the definition of $\hat{B}$ and the preceding inequality, we can obtain

$$
\begin{aligned}
&P_1\{\hat{B} - B > b|\overline{Y}_1 = \overline{y}_1\} \\
\leq &P_1\{\frac{S}{km(m-1)} - \frac{W(B+b)}{k} > 0|\overline{Y}_1 = \overline{y}_1\} \\
= &P_1\{\frac{S}{km(m-1)} - \frac{W_1(B+b)}{k} > \frac{(\overline{y}_1 - \theta_0)^2(B+b)}{k}|\overline{Y}_1 = \overline{y}_1\} \\
= &P_1\{\frac{S}{km(m-1)} - \frac{\sigma^2}{m}] - [W_1 - (k-1)(\frac{\sigma^2}{m} + \tau^2)]\frac{(B+b)}{k} > \frac{\sigma^2}{m}c_2(\overline{y}_1, b)|\overline{Y}_1 = \overline{y}_1\} \\
\leq &P_1\{\frac{S}{km(m-1)} - \frac{\sigma^2}{m} > \frac{\sigma^2}{2m}c_2(\overline{y}_1, b)|\overline{Y}_1 = \overline{y}_1\} \\
&+ P_1\{W_1 - (k-1)(\frac{\sigma^2}{m} + \tau^2) < -\frac{k\sigma^2}{2m(B+b)}c_2(\overline{y}_1, b)|\overline{Y}_1 = \overline{y}_1\} \\
= &P_1\{\frac{S}{\sigma^2 k(m-1)} - 1 > \frac{1}{2}c_2(\overline{y}_1, b)|\overline{Y}_1 = \overline{y}_1\} \\
&+ P_1\{\frac{W_1}{(\frac{\sigma^2}{m} + \tau^2)(k-1)} - 1 < -\frac{1}{2}c_2(\overline{y}_1, b)\frac{kB}{(k-1)(B+b)}|\overline{Y}_1 = \overline{y}_1\} \\
\leq &P_1\{\frac{S}{\sigma^2 k(m-1)} - 1 > \alpha_1(b)|\overline{Y}_1 = \overline{y}_1\} \\
&+ P_1\{\frac{W_1}{(\frac{\sigma^2}{m} + \tau^2)(k-1)} - 1 < \alpha_2(b)|\overline{Y}_1 = \overline{y}_1\} \\
\leq &\exp\{-\frac{k(m-1)}{2}[\alpha_1(b) - ln(1 + \alpha_1(b))]\} \\
&+ \exp\{-\frac{k-1}{2}[\alpha_2(b) - ln(1 + \alpha_2(b))]\} \\
= &\exp\{-\frac{k(m-1)}{2}h(\alpha_1(b))\} + \exp\{-\frac{k-1}{2}h(\alpha_2(b))\}.
\end{aligned}
\tag{4.5}
$$

In (4.5), the last inequality is obtained from an application of Lemma 4.1 by noting that $\alpha_1(b) > 0$ and $\alpha_2(b) < 0$ and S and W_1 are independent of $\overline{Y}_1$. □

Lemma 4.3 For each $\overline{y}_1 \in I$ and $b < 0$ such that $B + b > 0$ and $-b > 2mB^2(\overline{y}_1 - \theta_0)^2 \;/[(k-1)\sigma^2 + 2mB(\overline{y}_1 - \theta_0)^2]$, we have that $P_1\{\hat{B} - B < b|\overline{Y}_1 = \overline{y}_1\} \leq \exp\{-\frac{k(m-1)}{2}h(\alpha_1(b))\} + \exp\{-\frac{k-1}{2}h(\alpha_2(b))\}$.

Proof: Note that $c_2(\overline{y}_1, b) = c_3(\overline{y}_1, b) - \frac{1}{k} \le c_3(\overline{y}_1, b)$. Also, under the assumption of the Lemma, $c_3(\overline{y}_1, b) < 2\alpha_1(b) < 0$. Following an argument similar to the proof of Lemma 4.2, and by noting the preceding inequality, we obtain that

$$\begin{aligned}
& P_1\{\hat{B} - B < b | \overline{Y}_1 = \overline{y}_1\} \\
\le & P_1\{\frac{S}{\sigma^2 k(m-1)} - 1 < \frac{1}{2} c_2(\overline{y}_1, b) | \overline{Y}_1 = \overline{y}_1\} \\
& + P_1\{\frac{W_1}{(\frac{\sigma^2}{m} + \tau^2)(k-1)} > -\frac{1}{2} c_2(\overline{y}_1, b) \frac{kB}{(k-1)(B+b)} | \overline{Y}_1 = \overline{y}_1\} \\
\le & P_1\{\frac{S}{\sigma^2 k(m-1)} - 1 < \alpha_1(b) | \overline{Y}_1 = \overline{y}_1\} \\
& + P_1\{\frac{W_1}{(\frac{\sigma^2}{m} + \tau^2)(k-1)} - 1 > \alpha_2(b) | \overline{Y}_1 = \overline{y}_1\} \\
\le & \exp\{-\frac{k(m-1)}{2} [\alpha_1(b) - ln(1 + \alpha_1(b))]\} \\
& + \exp\{-\frac{k-1}{2} [\alpha_2(b) - ln(1 + \alpha_2(b))]\} \\
= & \exp\{-\frac{k(m-1)}{2} h(\alpha_1(b))\} + \exp\{-\frac{k-1}{2} h(\alpha_2(b))\}. \qquad \square
\end{aligned}$$

Lemma 4.4 For fixed $t_1 > 0$, and $n > 0$,

$$\int_0^{t_1} x \exp\{-\frac{n}{2}[x - ln(1+x)]\} dx = O(n^{-1}).$$

(b) For $0 < t_0 < 1$ and $n > 0$.

$$\int_0^{t_0} x \exp\{\frac{n}{2}[x + ln(1-x)]\} dx = O(n^{-1}).$$

Proof: These results can be obtained through straight forward computation. The details are omitted here. $\square$

Now we are going to investigate the asymptotic behavior of A_i and $B_i, i = 1, 2, 3$ for k being sufficiently large.

Lemma 4.5

$$\int_J P_1\{\hat{B} - B > \frac{2mc_1(\overline{y}_1)}{3} | \overline{Y}_1 = \overline{y}_1\} c_1(\overline{y}_1) g_1(\overline{y}_1) d\overline{y}_1 = O(k^{-1}).$$

Proof: Assume k being sufficiently large so that $2\sigma^2(\frac{1-B}{m}+c)/(1-B)^2 > \frac{2B}{k-1}$. Note that by the definition of $\hat{B}, P_1\{\hat{B}-B > \frac{2mc_1(\overline{y}_1)}{3} \,|\overline{Y}_1 = \overline{y}_1\} = 0$ if $\frac{2mc_1(\overline{y}_1)}{3} > 1-B$, which is equivalent to that $(\overline{y}_1-\theta_0)^2 \geq 2\sigma^2(\frac{1-B}{m}+c)/(1-B)^2$.

Let

$$J_1 =\{\overline{y}_1 \,\epsilon\, J|0 < \frac{2mc_1(\overline{y}_1)}{3} \leq \frac{2B}{k-1}\},$$
$$J_2 =\{\overline{y}_1 \,\epsilon\, J|\frac{2B}{k-1} < \frac{2mc_1(\overline{y}_1)}{3} < 1-B\}.$$

Then,

$$\begin{aligned} B_1 =& \int_{J_1} P_1\{\hat{B}-B > \frac{2mc_1(\overline{y}_1)}{3}|\overline{Y}_1 = \overline{y}_1\}c_1(\overline{y}_1)g_1(\overline{y}_1)d\overline{y}_1 \\ &+ \int_{J_2} P_1\{\hat{B}-B > \frac{2mc_1(\overline{y}_1)}{3}|\overline{Y}_1 = \overline{y}_1\}c_1(\overline{y}_1)g_1(\overline{y}_1)d\overline{y}_1 \\ \equiv& B_{11}+B_{12}, \end{aligned} \tag{4.6}$$

where

$$B_{11} \leq \int_{J_1} \frac{3B}{m(k-1)}g_1(\overline{y}_1)d\overline{y}_1 \leq \frac{3B}{m(k-1)} = O(k^{-1}); \tag{4.7}$$

and by Lemma 4.2,

$$\begin{aligned} B_{12} \leq& \int_{J_2} \exp\{-\frac{k(m-1)}{2}h(\alpha_1(\frac{2mc_1(\overline{y}_1)}{3}))\}c_1(\overline{y}_1)g_1(\overline{y}_1)d\overline{y}_1 \\ &+ \int_{J_2} \exp\{-\frac{k-1}{2}h(\alpha_2(\frac{2mc_1(\overline{y}_1)}{3}))\}c_1(\overline{y}_1)g_1(\overline{y}_1)d\overline{y}_1 \\ \equiv& B_{121}+B_{122}. \end{aligned} \tag{4.8}$$

For $\overline{y}_1 \,\epsilon\, J_2, \frac{g_1(\overline{y}_1)}{2|\overline{y}_1-\theta_0|} \leq [8c^*\pi(\frac{\sigma^2}{m}+\tau^2)]^{-\frac{1}{2}}\exp\{-\frac{c^*}{2(\frac{\sigma^2}{m}+\tau^2)}\} \equiv M^*$.

Let $M_1^* = M^* \times \frac{2\sigma^2}{(1-B)^2}$. Hence,

$$\begin{aligned} B_{121} =& \int_{J_2} \exp\{-\frac{k(m-1)}{2}h(\alpha_1(\frac{2mc_1(\overline{y}_1)}{3}))\}c_1(\overline{y}_1)\frac{g_1(\overline{y}_1)}{2|\overline{y}_1-\theta_0|} \times \frac{2\sigma^2}{(1-B)^2}dc_1(\overline{y}_1) \\ \leq& \int_{J_2} M_1^* c_1(\overline{y}_1)\exp\{-\frac{k(m-1)}{2}h(\alpha_1(\frac{2mc_1(\overline{y}_1)}{3}))\}dc_1(\overline{y}_1) \\ =& \int_{\frac{B}{m(k-1)}}^{1-B} \left(\frac{3}{2m}\right)^2 M_1^* z \exp(-\frac{k(m-1)}{2}h(\alpha_1(z)))dz \\ =& M^*\left[\frac{6kB}{m(k-1)}\right]^2 \int_{\frac{1}{4km}}^{\frac{(k-1)(1-B)}{4kB}} \alpha \exp\{-\frac{k(m-1)}{2}[\alpha - ln(1+\alpha)]\}d\alpha \\ =& O(k^{-1}) \text{ by Lemma 4.4(a).} \end{aligned} \tag{4.9}$$

Also,

$$
\begin{aligned}
B_{122} \le & M^*(\frac{3}{2m})^2 \int_{\frac{B}{m(k-1)}}^{1-B} \exp\{-\frac{k-1}{2} h(\alpha_2(z))\} z dz \\
= & M^*(\frac{6B}{m})^2 \int_{\frac{1}{4[m(k-1)+1]}}^{\frac{1-B}{4}} \frac{\alpha}{(1-4\alpha)^2} \exp\{\frac{k-1}{2}[\alpha + ln(1-\alpha)]\} d\alpha \\
\le & M^*(\frac{6B}{m})^2 \int_{\frac{1}{4[m(k-1)+1]}}^{\frac{1-B}{4}} \frac{\alpha}{B^2} \exp\{\frac{k-1}{2}[\alpha + ln(1-\alpha)]\} d\alpha \\
= & O(k^{-1}) \text{ by Lemma 4.4(b) .}
\end{aligned} \tag{4.10}
$$

Now, combining (4.6)-(4.10) together concludes the result of the lemma. □

Lemma 4.6

$$
\int_J P_1\{\hat{B} - B > \frac{2\sigma^2 c_1(\overline{y}_1)}{3(\overline{y}_1 - \theta_0)^2} | \overline{Y}_1 = \overline{y}_1\} c_1(\overline{y}_1) g_1(\overline{y}_1) d\overline{y}_1 = O(k^{-1}).
$$

Proof: For $\overline{y}_1 \, \epsilon \, J$, $0 < \frac{2\sigma^2 c_1(\overline{y}_1)}{3(\overline{y}_1-\theta_0)^2} = \frac{(1-B)^2}{9} + \frac{2\sigma^2}{9} \times \frac{\frac{1}{2m}(1-B)-c}{(\overline{y}_1-\theta_0)^2}$, which is increasing in $|\overline{y}_1 - \theta_0|$ and bounded above by $\frac{(1-B)^2}{9}$ since $\frac{1}{2m}(1-B) - c < 0$ by the assumption that $c > \frac{1}{2m}$ (see the end of Section 2). Assume k being sufficiently large so that $\frac{2\sigma^2 c_1(\theta_0 + 2\sqrt{c^*})}{12c^*} > \frac{2B}{k-1}$. Let

$$
\begin{aligned}
J_1^* = & \{\overline{y}_1 \epsilon J \,\Big|\, 0 < \frac{2\sigma^2 c_1(\overline{y}_1)}{3(\overline{y}_1 - \theta_0)^2} \le \frac{2B}{k-1}\}, \\
J_2^* = & \{\overline{y}_1 \epsilon J \,\Big|\, |\overline{y}_1 - \theta_0| < 2\sqrt{c^*} \text{and} \frac{2\sigma^2 c_1(\overline{y}_1)}{3(\overline{y}_1 - \theta_0)^2} > \frac{2B}{k-1}\}, \\
J_3^* = & \{\overline{y}_1 \epsilon J \,\Big|\, |\overline{y}_1 - \theta_0| \ge 2\sqrt{c^*}\}.
\end{aligned}
$$

Note that $\overline{y}_1 \epsilon J_3^*$ iff $\frac{2\sigma^2 c_1(\overline{y}_1)}{3(\overline{y}_1-\theta_0)^2} \ge \frac{2\sigma^2 c_1(\theta_0+2\sqrt{c^*})}{12c^*} \equiv \beta^*$. Therefore,

$$
\begin{aligned}
B_2 = & \int_{J_1^*} P_1\{\hat{B} - B > \frac{2\sigma^2 c_1(\overline{y}_1)}{3(\overline{y}_1 - \theta_0)^2} | \overline{Y}_1 = \overline{y}_1\} c_1(\overline{y}_1) g_1(\overline{y}_1) d\overline{y}_1 \\
& + \int_{J_2^*} P_1\{\hat{B} - B > \frac{2\sigma^2 c_1(\overline{y}_1)}{3(\overline{y}_1 - \theta_0)^2} | \overline{Y}_1 = \overline{y}_1\} c_1(\overline{y}_1) g_1(\overline{y}_1) d\overline{y}_1 \\
& + \int_{J_3^*} P_1\{\hat{B} - B > \frac{2\sigma^2 c_1(\overline{y}_1)}{3(\overline{y}_1 - \theta_0)^2} | \overline{Y}_1 = \overline{y}_1\} c_1(\overline{y}_1) g_1(\overline{y}_1) d\overline{y}_1 \\
& \equiv B_{21} + B_{22} + B_{23},
\end{aligned} \tag{4.11}
$$

where,

$$
\begin{aligned}
B_{21} \leq& \int_{J_1^*} \frac{3B}{(k-1)\sigma^2}(\overline{y}_1-\theta_0)^2 g_1(\overline{y}_1)d\overline{y}_1 \\
\leq& \frac{3B}{(k-1)\sigma^2}(\frac{\sigma^2}{m}+\tau^2) \\
=& O(k^{-1});
\end{aligned} \tag{4.12}
$$

$$
\begin{aligned}
B_{23} \leq& \int_{J_3^*} P_1\{\hat{B}-B>\beta^*|\overline{Y}_1=\overline{y}_1\}c_1(\overline{y}_1)g_1(\overline{y}_1)d\overline{y}_1 \\
\leq& \int_{J_3^*}\left[\exp\{-\frac{k(m-1)}{2}h(\alpha_1(\beta^*))\}+\exp\{-\frac{k-1}{2}h(\alpha_2(\beta^*))\}\right]\frac{(1-B)^2(\overline{y}_1-\theta_0)^2}{6\sigma^2}g_1(\overline{y}_1)d\overline{y}_1 \\
\leq& \left[\exp\{-\frac{k(m-1)}{2}h(\alpha_1(\beta^*))\}+\exp\{-\frac{k-1}{2}h(\alpha_2(\beta^*))\}\right]\frac{(1-B)^2(\frac{\sigma^2}{m}+\tau^2)}{6\sigma^2} \\
=& O(k^{-1})
\end{aligned} \tag{4.13}
$$

and by Lemma 4.2,

$$
\begin{aligned}
B_{22} \leq& \int_{J_2^*}\exp\{-\frac{k(m-1)}{2}h(\alpha_1(\frac{2\sigma^2 c_1(\overline{y}_1)}{3(\overline{y}_1-\theta_0)^2}))\}c_1(\overline{y}_1)g_1(\overline{y}_1)d\overline{y}_1 \\
&+\int_{J_2^*}\exp\{-\frac{k-1}{2}h(\alpha_2(\frac{2\sigma^2 c_1(\overline{y}_1)}{3(\overline{y}_1-\theta_0)^2}))\}c_1(\overline{y}_1)g_1(\overline{y}_1)d\overline{y}_1 \\
\leq& \int_{J_2^*}\exp\{-\frac{k(m-1)}{2}h(\alpha_1(\frac{2\sigma^2 c_1(\overline{y}_1)}{12c^*}))\}c_1(\overline{y}_1)g_1(\overline{y}_1)d\overline{y}_1 \\
&+\int_{J_2^*}\exp\{-\frac{k-1}{2}h(\alpha_2(\frac{2\sigma^2 c_1(\overline{y}_1)}{3c^*}))\}c_1(\overline{y}_1)g_1(\overline{y}_1)d\overline{y}_1 \\
=& O(k^{-1}).
\end{aligned} \tag{4.14}
$$

In (4.14), the second inequality is obtained by the fact that for $\overline{y}_1 \epsilon J_2^*, h(\alpha_1(\frac{2\sigma^2 c_1(\overline{y}_1)}{3(\overline{y}_1-\theta_0)^2})) \geq h(\alpha_1(\frac{2\sigma^2 c_1(\overline{y}_1)}{3(2\sqrt{c^*})^2}))$ and $h(\alpha_2(\frac{2\sigma^2 c_1(\overline{y}_1)}{3(\overline{y}_1-\theta_0)^2})) \geq h\ (\alpha_2(\frac{2\sigma^2 c_1(\overline{y}_1)}{3(\sqrt{c^*})^2}))$. Also, the last equality is obtained by an argument similar to that for B_{12} by noting that J_2^* is a bounded set. Combining (4.11)-(4.14) together leads to the result of the lemma. □

Lemma 4.7

$$
\int_J P_1\{\hat{\sigma}^2-\sigma^2>\frac{2\sigma^4 c_1(\overline{y}_1)}{2\sigma^2[-c_1(\overline{y}_1)]+3(\overline{y}_1-\theta_0)^2}|\overline{Y}_1=\overline{y}_1\}c_1(\overline{y}_1)g_1(\overline{y}_1)d\overline{y}_1 = O(k^{-1}).
$$

Proof: For $\overline{y}_1 \epsilon J$, $c_1(\overline{y}_1)>0$ and $0<2\sigma^2[-c_1(\overline{y}_1)]+3(\overline{y}_1-\theta_0)^2<3(\overline{y}_1-\theta_0)^2$.

Hence, $\frac{2\sigma^4 c_1(\overline{y}_1)}{2\sigma^2[-c_1(\overline{y}_1)]+3(\overline{y}_1-\theta_0)^2} > \frac{2\sigma^4 c_1(\overline{y}_1)}{3(\overline{y}_1-\theta_0)^2}$. Therefore,

$$\begin{aligned} B_3 \leq & \int_J P_1\{\hat{\sigma}^2 - \sigma^2 > \frac{2\sigma^4 c_1(\overline{y}_1)}{3(\overline{y}_1-\theta_0)^2} | \overline{Y}_1 = \overline{y}_1\} c_1(\overline{y}_1) g_1(\overline{y}_1) d\overline{y}_1 \\ = & \int_J P_1\{\frac{\hat{\sigma}^2}{\sigma^2} - 1 > \frac{2\sigma^2 c_1(\overline{y}_1)}{3(\overline{y}_1-\theta_0)^2} | \overline{Y}_1 = \overline{y}_1\} c_1(\overline{y}_1) g_1(\overline{y}_1) d\overline{y}_1 \\ \equiv & B_3^*, \end{aligned}$$

which is a form similar to that of B_2. Therefore, the technique used to treat B_2 can be applied here and one can conclude that $B_3^* = O(k^{-1})$. □

Lemma 4.8

$$\int_I P_1\{\hat{B} - B < \frac{2mc_1(\overline{y}_1)}{3} | \overline{Y}_1 = \overline{y}_1\}[-c_1(\overline{y}_1)] g_1(\overline{y}_1) d\overline{y}_1 = O(k^{-1}).$$

Proof: Note that $\ell(\overline{y}_1 - \theta_0) = \frac{2mB^2(\overline{y}_1-\theta_0)^2}{(k-1)\sigma^2+2m(\overline{y}_1-\theta_0)^2}$ is increasing in $|\overline{y}_1 - \theta_0|$, and $0 \leq \ell(\overline{y}_1 - \theta_0) \leq \frac{2mB^2c^*}{(k-1)\sigma^2+2mc^*}$ for all $\overline{y}_1 \in I$. Thus, for k being sufficiently large, $-b^* \equiv -\frac{2m}{3}c_1(\theta_0 + \frac{\sqrt{c^*}}{2}) > \ell(\overline{y}_1 - \theta_0)$ for all $\overline{y}_1 \in I$. Let

$$\begin{aligned} I_1 = & [\theta_0 - \frac{\sqrt{c^*}}{2}, \theta_0 + \frac{\sqrt{c^*}}{2}], \\ I_2 = & \{\overline{y}_1 \in I - I_1 | 0 \leq \frac{-2mc_1(\overline{y}_1)}{3} \leq \ell(\overline{y}_1 - \theta_0)\}, \\ \text{and } I_3 = & \{\overline{y}_1 \in I - I_1 | \frac{-2mc_1(\overline{y}_1)}{3} > \ell(\overline{y}_1 - \theta_0)\}. \end{aligned}$$

Then,

$$\begin{aligned} A_1 = & \int_{I_1} P_1\{\hat{B} - B < \frac{2mc_1(\overline{y}_1)}{3} | \overline{Y}_1 = \overline{y}_1\}[-c_1(\overline{y}_1)] g_1(\overline{y}_1) d\overline{y}_1 \\ & + \int_{I_2} P_1\{\hat{B} - B < \frac{2mc_1(\overline{y}_1)}{3} | \overline{Y}_1 = \overline{y}_1\}[-c_1(\overline{y}_1)] g_1(\overline{y}_1) d\overline{y}_1 \\ & + \int_{I_3} P_1\{\hat{B} - B < \frac{2mc_1(\overline{y}_1)}{3} | \overline{Y}_1 = \overline{y}_1\}[-c_1(\overline{y}_1)] g_1(\overline{y}_1) d\overline{y}_1 \\ & \equiv A_{11} + A_{12} + A_{13}. \end{aligned} \tag{4.15}$$

For $\overline{y}_1 \in I_1$, $\frac{2mc_1(\overline{y}_1)}{3} \leq \frac{2mc_1(\theta_0+\frac{\sqrt{c^*}}{2})}{3} = b^* < 0$. Also, $|c_1(\overline{y}_1)| \leq c$ for all $\overline{y}_1 \in I$.

Therefore, by Lemma 4.3,

$$\begin{aligned}
A_{11} \leq & \int_{I_1} P_1\{\hat{B} - B < b^* | \overline{Y}_1 = \overline{y}_1\} c g_1(\overline{y}_1) d\overline{y}_1 \\
\leq & \int_{I_1} c[\exp\{-\frac{k(m-1)}{2} h(\alpha_1(b^*))\} + \exp\{-\frac{k-1}{2} h(\alpha_2(b^*))\}] g_1(\overline{y}_1) d\overline{y}_1 \\
\leq & c \; [\; \exp \; \{-\frac{k(m-1)}{2} h(\alpha_1(b^*))\} + \exp\{\frac{k-1}{2} h(\alpha_1(b^*))\}] \\
= & O(k^{-1}).
\end{aligned} \tag{4.16}$$

Also,

$$\begin{aligned}
A_{12} \leq & \int_{I_2} \frac{3\ell(\overline{y}_1 - \theta_0)}{2m} g_1(\overline{y}_1) d\overline{y}_1 \\
\leq & \int_{I_2} \frac{2mB^2 c^*}{(k-1)\sigma^2 + 2mc^*} g_1(\overline{y}_1) d\overline{y}_1 \\
= & O(k^{-1}).
\end{aligned} \tag{4.17}$$

By Lemma 4.3 and by a proof similar to that of (4.9) and (4.10), we have

$$\begin{aligned}
A_{13} \leq & \int_{I_3} \exp\{-\frac{k(m-1)}{2} h(\alpha_1(\frac{2mc_1(\overline{y}_1)}{3}))\}[-c_1(\overline{y}_1)] g_1(\overline{y}_1) d\overline{y}_1 \\
& + \int_{I_3} \exp\{-\frac{k-1}{2} h(\alpha_2(\frac{2mc_1(\overline{y}_1)}{3}))\}[-c_1(\overline{y}_1)] g_1(\overline{y}_1) d\overline{y}_1 \\
= & O(k^{-1}).
\end{aligned} \tag{4.18}$$

Combining (4.15)-(4.18) together leads to the conclusion of the lemma. □.

Lemma 4.9

$$\int_I P_1\{\hat{B} - B < \frac{2\sigma^2 c_1(\overline{y}_1)}{3(\overline{y}_1 - \theta_0)^2} | \overline{Y}_1 = \overline{y}_1\}[-c_1(\overline{y}_1)] g_1(\overline{y}_1) d\overline{y}_1 = O(k^{-1}).$$

Proof: For $\overline{y}_1 \, \epsilon \, I$, $\frac{2\sigma^2 c_1(\overline{y}_1)}{3(\overline{y}_1 - \theta_0)^2} \leq \frac{2\sigma^2 c_1(\overline{y}_1)}{3c^*} < 0$. Therefore,

$$\begin{aligned}
A_2 \leq & \int_I P_1\{\hat{B} - B < \frac{2\sigma^2 c_1(\overline{y}_1)}{3c^*} | \overline{Y}_1 = \overline{y}_1\}[-c_1(\overline{y}_1)] g_1(\overline{y}_1) d\overline{y}_1 \\
\equiv & A_2^*
\end{aligned}$$

which has a form similar to that of A_1. Hence we conclude that $A_2 = O(k^{-1})$ by Lemma 4.8. □.

Lemma 4.10

$$\int_I P_1\{\hat{\sigma}^2 - \sigma^2 < \frac{\sigma^4 c_1(\overline{y}_1)}{2\sigma^2[-c_1(\overline{y}_1)] + 3(\overline{y}_1 - \theta_0)^2} | \overline{Y}_1 = \overline{y}_1\}[-c_1(\overline{y}_1)] g_1(\overline{y}_1) d\overline{y}_1 = O(k^{-1}).$$

Proof: For $\overline{y}_1 \in I$, $\frac{\sigma^4 c_1(\overline{y}_1)}{2\sigma^2[-c_1(\overline{y}_1)]+3(\overline{y}_1-\theta_0)^2} \leq \frac{\sigma^4 c_1(\overline{y}_1)}{2\sigma^2 c+3c^*} < 0$. Hence, similar to that of Lemma 4.8, we can conclude that

$$\begin{aligned}A_3 \leq &\int_I P\{\hat{\sigma}^2 - \sigma^2 < \frac{\sigma^4 c_1(\overline{y}_1)}{2c\sigma^2 + 3c^*} | \overline{Y}_1 = \overline{y}_1\}[-c_1(\overline{y}_1)]g_1(\overline{y}_1)d\overline{y}_1 \\ =&O(k^{-1}).\end{aligned}$$

□

We summarize the preceding results as a theorem as follows.

Theorem 4.1 Under the statistical model described in Section 2, the empirical Bayes selection rule $\underset{\sim}{d}_k^*$ is asymptotically optimal with $\rho_k(\underset{\sim}{d}_k^*) = O(k^{-1})$.

5. Small Sample Performance: Simulation Study

We carried out a simulation study to investigate the small sample performance of the empirical Bayes selection rule $\underset{\sim}{d}_k^*$. Note that

$$\rho_k(\underset{\sim}{d}_k^*) = \frac{R_k(\underset{\sim}{d}_k^*) - R_k(\underset{\sim}{d}_B)}{kR_{k1}(d_{B1})},$$

where $R_{k1}(d_{B1})$ is a fixed value, independent of the value k. Let

$$B_k(\underset{\sim}{Y}) = \sum_{i=1}^{k} [d_{ki}^*(\overline{Y}_i | \hat{B}, \hat{\sigma}^2) - d_{Bi}(\overline{Y}_i)][\psi_i(\overline{Y}_i) - c].$$

Then,

$$\begin{aligned}EB_k(\underset{\sim}{Y}) =& \sum_{i=1}^{k} E\{[d_{ki}^*(\overline{Y}_i | \hat{B}, \hat{\sigma}^2) - d_{Bi}(\overline{Y}_i)][\psi_i(\overline{Y}_i) - c]\} \\ =& \sum_{i=1}^{k} E_{(i)}\{E_i\{[d_{ki}^*(\overline{Y}_i | \hat{B}, \hat{\sigma}^2) - d_{Bi}(\overline{Y}_i)][\psi_i(\overline{Y}_i) - c]\}\} \\ =& R_k(\underset{\sim}{d}_k^*) - R_k(\underset{\sim}{d}_B),\end{aligned}$$

where the expectation $E_{(i)}$ is taken with respect to the probability measure generated by $\overline{Y}_i$, and E_i is the expectation taken with respect to the conditional probability measure generated by $\hat{B}$ and $\hat{\sigma}^2$ given $\overline{Y}_i$. Hence,

$$\rho_k(\underset{\sim}{d}_k^*) = E[\frac{B_k(\underset{\sim}{Y})}{kR_{k1}(d_{B1})}] = \frac{1}{R_{k1}(d_{B1})} E[\frac{B_k(\underset{\sim}{Y})}{k}].$$

Since $R_{k1}(d_{B1})$ is independent of the value k, the relative regret Bayes risk $\rho_k(\underset{\sim}{d}_k^*)$ depends on k only through the part $E\left[\frac{B_k(\underset{\sim}{Y})}{k}\right]$.

By the law of large numbers, the sample mean $\overline{B}_k = \frac{1}{n}\sum_{\ell=1}^{n} B_k(\underset{\sim}{Y}(\ell))$ can be used as an estimator of the regret Bayes risk $R_k(\underset{\sim}{d}_k^*) - R_k(\underset{\sim}{d}_B)$, where $\underset{\sim}{Y}(\ell), \ell = 1, 2, \ldots, n$, are iid random vectors, identically distributed with $\underset{\sim}{Y}$. Therefore, we use $\overline{B}_k/k = \frac{1}{n}\sum_{\ell=1}^{n} B_k(\underset{\sim}{Y}(\ell))/k$ to estimate the relationship between $\rho_k(\underset{\sim}{d}_k^*)$ and k.

The simulation scheme used in this paper is described as follows.

(1) For each $i = 1, \ldots, k$, generate the independent random vector $\underset{\sim}{Y}_i = (Y_{i1}, \ldots, Y_{im})$ by the following:

  (a) Generate Θ_i from a $N(\theta_0, \tau^2)$ distribution.

  (b) Given $\Theta_i = \theta_i$, generate random sample $Y_{i1} \ldots Y_{im}$ from a $N(\theta_i, \sigma^2)$ distribution.

2. Based on the data $\underset{\sim}{Y} = (\underset{\sim}{Y}_1, \ldots, \underset{\sim}{Y}_k)$, construct the Bayes and empirical Bayes selection rules $\underset{\sim}{d}_B$ and $\underset{\sim}{d}_k^*$, respectively, and compute the $B_k(\underset{\sim}{Y})$ value.

3. For each k, steps (1) and (2) were repeated 1000 times. The average $\overline{B}_k$ of $B_k(\underset{\sim}{Y}(\ell)), \ell = 1, \ldots, 1000$, based on the 1000 repetitions is used as an estimator of the regret Bayes risk $R_k(\underset{\sim}{d}_k^*) - R_k(\underset{\sim}{d}_B)$ and $\overline{B}_k/k$ as an estimator of $R_{k1}(d_{B1})\rho_k(\underset{\sim}{d}_k^*) = [R_k(\underset{\sim}{d}_k^*) - R_k(\underset{\sim}{d}_B)]/k$. Also, $SE(\overline{B}_k/k)$, the estimated standard error of $\overline{B}_k/k$, is computed.

Table 1 lists a simulation result on the performance of the proposed empirical Bayes selection rule $\underset{\sim}{d}_k^*$ for the case where $m = 10, \sigma^2 = 2, \tau^2 = 1.5, \theta_0 = 0$ and $c = 0.3$.

From Table 1, we learn that the values of $\overline{B}_k/k$ decrease quite rapidly as k increases. Note that for $k \geq 40$, the estimated regret Bayes risk values $\overline{B}_k$ oscilate about the value 0.0170, which indicates that $\overline{B}_k/k$ converges to 0 with a rate of convergence of order $O(k^{-1})$, same as the conclusion in Theorem 4.1.

Table 1. Small Sample Performance of $\underset{\sim}{d}_k^*$ for $m = 10$, $\sigma^2 = 2$, $\tau^2 = 1.5$, $\theta_0 = 0$ and $c = 0.3$

k	$\overline{B}_k$	$\overline{B}_k/k$	$SE(\overline{B}_k/k)$
10	0.0238	0.00238	0.00538
20	0.0242	0.00121	0.00290
30	0.0236	0.00079	0.00186
40	0.0171	0.00043	0.00096
50	0.0159	0.00032	0.00060
60	0.0194	0.00032	0.00062
70	0.0186	0.00027	0.00057
80	0.0185	0.00023	0.00043
90	0.0187	0.00021	0.00037
100	0.0165	0.00017	0.00034
110	0.0171	0.00016	0.00034
120	0.0196	0.00016	0.00027
130	0.0168	0.00013	0.00023
140	0.0201	0.00015	0.00028
150	0.0194	0.00013	0.00022
160	0.0168	0.00010	0.00019
170	0.0164	0.00010	0.00016
180	0.0159	0.00009	0.00015
190	0.0175	0.00009	0.00016
200	0.0177	0.00009	0.00015

References

Bechhofer, R. E. and Turnbull, B. W. (1978). Two $(k+1)$-decision selection procedures for comparing k normal means with a specified standard. *J. Amer. Statist. Assoc.*, **73**, 385–392.

Burr, I. W. (1976). *Statistical Quality Control Methods.* Marcel Dekker, Inc., New York.

Dunnett, Charles, W. (1984). Selection of the best treatment in comparison to a control with an application to a medical trial. A *Design of Experiments: Ranking and Selection*, (Eds. A. C. Tamhane and T. Santner), Marcel Dekker, 47–66.

Giani, G. and Straßburger, K. (1994). Testing and selecting for equivalence with respect to a control. *J. Amer. Statist. Assoc.*, **89**, 320–329.

Gupta, S. S. and Hsiao, P. (1981). On Γ-minimax, minimax, and Bayes procedures for selecting populations close to a control. *Sankhya*, **B, 43**, 291–318.

Gupta, S. S., Liang, T. and Rau, R. B. (1994). Empirical Bayes rules for selecting the best normal population compared with a control. *Statistics & Decisions*, **12**, 125–147.

Gupta, S. S. and Singh, A. K. (1979). On selection rules for treatments versus control problems. *Proceedings of the 42nd Session of the International Statistical Institute.* 229–232.

Gupta, S. S. and Sobel, M. (1958). On selecting a subset which contains all populations better than a control. *Ann. Math. Statist.*, **29**, 235–244.

Huang, W. T. (1975). Bayes approach to a problem of partitioning k normal populations. *Bull. Inst. Math. Acad. Sinica*, **3**, 87–97.

Mee, R. W., Shah, A. K. and Lefante, J. J. (1987). Comparing k independent sample means with a known standard. *J. Quality Technology*, **19**, 75–81.

Miescke, K. J. (1981). Γ-minimax selection procedures in simultaneous testing problems. *Ann. Statist.*, **9**, 215–220.

Paulson, E. (1952). On the comparison of several experimental categories with a control. *Ann. Math. Statist.*, **23**, 239–246.

Randles, R. H. and Hollander, M. (1971). Γ-minimax selection procedures in treatments versus control problems. *Ann. Math. Statist.*, **42**, 330–341.

Romano, A. (1977). *Applied Statistics For Science and Industry.* Allyn and Bacon, Boston, MA.

Wellek, S. and Michaelis, J. (1991). Element of significance testing with equivalence problem. *Methods of Information in Medicine*, **30**, 194–198.

Wilcox, R. R. (1984). Selecting the best population, provided it is better than a standard: the unequal variance case. *J. Amer. Statist. Assoc.*, **79**, 887–891.

SF 298 MASTER COPY KEEP THIS COPY FOR REPRODUCTION PURPOSES

REPORT DOCUMENTATION PAGE

Form Approved
OMB NO. 0704-0188

1. AGENCY USE ONLY (Leave blank)

2. REPORT DATE
June, 1995

3. REPORT TYPE AND DATES COVERED
Final Technical Report, June 1995

4. TITLE AND SUBTITLE
Simultaneousley Selecting Normal Populations Close to A Control

5. FUNDING NUMBERS
DAAH04-95-01-0165

6. AUTHOR(S)
TaChen Liang

7. PERFORMING ORGANIZATION NAMES(S) AND ADDRESS(ES)
Purdue University
West Lafayette IN 47907

8. PERFORMING ORGANIZATION REPORT NUMBER
Technical Report #95-26C

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)
U.S. Army Research Office
P.O. Box 12211
Research Triangle Park, NC 27709-2211

10. SPONSORING / MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES
The views, opinions and/or findings contained in this report are those of the author(s) and should not be construed as an official Department of the Army position, policy or decision, unless so designated by other documentation.

12a. DISTRIBUTION / AVAILABILITY STATEMENT
Approved for public release; distribution unlimited.

12 b. DISTRIBUTION CODE

13. ABSTRACT (Maximum 200 words)

We study the problem of selecting populations close to a control from among k normal populations using the parametric empirical Bayes approach. A Bayes selection rule is derived, which depends on certain parameters. When those parameters are unknown, using the empirical Bayes idea, we first present estimators, based on information collected from the k populations for the unknown parameters. Then, mimicking the behavior of the Bayes selection rule, an empirical Bayes selection rule is constructed. The relative regret Bayes risk is used as a measure of performance of the empirical Bayes selection rule. It is shown that the relative regret Bayes risk of the proposed empirical Bayes selection rule converges to zero at a rate of order $O(k^{-1})$. A simulation study is also carried out to investigate the performance of the proposed empirical Bayes selection rule for small to moderate values of k.

14. SUBJECT TERMS
Asymptotically optimal, empirical Bayes, rate of convergence, relative regret Bayes risk, simultanious selection

15. NUMBER IF PAGES
22

16. PRICE CODE

17. SECURITY CLASSIFICATION OR REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
UNCLASSIFIED	UNCLASSIFIED	UNCLASSIFIED	UL

NSN 7540-01-280-5500

Standard Form 298 (Rev. 2-89)
Prescribed by ANSI Std. 239-18
298-102

GENERAL INSTRUCTIONS FOR COMPLETING SF 298

The Report Documentation Page (RDP) is used in announcing and cataloging reports. It is important that this information be consistent with the rest of the report, particularly the cover and title page. Instructions for filling in each block of the form follow. It is important to ***stay within the lines*** to meet ***optical scanning requirements.***

Block 1. Agency Use Only *(Leave blank)*

Block 2. Report Date. Full publication date including day, month, and year, if available (e.g. 1 Jan 88). Must cite at least year.

Block 3. Type of Report and Dates Covered. State whether report is interim, final, etc. If applicable, enter inclusive report dates (e.g. 10 Jun 87 - 30 Jun 88).

Block 4. Title and Subtitle. A title is taken from the part of the report that provides the most meaningful and complete information. When a report is prepared in more than one volume, repeat the primary title, add volume number, and include subtitle for the specific volume. On classified documents enter the title classification in parentheses.

Block 5. Funding Numbers. To include contract and grant numbers; may include program element number(s), project number(s), task number(s), and work unit number(s). Use the following labels:

C	- Contract	**PR**	- Project
G	- Grant	**TA**	- Task
PE	- Program Element	**WU**	- Work Unit Accession No.

Block 6. Author(s). Name(s) of person(s) responsible for writing the report, performing the research, or credited with the content of the report. If editor or compiler, this should follow the name(s).

Block 7. Performing Organization Name(s) and Address(es). Self-explanatory.

Block 8. Performing Organization Report Number. Enter the unique alphanumeric report number(s) assigned by the organization performing the report.

Block 9. Sponsoring/Monitoring Agency Name(s) and Address(es). Self-explanatory.

Block 10. Sponsoring/Monitoring Agency Report Number. *(If known)*

Block 11. Supplementary Notes. Enter information not included elsewhere such as; prepared in cooperation with...; Trans. of...; To be published in.... When a report is revised, include a statement whether the new report supersedes or supplements the older report.

Block 12a. Distribution/Availability Statement. Denotes public availability or limitations. Cite any availability to the public. Enter additional limitations or special markings in all capitals (e.g. NORFORN, REL, ITAR).

- **DOD** - See DoDD 4230.25, "Distribution Statements on Technical Documents."
- **DOE** - See authorities.
- **NASA** - See Handbook NHB 2200.2.
- **NTIS** - Leave blank.

Block 12b. Distribution Code.

- **DOD** - Leave blank
- **DOE** - Enter DOE distribution categories from the Standard Distribution for Unclassified Scientific and Technical Reports
- **NASA** - Leave blank.
- **NTIS** - Leave blank.

Block 13. Abstract. Include a brief *(Maximum 200 words)* factual summary of the most significant information contained in the report.

Block 14. Subject Terms. Keywords or phrases identifying major subjects in the report.

Block 15. Number of Pages. Enter the total number of pages.

Block 16. Price Code. Enter appropriate price code *(NTIS only).*

Block 17. - 19. Security Classifications. Self-explanatory. Enter U.S. Security Classification in accordance with U.S. Security Regulations (i.e., UNCLASSIFIED). If form contains classified information, stamp classification on the top and bottom of the page.

Block 20. Limitation of Abstract. This block must be completed to assign a limitation to the abstract. Enter either UL (unlimited) or SAR (same as report). An entry in this block is necessary if the abstract is to be limited. If blank, the abstract is assumed to be unlimited.