EMPIRICAL BAYES TWO-STAGE SELECTION PROCEDURES FOR SELECTING THE BEST BERNOULLI TREATMENT USING THE INVERSE BINOMIAL SAMPLING*

by

Shanti S. Gupta
Purdue University

and

TaChen Liang
Wayne State University

Xun Lin
Purdue University

Technical Report #97-16C

Department of Statistics
Purdue University
West Lafayette, IN USA

September 1997

EMPIRICAL BAYES TWO-STAGE SELECTION PROCEDURES FOR SELECTING THE BEST BERNOULLI TREATMENT USING THE INVERSE BINOMIAL SAMPLING*

by

Shanti S. Gupta
Department of Statistics
Purdue University
West Lafayette, IN 47907

and

TaChen Liang
Department of Mathematics
Wayne State University
Detroit, MI 48202

Xun Lin
Purdue University
Department of Statistics
West Lafayette, IN 47907

Abstract

In this paper we investigate the problem of selecting the treatment with the largest probability of success from $k(\geq 2)$ independent Bernoulli treatments. The selected treatment must also be better than a given control. We employ the empirical Bayes approach and develop a two-stage selection procedure. We prove that the proposed selection rule is asymptotically optimal at the rate of convergence of order $O(\exp(-c^*n))$, for some positive constant c^*, where n is the number of the historical data at hand. We also carry out a simulation study to investigate the performance of the proposed empirical Bayes selection procedure for small to moderate values of n. The simulated results are provided in the paper.

AMS Classification: primary 62F07; secondary 62C12.

Keywords: Asymptotically optimal; empirical Bayes; two-stage selection procedure; rate of convergence; inverse binomial sampling.

---

*This research was supported in part by US Army Research Office, Grant DAAH04-95-1-0165 at Purdue University.

1. Introduction

Consider k independent Bernoulli treatments $\pi_1, \ldots, \pi_k$, where for each i, treatment π_i is characterized by the probability of success θ_i. Let $\theta_{[1]} \leq \ldots \leq \theta_{[k]}$ denote the ordered values of the parameters $\theta_1, \ldots, \theta_k$. It is assumed that the exact pairing between the ordered and the unordered parameters is unknown. A treatment π_i with $\theta_i = \theta_{[k]}$ is called the best among the k underlying treatments. The Bernoulli treatment occurs in many fields, such as medicine, engineering and sociology. A number of statistical procedures have been studied to find the best Bernoulli treatment. Tamhane (1980) considered the problem of selecting the better of two Bernoulli treatments using a matched pair sample design. Sanchez (1987) investigated a modified least-failure sampling method for Bernoulli subset selection. Yang (1989) studied the selection procedures through a Bayes approach. Gupta and Liang (1986, 1988, 1989) developed empirical Bayes procedures for selecting the best Bernoulli treatment using binomial sampling design. In two recent papers, Gupta, Liang and Rau (1994, 1995) have specifically, investigated the empirical Bayes selection procedures for the Bernoulli and normal populations using two-stage procedures. A good review in this area is contained in Gupta and Panchapakesan (1996). For further references in the area, see also Gupta and Panchapakesan (1979) and Bechhofer, Santner and Goldsman (1996).

In this paper, we are concerned with the problem of selecting the best Bernoulli treatment compared with a specific standard using the inverse binomial sampling scheme. Unlike binomial sampling, in which the sample size (number of trials) from each treatment is fixed and therefore the pertaining random variable is always being bounded, the sample size in inverse binomial sampling is not fixed and its associated random observation is unbounded.

The formulation of the selection problem is described in Section 2. A Bayes two-stage selection procedure is derived in Section 3. We construct an empirical Bayes two-stage selection procedure in Section 4. The asymptotic optimality of the proposed empirical Bayes two-stage selection procedure is investigated in Section 5. We prove that the convergence rate of the proposed selection procedure is of exponential order, that is, $O(\exp(-c^* n))$, for some positive constant c^*, where n is the number of the historical data at hand. In

Section 6, we carry out a simulation study to investigate the performance of the proposed selection procedure for small to moderate values of n. The simulated results are consistent with the rate of convergence obtained in Section 5.

2. Formulation of the Selection Problem

Consider k independent Bernoulli treatments $\pi_1, \ldots, \pi_k$, with unknown success probabilities $\theta_1, \ldots, \theta_k$, respectively. Let $\theta_{[1]} \leq \ldots \theta_{[k]}$ be the ordered values of the parameters $\theta_1, \ldots, \theta_k$. We assume that the exact pairing between the ordered and unordered parameters is unknown. Any treatment associated with the largest success probability $\theta_{[k]}$ is defined as the best treatment. For a given standard $\theta_0 (0 < \theta_0 < 1)$, treatment π_i is defined to be good if the corresponding $\theta_i > \theta_0$, and bad otherwise. Our goal is to select a treatment which is the best among the k treatments and also good compared with the standard θ_0. If there is no such treatment, we select none.

A two-stage selection procedure is defined as follows. First, from each of the k treatments, a sequence of Bernoulli trials are carried out until a fixed number m_1 of successes are achieved. Let X_i be the number of failures before observing the m_1-th success from π_i. Based on the observations $\underset{\sim}{X} = (X_1, \ldots, X_k)$, we decide whether the selection should be made immediately or not. If we decide to make the selection immediately, we may select a treatment from the k treatments based on $\underset{\sim}{X}$. Or we may select none in which case the k treatments are all excluded as bad. If we decide not to make the selection immediately, then based on $\underset{\sim}{X}$, one (and only one) treatment, say, π_i is chosen. And from this chosen treatment π_i, a sequence of Bernoulli trials are carried out until m_2 additional successes are achieved. We let Y_i be the number of failures before observing the m_2-th success at the second-stage sampling. Then based on $\underset{\sim}{X}$ and Y_i, we decide either to select treatment π_i as the best treatment and consider it to be good, or select none and thus exclude all k treatments as bad.

Let $\Omega = \{\underset{\sim}{\theta} = (\theta_1, \ldots, \theta_k) | 0 \leq \theta_i \leq 1, i = 1, \ldots, k\}$ be the parameter space. Let $\underset{\sim}{a} = (a_0, \ldots, a_k)$ be an action, where $a_i = 0, 1; i = 0, 1, \ldots, k$ and $\sum_{i=0}^{k} a_i = 1$. For each $i = 1, \ldots, k, a_i = 1$ means that treatment π_i is selected as the best and also considered to be good compared with θ_0. $a_0 = 1$ means that all the k treatments are excluded as bad and none is selected. We define the termination action t as follows:

$$t = \begin{cases} 1, & \text{if the selection is made immediately after } \underset{\sim}{X} \text{ is observed,} \\ 0, & \text{otherwise.} \end{cases}$$

When $t = 0$, let $\underset{\sim}{\Delta} = (\Delta_1, \ldots, \Delta_k)$ be the identity action, where $\Delta_i = 0, 1; i = 1, \ldots, k$, and $\sum_{i=1}^{k} \Delta_i = 1$. $\Delta_i = 1$ means that the second stage sampling is taken from treatment π_i.

The loss function is defined to be

$$L(\underset{\sim}{\theta}, (\underset{\sim}{a}, t, \underset{\sim}{\Delta})) = \max(\theta_{[k]}, \theta_0) - t \sum_{i=0}^{k} a_i \theta_i + kc_1$$

$$+(1-t)\Big\{ - \sum_{i=1}^{k} \Delta_i \Big(a_i\theta_i + (1-a_i)\theta_0 \Big) + c_2 \Big\}, \tag{2.1}$$

where $c_1 > 0$ is the cost of sampling from each of the k treatments at the first stage and c_2 is the cost of the sampling at the second stage.

Conditional on θ_i, we have $X_i \sim \text{NB}(m_1, \theta_i)$ and $Y_i \sim \text{NB}(m_2, \theta_i)$. Here, $\text{NB}(\cdot, \cdot)$ denotes the negative binomial distribution with the two parameters. Given θ_i, let $f_i(x|\theta_i)$ and $g_i(y|\theta_i)$ be the conditional probability functions of X_i and Y_i, respectively. That is,

$$f_i(x|\theta_i) = \binom{m_1 + x - 1}{x} \theta_i^{m_1} (1-\theta_i)^x; \qquad x = 0, 1, \ldots$$

and

$$g_i(y|\theta_i) = \binom{m_2 + y - 1}{y} \theta_i^{m_2} (1-\theta_i)^y; \qquad y = 0, 1, \ldots$$

It is assumed that for each $i = 1, \ldots, k, \theta_i$ is a realization of a random variable Θ_i which has a Beta distribution with probabiliity density function $h_i(\theta_i|\alpha_i, \mu_i)$, where

$$h_i(\theta_i|\alpha_i, \mu_i) = \frac{\Gamma(\alpha_i)}{\Gamma(\alpha_i\mu_i)\Gamma(\alpha_i(1-\mu_i))} \theta_i^{\alpha_i\mu_i - 1} (1-\theta_i)^{\alpha_i(1-\mu_i)-1}, \qquad 0 < \theta_i < 1.$$

Both α_i and μ_i are unknown, $0 < \mu_i < 1$ and $\alpha_i > 0$. Note that $E(\Theta_i) = \mu_i$ and $\text{Var}(\Theta_i) = \frac{\mu_i(1-\mu_i)}{\alpha_i + 1}$. It is also assumed that $\Theta_1, \ldots, \Theta_k$ are mutually independent.

Denote $\underset{\sim}{Y} = (Y_1, \ldots, Y_k)$. Let $\mathcal{X}$ and $\mathcal{Y}$ be the sample spaces generated by $\underset{\sim}{X}$ and $\underset{\sim}{Y}$, respectively. A two-stage selection procedure, in general, consists of the following rules:

a. Stopping rule τ: For each $\underset{\sim}{x} \in \mathcal{X}, \tau(\underset{\sim}{x})$ is the probability of terminating the sampling after observing $\underset{\sim}{x}$ and making a selection decision immediately based on $\underset{\sim}{x}$.

b. Identity rule $\underset{\sim}{\delta} = (\delta_1, \ldots, \delta_k)$: For each $\underset{\sim}{x} \in \mathcal{X}, \delta_i(\underset{\sim}{x})$ is the probability of taking the second-stage inverse binominal sample from treatment π_i when we decide not to make a selection decision immediately after observing $\underset{\sim}{x}$. Note that $\underset{\sim}{\delta}$ satisfies $\sum_{i=1}^{k} \delta_i(\underset{\sim}{x}) = 1$ for each $\underset{\sim}{x} \in \mathcal{X}$.

c. First-stage selection rule $\underset{\sim}{d}_1 = (d_{10}, \ldots, d_{1k})$: For each $\underset{\sim}{x} \in \mathcal{X}$ and $i = 1, \ldots, k$, let $d_{1i}(\underset{\sim}{x})$ be the probability of selecting treatment π_i as the best from among the k treatments and also good compared with θ_0, and let $d_{10}(\underset{\sim}{x})$ be the probability of excluding all the k treatments and selecting none. For all $\underset{\sim}{x} \in \mathcal{X}$, $\sum_{i=0}^{k} d_{1i}(\underset{\sim}{x}) = 1$.

d. Second-stage selection rule $\underset{\sim}{d}_2 = (d_{20}, \ldots, d_{2k})$. For each $\underset{\sim}{x} \in \mathcal{X}$ and $y \in \mathcal{Y}$, when the decision of going to take the second-stage sample from π_i is made, $d_{2i}(\underset{\sim}{x}, y)$ is the probability of selecting π_i as the best from among the k treatments and also considered to be good compared with θ_0, for $i = 1, \ldots, k$. Note that $d_{2i}(\underset{\sim}{x}, y)$ depends on $\underset{\sim}{Y}$ only through Y_i since there are no second-stage observations from any other treatments. Therefore, we denote $d_{2i}(\underset{\sim}{x}, y_i) = d_{2i}(\underset{\sim}{x}, y), 1 \leq i \leq k$. Let $d_{20}(\underset{\sim}{x}, y)$ $= \sum_{i=1}^{k} \delta_i(\underset{\sim}{x})(1 - d_{2i}(\underset{\sim}{x}, y_i))$ be the probability of excluding all the k treatments and selecting none, based on $\underset{\sim}{x}$ and y.

Under the preceding statistical model, the Bayes risk of the two-stage selection procedure $(\tau, \underset{\sim}{\delta}, \underset{\sim}{d}_1, \underset{\sim}{d}_2)$ is denoted by $R(\tau, \underset{\sim}{\delta}, \underset{\sim}{d}_1, \underset{\sim}{d}_2)$.

We have

$$
\begin{aligned}
R(\tau, \underset{\sim}{\delta}, \underset{\sim}{d}_1, \underset{\sim}{d}_2) &= C - \sum_{\underset{\sim}{x} \in \mathcal{X}} \tau(\underset{\sim}{x}) \left(\sum_{i=0}^{k} d_{1i}(\underset{\sim}{x}) \varphi_i(x_i) \right) f(\underset{\sim}{x}) + \sum_{\underset{\sim}{x} \in \mathcal{X}} (1 - \tau(\underset{\sim}{x})) \\
&\quad \times \left\{ c_2 - \theta_0 + \sum_{i=1}^{k} \delta_i(\underset{\sim}{x}) \left(\sum_{y_i=0}^{+\infty} d_{2i}(\underset{\sim}{x}, y_i)(\theta_0 - \psi_i(x_i + y_i)) f_i(y_i | x_i, \alpha_i, \mu_i) \right) \right\} f(\underset{\sim}{x}) \qquad (2.2) \\
&= \sum_{\underset{\sim}{x} \in \mathcal{X}} \tau(\underset{\sim}{x}) \left\{ \begin{array}{l} \sum_{i=0}^{k} d_{1i}(\underset{\sim}{x})(\theta_0 - \varphi_i(x_i)) - c_2 + \theta_0 \\ - \sum_{i=1}^{k} \delta_i(\underset{\sim}{x}) \left(\sum_{y_i=0}^{+\infty} d_{2i}(\underset{\sim}{x}, y_i)(\theta_0 - \psi_i(x_i + y_i)) f_i(y_i | x_i, \alpha_i, \mu_i) \right) \end{array} \right\} f(\underset{\sim}{x})
\end{aligned}
$$

$$+\sum_{\underset{\sim}{x}\in\mathcal{X}}\left\{c_2-\theta_0+\sum_{i=1}^{k}\delta_i(\underset{\sim}{x})\left(\sum_{y_i=0}^{+\infty}d_{2i}(\underset{\sim}{x},y_i)(\theta_0-\psi_i(x_i+y_i))f_i(y_i|x_i,\alpha_i,\mu_i)\right)\right\}f(\underset{\sim}{x})$$
$$+C. \tag{2.3}$$

where $\varphi_i(x_i)=E(\Theta_i|X_i=x_i)=\frac{\alpha_i\mu_i+m_i}{\alpha_i+m_1+x_i}$ is the posterior mean of Θ_i given $X_i=x_i, i=1,\ldots,k, \varphi_0(x_0)=\theta_0; \psi_i(x_i+y_i)=\frac{\alpha_i\mu_i+m_1+m_2}{\alpha_i+m_1+m_2+x_i+y_i}$ is the posterior mean of Θ_i given $X_i=x_i$ and $Y_i=y_i$ for each $i=1,\ldots,k$. Also $C=\int_\Omega \max(\theta_{[k]},\theta_0)dH(\underset{\sim}{\theta})+kc_1$ and $H(\underset{\sim}{\theta})$ is the joint distribution of $\underset{\sim}{\Theta}=(\Theta_1,\ldots,\Theta_k); f(\underset{\sim}{x})=\prod_{i=1}^{k}f_i(x_i)$ where

$$\begin{aligned}f_i(x_i)&=\int_0^1 f_i(x_i|\theta_i)h_i(\theta_i|\alpha_i,\mu_i)d\theta_i\\&=\binom{m_1+x_i-1}{x_i}\frac{\Gamma(\alpha_i)\Gamma(m_1+\alpha_i+\mu_i)\Gamma(x_i)+\alpha_i(1-\mu_i))}{\Gamma(\alpha_i\mu_i)\Gamma(\alpha_i(1-\mu_i))\Gamma(m_1+x_i+\alpha_i)}\end{aligned}$$

is the marginal probability function of X_i.

$$f_i(y_i|x_i,\alpha_i,\mu_i)=\frac{f_i(x_i,y_i)}{f_i(x_i)},\quad\text{where}\quad f_i(x_i,y_i)=\int_0^1 f_i(x_i|\theta_i)g_i(y_i|\theta_i)h_i(\theta_i|\alpha_i,\mu_i)d\theta_i$$

Note that $f_i(y_i|x_i,\alpha_i,\mu_i)$ is the marginal conditional probability function of Y_i given $X_i=x_i$ and (α_i,μ_i). A direct computation yields

$$\begin{aligned}f_i(y_i|x_i,\alpha_i,\mu_i)=&\binom{m_2+y_i-1}{y_i}\\&\times\frac{\Gamma(m_1+x_i+\alpha_i)\Gamma(m_1+m_2+\alpha_i\mu_i)\Gamma(x_i+y_i+\alpha_i(1-\mu_i))}{\Gamma(m_1+m_2+x_i+y_i+\alpha_i)\Gamma(m_1+\alpha_i\mu_i)\Gamma(x_i+\alpha_i(1-\mu_i))}.\end{aligned}$$

3. Derivation of a Bayes Two-stage Selection Procedure

We derive a Bayes two-stage selection procedure as follows.

a. First-stage selection rule

For each $\underset{\sim}{x}\in\mathcal{X}$, let $I(\underset{\sim}{x})=\{i|\varphi_i(x_i)=\max_{0\le j\le k}\varphi_j(x_j), i=0,1,\ldots,k\}$. Define $i^*=\min\{i|i\in I(\underset{\sim}{x})\}$. Then we have a first-stage selection rule $\underset{\sim}{d}_1^B=(d_{10}^B,d_{11}^B,\ldots,d_{1k}^B)$ as follows:

$$\begin{cases} d^B_{1i^*}(\underset{\sim}{x}) = 1, \\ d^B_{1j}(\underset{\sim}{x}) = 0, \quad \text{for} \quad j \neq i^*. \end{cases} \tag{3.1}$$

b. Second-stage selection rule

We define a second-stage selection rule $\underset{\sim}{d}^B_2 = (d^B_{20}, d^B_{21}, \dots d^B_{2k})$ as follows: for $i = 1, \dots, k$,

$$d^B_{2i}(\underset{\sim}{x}, \underset{\sim}{y}) = d^B_{2i}(\underset{\sim}{x}, y_i) = \begin{cases} 1, & \text{if} \quad \psi_i(x_i + y_i) > \theta_0; \\ 0, & \text{otherwise,} \end{cases} \tag{3.2}$$

and

$$d^B_{20}(\underset{\sim}{x}, \underset{\sim}{y}) = \sum_{i=1}^{k} \delta^B_i(\underset{\sim}{x})(1 - d^B_{2i}(\underset{\sim}{x}, y_i))$$

where $\underset{\sim}{\delta}^B(\underset{\sim}{x}) = (\delta^B_1(\underset{\sim}{x}), \dots, \delta^B_k(\underset{\sim}{x}))$ is the identity rule defined below.

c. Identity rule

For each $i = 1, \dots, k$, and $\underset{\sim}{x} \in \mathcal{X}$, define

$$T_i(\underset{\sim}{x}) = \sum_{y_i=0}^{\infty} d^B_{2i}(\underset{\sim}{x}, y_i)(\theta_0 - \psi_i(x_i + y_i)) f_i(y_i | x_i, \alpha_i, \mu_i). \tag{3.3}$$

Let $J(\underset{\sim}{x}) = \{j | T_j(\underset{\sim}{x}) = \min_{1 \le j \le k} T_i(\underset{\sim}{x})\}$ and $j^* = \min\{j | j \in J(\underset{\sim}{x})\}$.

We then define an identity rule $\underset{\sim}{\delta}^B = (\delta^B_1, \dots, \delta^B_k)$ as follows:

$$\delta^B_j(\underset{\sim}{x}) = \begin{cases} 1, & \text{if} \quad j = j^*; \\ 0, & \text{otherwise.} \end{cases} \tag{3.4}$$

d. Stopping rule

For each $\underset{\sim}{x} \in \mathcal{X}$, let

$$Q(\underset{\sim}{x}) = \sum_{i=0}^{k} d^B_{1i}(\underset{\sim}{x})(\theta_0 - \varphi_i(x_i)) - c_2 - \sum_{i=1}^{k} \delta^B_i(\underset{\sim}{x}) T_i(\underset{\sim}{x}). \tag{3.5}$$

The stopping rule τ^B is defined to be

$$\tau^B(\underset{\sim}{x}) = \begin{cases} 1, & \text{if} \quad Q(\underset{\sim}{x}) \le 0; \\ 0, & \text{otherwise.} \end{cases} \tag{3.6}$$

Next we will show that the selection procedure $(\tau^B, \underset{\sim}{\delta}^B, \underset{\sim}{d}_1^B, \underset{\sim}{d}_2^B)$ is a Bayes two-stage selection procedure.

Theorem 3.1 The two-stage selection procedure $(\tau^B, \underset{\sim}{\delta}^B, \underset{\sim}{d}_1^B, \underset{\sim}{d}_2^B)$ is a Bayes two-stage selection procedure.

Proof: Let $(\tau, \underset{\sim}{\delta}, \underset{\sim}{\delta}_1, \underset{\sim}{\delta}_2)$ be any two-stage selection procedure. It suffices to show $R(\tau, \underset{\sim}{\delta}, \underset{\sim}{d}_1, \underset{\sim}{d}_2) - R(\tau^B, \underset{\sim}{\delta}^B, \underset{\sim}{d}_1^B, \underset{\sim}{d}_2^B) \ge 0$. We have

$$R(\tau, \underset{\sim}{\delta}, \underset{\sim}{d}_1, \underset{\sim}{d}_2) - R(\tau^B, \underset{\sim}{\delta}^B, \underset{\sim}{d}_1^B, \underset{\sim}{d}_2^B) = I + II + III, \tag{3.7}$$

where

$$\begin{cases} I = R(\tau, \underset{\sim}{\delta}, \underset{\sim}{d}_1, \underset{\sim}{d}_2) - R(\tau, \underset{\sim}{\delta}, \underset{\sim}{d}_1^B, \underset{\sim}{d}_2^B), \\ II = R(\tau, \underset{\sim}{\delta}, \underset{\sim}{d}_1^B, \underset{\sim}{d}_2^B) - R(\tau, \underset{\sim}{\delta}^B, \underset{\sim}{d}_1^B, \underset{\sim}{d}_2^B), \\ III = R(\tau, \underset{\sim}{\delta}^B, \underset{\sim}{d}_1^B, \underset{\sim}{d}_2^B) - R(\tau^B, \underset{\sim}{\delta}^B, \underset{\sim}{d}_1^B, \underset{\sim}{d}_2^B). \end{cases} \tag{3.8}$$

From (2.2),

$$\begin{aligned} I = & \sum_{\underset{\sim}{x} \in \mathcal{X}} \tau(\underset{\sim}{x}) \Big(\sum_{i=0}^{k} (d_{1i}^B(\underset{\sim}{x}) - d_{1i}(\underset{\sim}{x})) \varphi_i(x_i) \Big) f(\underset{\sim}{x}) \\ & + \sum_{\underset{\sim}{x} \in \mathcal{X}} (1 - \tau(\underset{\sim}{x})) \Big\{ \sum_{i=1}^{k} \delta_i(\underset{\sim}{x}) \Big(\sum_{y_i=0}^{+\infty} (d_{2i}(\underset{\sim}{x}, y_i) - d_{2i}^B(\underset{\sim}{x}, y_i))(\theta_0 - \psi_i(x_i + y_i)) \\ & \qquad \times \; f_i(y_i | x_i, \alpha_i, \mu_i) \Big) \Big\} f(\underset{\sim}{x}) \end{aligned} \tag{3.9}$$

By the definition of $\underset{\sim}{d}_1^B$ and $\underset{\sim}{d}_1$, for each $i = 1, \dots, k$,

$$\sum_{i=0}^{k} (d_i^B(\underset{\sim}{x}) - d_{1i}(\underset{\sim}{x})) \varphi_i(x_i) = \max_{0 \le i \le k} (\varphi_i(x_i)) - \sum_{i=1}^{k} d_{1i}(\underset{\sim}{x}) \varphi_i(x_i) \ge 0,$$

and by the definition of $\underset{\sim}{d}_2^B$, we have $(d_{2i}(\underset{\sim}{x}, y_i) - d_{2i}^B(\underset{\sim}{x}, y_i))(\theta_0 - \psi_i(x_i + y_i)) \ge 0$. Therefore, $I \ge 0$ since all the other terms in (3.9) are nonnegative.

From (2.2) again, since

$$\sum_{i=1}^{k}(\delta_i(\underset{\sim}{x}) - \delta_i^B(\underset{\sim}{x}))T_i(\underset{\sim}{x}) = \sum_{i=1}^{k}\delta_i(\underset{\sim}{x})T_i(\underset{\sim}{x}) - \min_{1\le i\le k} T_i(\underset{\sim}{x}) \ge 0, \tag{3.10}$$

we have

$$II = \sum_{\underset{\sim}{x}\in\mathcal{X}}(1-\tau(\underset{\sim}{x}))\left(\sum_{i=1}^{k}(\delta_i(\underset{\sim}{x}) - \delta_i^B(\underset{\sim}{x}))T_i(\underset{\sim}{x})\right)f(\underset{\sim}{x}) \ge 0. \tag{3.11}$$

From (2.3), by the definition of τ^B,

$$III = \sum_{\underset{\sim}{x}\in\mathcal{X}}(\tau(\underset{\sim}{x}) - \tau^B(\underset{\sim}{x}))Q(\underset{\sim}{x})f(\underset{\sim}{x}) \ge 0. \tag{3.12}$$

The proof of Theorem 3.1 is completed by combining (3.7), (3.9), (3.11) and (3.12).

4. The Construction of an Empirical Bayes Two-Stage Selection Procedure

4.1. Empirical Bayes Framework

The Bayes two-stage selection procedure $(\tau^B, \underset{\sim}{\delta}^B, \underset{\sim}{d}_1^B, \underset{\sim}{d}_2^B)$ defined in Section 3 depends on the unknown parameters $(\alpha_i, \mu_i), i = 1, \ldots, k$. Since the parameters are unknown, it is impossible to implement the Bayes two-stage selection procedure for the selection problem in practice. In the empirical Bayes framework, it is generally assumed that there are some past observations when the present selection is to be made. At time $j = 1, \ldots, n$, let X_{ij} be the number of failures before observing the m_1-th success from π_i. Let $\underset{\sim}{\Theta}_j = (\Theta_{1j}, \ldots, \Theta_{kj})$ be a random vector where Θ_{ij} stands for the probability of success from treatment π_i at time j. We assume that $\underset{\sim}{\Theta}_j, j = 1, \ldots, n$, are i.i.d with prior density $h(\underset{\sim}{\theta}_j) = \prod_{i=1}^{k} h_i(\theta_{ij}|\alpha_i, \mu_i)$. Therefore, conditional on $\Theta_{ij} = \theta_{ij}, X_{ij}|\theta_{ij} \sim \mathrm{NB}(m_1, \theta_{ij})$ and X_{ij} has a marginal probability function $f_i(x)$. Let $\underset{\sim}{X}_j = (X_{ij}, \ldots, X_{kj})$ be the random observations of the first-stage sampling taken at time $j = 1, \ldots, n$. Let $\underset{\sim}{X}_{n+1} = \underset{\sim}{X} = (X_1, \ldots, X_k)$ be the current (present) observations taken at the first stage..

4.2 Certain Useful Empirical Bayes Estimates

In order to construct an empirical Bayes two-stage selection procedure, we first introduce certain properties related to the parameters (α_i, μ_i). Then based on these properties, the empirical Bayes estimators are derived. We then construct an empirical Bayes two-stage selection procedure in the next subsection.

For each $i = 1, \ldots, k$ and $j = 1, 2, \ldots, n$, we have

$$E\left(\frac{\binom{m_1+X_{ij}-2}{m_1-2}}{\binom{m_1+X_{ij}-1}{m_1-1}}\right) = E\left(E\left[\frac{\binom{m_1+X_{ij}-2}{m_1-2}}{\binom{m_1+X_{ij}-1}{m_1-1}}\Big|\Theta_{ij}\right]\right) = E\Theta_{ij} = \mu_i, \tag{4.1}$$

and

$$E\left(\frac{\binom{m_1+X_{ij}-3}{m_1-1}}{\binom{m_1+X_{ij}-1}{m_1-1}} I(X_{ij} \geq 2)\right) = E\left(E\left[\frac{\binom{m_1+X_{ij}-3}{m_1-1}}{\binom{m_1+X_{ij}-1}{m_1-1}} I(X_{ij} \geq 2)\Big|\Theta_{ij}\right]\right)$$

$$= E(1-\Theta_{ij})^2 = \frac{(\alpha_i(1-\mu_i)+1)(1-\mu_i)}{\alpha_i+1} \equiv \nu_i \tag{4.2}$$

Let $C_i = \nu_i - (1-\mu_i)^2$ and $D_i = 1 - \mu_i - \nu_i$, for $i = 1, \ldots, k$. From (4.2), we have,

$$\begin{cases} C_i = \frac{(1-\mu_i)\mu_i}{\alpha_i+1}, \\ D_i = \frac{\alpha_i(1-\mu_i)\mu_i}{\alpha_i+1}. \end{cases} \tag{4.3}$$

Note that for $i = 1, \ldots, k, \alpha_i = D_i/C_i$. Since $\alpha_i > 0$ and $0 < \mu_i < 1$. both C_i and D_i are positive.

For $i = 1, \ldots, k$, based on $(X_{i1}, \ldots, X_{in})$, let,

$$\begin{cases} \mu_{in} = \frac{1}{n}\sum_{j=1}^{n} \frac{\binom{m_1+X_{ij}-2}{m_1-2}}{\binom{m_1+X_{ij}-1}{m_1-1}}, \\ \nu_{in} = \frac{1}{n}\sum_{j=1}^{n} \frac{\binom{m_1+X_{ij}-3}{m_i-1}}{\binom{m_1+X_{ij}-1}{m_1-1}} I(X_{ij} \geq 2), \end{cases} \tag{4.4}$$

and

$$\begin{cases} C_{in} = \nu_{in} - (1-\mu_{in})^2, \\ D_{in} = 1 - \mu_{in} - \nu_{in}. \end{cases} \tag{4.5}$$

Then we have

$$\begin{aligned} D_{in} &= \frac{1}{n}\sum_{j=1}^{n}\left(1 - \frac{\binom{m_1+X_{ij}-2}{m_1-2}}{\binom{m_1+X_{ij}-1}{m_1-1}} - \frac{\binom{m_1+X_{ij}-3}{m_1-1}}{\binom{m_1+X_{ij}-1}{m_1-1}} I(X_{ij} \geq 2)\right) \\ &= \frac{1}{n}\sum_{j=1}^{n}\left(1 - \frac{m_1-2}{m_1+X_{ij}-2} - \frac{(X_{ij}-1)(X_{ij}-2)}{(m_1+X_{ij}-2)(m_1+X_{ij}-3)} I(X_{ij} \geq 2)\right) \\ &\geq 0. \end{aligned}$$

However, C_{in} may be nonpositive while $C_i > 0$. From $\alpha_i = D_i/C_i$, we can see that $\alpha_i \to +\infty$ as $C_i \to 0$. So we define

$$\alpha_{in} = \begin{cases} D_{in}/C_{in}, & \text{if} \quad C_{in} > 0, \\ \infty, & \text{otherwise.} \end{cases} \tag{4.6}$$

We can also see that $\lim_{\alpha_i \to +\infty} \varphi_i(x_i) = \mu_i$ and $\lim_{\alpha_i \to +\infty} \psi_i(x_i + y_i) = \mu_i$. Therefore, we define empirical Bayes estimators $\varphi_{in}(x_i)$ and $\psi_{in}(x_i + y_i)$ for the posterior means $\varphi_i(x_i)$ and $\psi_i(x_i + y_i)$, respectively, as follows:

$$\begin{cases} \varphi_{in}(x_i) = \begin{cases} \frac{\alpha_{in}\mu_{in}+m_1}{\alpha_{in}+x_i+m_1} & \text{if} \quad C_{in} > 0; \\ \mu_{in}, & \text{otherwise,} \end{cases} \\ \psi_{in}(x_i + y_i) = \begin{cases} \frac{\alpha_{in}\mu_{in}+m_1+m_2}{\alpha_{in}+x_i+y_i+m_1+m_2}, & \text{if} \quad C_{in} > 0; \\ \mu_{in}, & \text{otherwise.} \end{cases} \end{cases} \tag{4.7}$$

Let $\varphi_{on}(x_0) = \theta_0$. Since $\lim_{\alpha_i \to +\infty} f_i(y_i|x_i, \alpha_i, \mu_i) = \binom{m_2+y_i-1}{y_i}\mu_i^{m_2}(1-\mu_i)^{y_i}$, we define an empirical Bayes estimator for the marginal conditional probability function $f_i(y_i|x_i, \alpha_i, \mu_i)$:

$$f_{in}(y_i|x_i, \alpha_{in}, \mu_{in}) = \begin{cases} f_i(y_i|x_i, \alpha_{in}, \mu_{in}), & \text{if} \quad C_{in} > 0; \\ \binom{m_2+y_i-1}{y_i}\mu_{in}^{m_2}(1-\mu_{in})^{y_i}, & \text{otherwise.} \end{cases} \tag{4.8}$$

4.3 The Proposed Empirical Bayes Two-Stage Selection Procedure

We now propose an empirical Bayes two-stage selection procedure $(\tau^{*n}, \underset{\sim}{\delta}^{*n}, \underset{\sim}{d}_1^{*n}, \underset{\sim}{d}_2^{*n})$ as follows:

a. Empirical Bayes first-stage selection rule $\underset{\sim}{d}_1^{*n} = (d_{10}^{*n}, d_{11}^{*n}, \ldots, d_{1k}^{*n})$.

Let $I_n(\underset{\sim}{x}) = \{i|\varphi_{in}(\underset{\sim}{x}) = \max_{0 \le j \le k} \varphi_{jn}(x_j), i = 0, 1, \ldots, k\}$. Define $i_n^* = \min\{i|i \in I_n(\underset{\sim}{x})\}$. Then we define $\underset{\sim}{d}_1^{*n}$ as follows:

$$\begin{cases} d_{1i_n^*}^{*n}(\underset{\sim}{x}) = 1, \\ d_{1j}^{*n}(\underset{\sim}{x}) = 0, \quad \text{for} \quad j \ne i_n^*. \end{cases} \tag{4.9}$$

b. Empirical Bayes second-stage selection rule $\underset{\sim}{d}_2^{*n} = (d_{20}^{*n}, d_{21}^{*n}, \ldots, d_{2k}^{*n})$

For $i = 1, \ldots, k$, let

$$d_{2i}^{*n}(\underset{\sim}{x}, y) = d_{2i}^{*n}(\underset{\sim}{x}, y_i) = \begin{cases} 1, & \text{if} \quad \psi_{in}(x_i + y_i) > \theta_0; \\ 0, & \text{otherwise,} \end{cases} \tag{4.10}$$

and

$$d_{20}^{*n}(\underset{\sim}{x}, y) = \sum_{i=1}^{k} \delta_i^{*n}(\underset{\sim}{x})(1 - d_{2i}^{*n}(\underset{\sim}{x}, y_i))$$

where $\underset{\sim}{\delta}^{*n} = (\delta_1^{*n}, \ldots, \delta_k^{*n})$ is the empirical Bayes identity rule defined below.

c. Empirical Bayes identity rule $\underset{\sim}{\delta}^{*n} = (\delta_1^{*n}, \ldots, \delta_k^{*n})$

For $i = 1, \ldots, k$, define

$$T_{in}(\underset{\sim}{x}) = \sum_{y_i=0}^{+\infty} d_{2i}^{*n}(\underset{\sim}{x}, y_i)(\theta_0 - \psi_{in}(x_i + y_i)) f_{in}(y_i | x_i, \alpha_{in}, \mu_{in}) \tag{4.11}$$

Let $J_n(\underset{\sim}{x}) = \{j | T_{jn}(\underset{\sim}{x}) = \min_{1 \le i \le k} T_{in}(\underset{\sim}{x}), j = 1, \ldots, k\}$ and $j^* = \min\{j | j \in J_n(\underset{\sim}{x})\}$. Then the empirical Bayes identity rule $\underset{\sim}{\delta}^{*n}$ is

$$\delta_j^{*n}(\underset{\sim}{x}) = \begin{cases} 1, & \text{if} \;\; j = j_n^*, \\ 0, & \text{otherwise.} \end{cases} \tag{4.12}$$

d. Empirical Bayes Stopping Rule

For each $\underset{\sim}{x} \in \mathcal{X}$, let

$$Q_n(\underset{\sim}{x}) = \sum_{i=0}^{k} d_{1i}^{*n}(\underset{\sim}{x})(\theta_0 - \varphi_{in}(x_i)) - c_2 - \sum_{i=1}^{k} \delta_i^{*n}(\underset{\sim}{x}) T_{in}(\underset{\sim}{x}) \tag{4.13}$$

and

$$\tau^{*n}(\underset{\sim}{x}) = \begin{cases} 1, & \text{if} \quad Q_n(\underset{\sim}{x}) \le 0, \\ 0, & \text{otherwise.} \end{cases} \tag{4.14}$$

5. Asymptotic Optimality of $(\tau^{*n}, \underset{\sim}{\delta}^{*n}, \underset{\sim}{d}_1^{*n}, d_2^{*n})$

Consider the empirical Bayes two-stage selection procedure $(\tau^{*n}, \underset{\sim}{\delta}^{*n}, \underset{\sim}{d}_1^{*n}, \underset{\sim}{d}_2^{*n})$ constructed in Section 4. Let $R(\tau^{*n}, \underset{\sim}{\delta}^{*n}, \underset{\sim}{d}_1^{*n}, \underset{\sim}{d}_2^{*n})$ be the conditional Bayes risk given $(\underset{\sim}{X}_1, \underset{\sim}{X}_2, \ldots, \underset{\sim}{X}_n)$ and $E_n R(\tau^{*n}, \underset{\sim}{\delta}^{*n}, \underset{\sim}{d}_1^{*n}, \underset{\sim}{d}_2^{*n})$ the Bayes risk of the empirical Bayes two-stage selection procedure $(\tau^{*n}, \underset{\sim}{\delta}^{*n}, \underset{\sim}{d}_1^{*n}, \underset{\sim}{d}_2^{*n})$, respectively, where E_n is the expectation taken with respect to $(\underset{\sim}{X}_1, \ldots \underset{\sim}{X}_n)$. Note that $R(\tau^{*n}, \underset{\sim}{\delta}^{*n}, \underset{\sim}{d}_1^{*n}, \underset{\sim}{d}_2^{*n}) - R(\tau^B, \underset{\sim}{\delta}^B, \underset{\sim}{d}_1^B, \underset{\sim}{d}_2^B) \geq 0$ since $(\tau^B, \underset{\sim}{\delta}^B, \underset{\sim}{d}_1^B, \underset{\sim}{d}_2^B)$ is the Bayes two-stage selection procedure. Therefore, we have $E_n R(\tau^{*n}, \underset{\sim}{\delta}^{*n}, \underset{\sim}{d}_1^{*n}, \underset{\sim}{d}_2^{*n}) - R(\tau^B, \underset{\sim}{\delta}^B, \underset{\sim}{d}_1^B, \underset{\sim}{d}_2^B) \geq 0$. We use the nonnegative regret risk $E_n R(\tau^{*n}, \underset{\sim}{\delta}^{*n}, \underset{\sim}{d}_1^{*n}, \underset{\sim}{d}_2^{*n}) - R(\tau^B, \underset{\sim}{\delta}^B, \underset{\sim}{d}_1^B, \underset{\sim}{d}_2^B)$ as a measure of the performance of the empirical Bayes two-stage procedure $(\tau^{*n}, \underset{\sim}{\delta}^{*n}, \underset{\sim}{d}_1^{*n}, \underset{\sim}{d}_2^{*n})$. Following (3.7) – (3.12), we have

$$R(\tau^{*n}, \underset{\sim}{\delta}^{*n}, \underset{\sim}{d}_1^{*n}, \underset{\sim}{d}_2^{*n}) - R(\tau^B, \underset{\sim}{\delta}^B, \underset{\sim}{d}_1^B, \underset{\sim}{d}_2^B) = I_n + II_n + III_n \tag{5.1}$$

where

$$\begin{aligned}
0 \leq I_n &= R(\tau^{*n}, \underset{\sim}{\delta}^{*n}, \underset{\sim}{d}_1^{*n}, \underset{\sim}{d}_2^{*n}) - R(\tau^{*n}, \underset{\sim}{\delta}^{*n}, \underset{\sim}{d}_1^B, \underset{\sim}{d}_2^B) \\
&= \sum_{\underset{\sim}{x} \in \mathcal{X}} \tau^{*n}(\underset{\sim}{x})(d_{1i^*}^B(\underset{\sim}{x})\varphi_{i^*}(x_{i^*}) - d_{i_n^*}^{*n}(\underset{\sim}{x})\varphi_{i_n^*}(x_{i_n^*}))f(\underset{\sim}{x}) \\
&+ \sum_{\underset{\sim}{x} \in \mathcal{X}} (1 - \tau^{*n}(\underset{\sim}{x}))\Bigg\{\sum_{i=1}^{k} \delta_i^{*n}(\underset{\sim}{x})\Bigg(\sum_{y_i=0}^{+\infty}(d_{2i}^{*n}(\underset{\sim}{x}, y_i) - d_{2i}^B(\underset{\sim}{x}, y_i)) \\
&\qquad \times (\theta_0 - \psi_i(x_i + y_i))f_i(y_i|\underset{\sim}{x}, \alpha_i, \mu_i)\Bigg)\Bigg\}f(\underset{\sim}{x}) \\
&= I_{n,1} + I_{n,2},
\end{aligned} \tag{5.2}$$

$$\begin{aligned}
0 \leq II_n &= R(\tau^{*n}, \underset{\sim}{\delta}^{*n}, \underset{\sim}{d}_1^B, \underset{\sim}{d}_1^B) - R(\tau^{*n}, \underset{\sim}{\delta}^B, \underset{\sim}{d}_1^B, \underset{\sim}{d}_2^B) \\
&= \sum_{\underset{\sim}{x} \in \mathcal{X}} (1 - \tau^{*n}(\underset{\sim}{x}))(\delta_{j_n^*}^{*n}(\underset{\sim}{x})T_{j_n^*}(\underset{\sim}{x}) - \delta_{j^*}^B(\underset{\sim}{x})T_{j^*}(\underset{\sim}{x}))f(\underset{\sim}{x})
\end{aligned} \tag{5.3}$$

and

$$\begin{aligned}
0 \leq III_n &= R(\tau^{*n}, \underset{\sim}{\delta}^B, \underset{\sim}{d}_1^B, \underset{\sim}{d}_i^B) - R(\tau^B, \underset{\sim}{\delta}^B, \underset{\sim}{d}_1^B, \underset{\sim}{d}_2^B) \\
&= \sum_{\underset{\sim}{x} \in \mathcal{X}} (\tau^{*n}(\underset{\sim}{x}) - \tau^B(\underset{\sim}{x}))Q(\underset{\sim}{x})f(\underset{\sim}{x}).
\end{aligned} \tag{5.4}$$

5.1 Preliminary Analysis

To investigate the asymptotic optimality of $(\tau^{*n}, \underset{\sim}{\delta}^{*n}, \underset{\sim}{d}_1^{*n}, \underset{\sim}{d}_2^{*n})$, we first do some preliminary analysis. Define for $i = 1, \ldots, k$,

$$L_{1i} = \begin{cases} \max\{N | \frac{\alpha_i\mu_i + m_1}{\alpha_i + m_1 + N} > \theta_0, & N \text{ is a nonnegative integer}\}, \\ -1, & \text{if no such nonnegative integer exists.} \end{cases} \tag{5.5}$$

and

$$L_{2i} = \begin{cases} \max\{N | \frac{\alpha_i\mu_i + m_1 + m_2}{\alpha_i + m_1 + m_2 + N} > \theta_0, & N \text{ is a nonnegative interger}\}, \\ -1, & \text{if no such nonnegative integer exists.} \end{cases} \tag{5.6}$$

From (5.5) and (5.6), we can immediately see that for $1 \leq i \leq k$.

$$\begin{cases} \text{If } L_{1i} \geq 0, & \text{then } \varphi_i(x_i) > \theta_0 \quad \text{for } \ 0 \leq x_i \leq L_{1i}, \\ \text{If } L_{1i} = -1, & \text{then } \varphi_i(x_i) \leq \theta_0 \quad \text{for all } \ x_i = 0, 1, \ldots \end{cases} \tag{5.7}$$

and

$$\begin{cases} \text{If } L_{2i} \geq 0, & \text{then } \psi_i(x_i + y_i) > \theta_0 \quad \text{for } \ 0 \leq (x_i + y_i) \leq L_{2i}; \\ \text{If } L_{2i} = -1, & \text{then } \psi_i(x_i + y_i) \leq \theta_0 \quad \text{for all } \ x_i(y_i) = 0, 1, \ldots \end{cases} \tag{5.8}$$

Note that both L_{1i} and L_{2i} are finite whole numbers given α_i and μ_i, and we always have $L_{2i} \geq L_{1i}$, for $i = 1, \ldots, k$.

For $i = 1, \ldots, k$, let

$$S_{1i} = \begin{cases} \{\underset{\sim}{x} \in \mathcal{X} | 0 \leq x_i \leq L_{1i}\}, & \text{if } \ L_{1i} \geq 0; \\ \phi, & \text{otherwise.} \end{cases} \tag{5.9}$$

$$S_{2i} = \begin{cases} \{(\underset{\sim}{x}, \underset{\sim}{y}) \in \mathcal{X} \times \mathcal{Y} | 0 \leq x_i + y_i \leq L_{2i}\}, & \text{if } \ L_{2i} \geq 0; \\ \phi, & \text{otherwise.} \end{cases} \tag{5.10}$$

Then from (5.7) and (5.9), for each $\underset{\sim}{x} \in S_{1i}, \varphi_i(x_i) - \theta_0 > 0$. Similarly, from (5.8) and (5.10), for $(\underset{\sim}{x}, \underset{\sim}{y}) \in S_{2i}, \psi_i(x_i + y_i) - \theta_0 > 0$.

Also, for $i = 1, \ldots, k$, let

$$S_{3i} = \{\underset{\sim}{x} \in \mathcal{X} | T_i(\underset{\sim}{x}) < 0\}. \tag{5.11}$$

From (3.2) and (3.3), we have, for $i = 1, \ldots, k$,

$$\begin{cases} T_i(\underset{\sim}{x}) = 0, & \text{if } \ x_i > L_{2i}; \\ T_i(\underset{\sim}{x}) < 0, & \text{if } \ x_i \leq L_{2i} \ \text{ and } \ L_{2i} \geq 0. \end{cases} \tag{5.12}$$

Next, for $j = 1, \ldots, k$, let

$$D_{1j} = \{\underset{\sim}{x} \in \mathcal{X} | \theta_0 - \varphi_j(x_j) > 0\} \tag{5.13}$$

and

$$D_{2j} = \{(\underset{\sim}{x}, \underset{\sim}{y}) \in \mathcal{X} \times \mathcal{Y} | \theta_0 - \psi_j(x_j + y_j) > 0\} \tag{5.14}$$

Since $\varphi_j(x_j) \to 0$ as $x_j \to +\infty$, and $\psi_j(x_j + y_j) \to 0$ as $(x_i + y_i) \to +\infty$, we can see that both D_{1j} and D_{2j} are always nonempty.

(I) Analysis for I_n

We have, from (5.2),

$$
\begin{aligned}
I_{n,1} &= \sum_{\underset{\sim}{x} \in \mathcal{X}} \tau^{*n}(\underset{\sim}{x})(d^B_{1i^*}(\underset{\sim}{x})\varphi_{i^*}(x_{i^*}) - d^{*n}_{i^*_n}(\underset{\sim}{x})\varphi_{i^*_n}(x_{i^*_n}))f(\underset{\sim}{x}) \\
&= \sum_{i=1}^{k} \sum_{\underset{\sim}{x} \in \mathcal{X}} \tau^{*n}(\underset{\sim}{x}) I\{i^* = i, i^*_n = 0\}(\varphi_i(x_i) - \theta_0) f(\underset{\sim}{x}) \\
&\quad + \sum_{j=1}^{k} \sum_{\underset{\sim}{x} \in \mathcal{X}} \tau^{*n}(\underset{\sim}{x}) I\{i^* = 0, i^*_n = j\}(\theta_0 - \varphi_j(x_j)) f(\underset{\sim}{x}) \\
&\quad + \sum_{i=1}^{k} \sum_{j=1}^{k} \sum_{\underset{\sim}{x} \in \mathcal{X}} \tau^{*n}(\underset{\sim}{x}) I\{i^* = i, i^*_n = j\}(\varphi_i(x_j) - \varphi_j(x_j)) f(\underset{\sim}{x}) \\
&= I_{n,1,1} + I_{n,1,2} + I_{n,1,3}
\end{aligned}
\tag{5.15}
$$

Let $b_1 = \min_{1 \le i \le k} \min_{\substack{0 \le x_i \le L_{1i} \\ S_{1i} \ne \phi}} \{\varphi_i(x_i) - \theta_0\}$. Note that when $S_{1i} \ne \phi$ and $0 \le x_i \le L_{1i}$, $\varphi_i(x_i) - \theta_0 > 0$. Since L_{1i} is finite, minimum operation in b_1 is taken on a finite set. Therefore, $b_1 > 0$.

Since $0 \le \tau^{*n}(\underset{\sim}{x}) \le 1$ and $|\varphi_i(x_i) - \theta_0| \le 1$ for all $\underset{\sim}{x} \in \mathcal{X}$,

$$
\begin{aligned}
E_n I_{n,1,1} &\le \sum_{i=1}^{k} \sum_{\underset{\sim}{x} \in \mathcal{X}} P_n\{i^* = i, i^*_n = 0\} f(\underset{\sim}{x}) \\
&= \sum_{i=1}^{k} \sum_{\underset{\sim}{x} \in S_{1i}} P_n\{i^* = i, i^*_n = 0\} f(\underset{\sim}{x}) \\
&= \sum_{i=1}^{k} \sum_{\underset{\sim}{x} \in S_{1i}} P_n\{\varphi_i(x_i) \ge \varphi_l(x_l), \forall l \ne i, \varphi_i(x_i) > \theta_0, \\
&\qquad\qquad \text{and } \varphi_{ln}(x_l) \le \theta_0, \forall l \ne 0\} f(\underset{\sim}{x}) \\
&\le \sum_{i=1}^{k} \sum_{\underset{\sim}{x} \in S_{1i}} P_n\{|\varphi_{in}(x_i) - \varphi_i(x_i)| \ge \varphi_i(x_i) - \theta_0\} f(\underset{\sim}{x}) \\
&\le \sum_{i=1}^{k} \sum_{\underset{\sim}{x} \in S_{1i}} P_n\{|\varphi_{in}(x_i) - \varphi_i(x_i)| \ge b_1\} f(\underset{\sim}{x})
\end{aligned}
\tag{5.16}
$$

Denote $b_{2j} = \min_{\underset{\sim}{x} \in D_{1j}} \{\theta_0 - \varphi_j(x_j)\}$ and $b_2 = \min_{1 \le j \le k} b_{2j}$. Since $\varphi_j(x_j)$ is strictly decreasing as x_j increases, we have

$$b_{2j} = \begin{cases} \theta_0 - \varphi_j(L_{1j}+1), & \text{if } \theta_0 > \varphi_j(L_{1j}+1); \\ \theta_0 - \varphi_j(L_{1j}+2), & \text{if } \theta_0 = \varphi_j(L_{1j}+1). \end{cases} \tag{5.17}$$

Therefore, $b_2 = \min_{1 \le j \le k} b_{2j} > 0$.

$$\begin{aligned} E_n I_{n,1,2} &\le \sum_{j=1}^{k} \sum_{\underset{\sim}{x} \in D_{1j}} P_n\{i^* = 0, i_n^* = j\} f(\underset{\sim}{x}) \\ &\le \sum_{j=1}^{k} \sum_{\underset{\sim}{x} \in D_{1j}} P_n\{|\varphi_{jn}(x_j) - \varphi_j(x_j)| \ge \theta_0 - \varphi_j(x_j)\} f(\underset{\sim}{x}) \\ &\le \sum_{j=1}^{k} \sum_{\underset{\sim}{x} \in D_{1j}} P_n\{|\varphi_{jn}(x_j) - \varphi_j(x_j)| \ge b_2\} f(\underset{\sim}{x}) \end{aligned} \tag{5.18}$$

Let $b_3 = \frac{1}{2} \min_{\substack{1 \le i \le k \\ 1 \le j \le k \\ j \ne i}} \min_{\underset{\sim}{x} \in S_{1i} \cap S_{1j}} \{| \varphi_i(x_i) - \varphi_j(x_j)| \ne 0\}$

and

$$b_4 = \tfrac{1}{2} \min_{\substack{1 \le i \le k \\ 1 \le j \le k \\ j \ne i}} \min_{\underset{\sim}{x} \in S_{1i} \cap S_{1j^c}} \{\varphi_i(x_i) - \varphi_j(x_j)\}.$$

We can see that $b_3 > 0$ because the minimum operation is taken on a finite set. As for b_4, since $\varphi_j(x_j)$ is strictly decreasing as x_j increases, and $\min_{\substack{0 \le x_i \le L_{1i} \\ S_{1i} \ne \phi}} \varphi_i(x_i) = \varphi_i(L_{1i}) > \theta_0$. $\max_{\underset{\sim}{x} \in S_{1j}^c} \varphi_j(x_j) = \varphi_j(L_{1j}+1) \le \theta_0$, we have

$$b_4 = \frac{1}{2} \min_{i \ne j} \min_{\substack{S_{1i} \ne \phi \\ 0 \le x_i \le L_{1i}}} \{\varphi_i(x_i) - \varphi_j(L_{1j}+1)\} > 0.$$

Thus,

$$\begin{aligned} E_n I_{n,1,3} &\le \sum_{i=1}^{k} \sum_{j=1}^{k} \sum_{\underset{\sim}{x} \in S_{1i}} P_n\{i^* = i, i_n^* = j\}(\varphi_i(x_i) - \varphi_j(x_j)) f(\underset{\sim}{x}) \\ &= \sum_{i=1}^{k} \sum_{j=1}^{k} \sum_{\underset{\sim}{x} \in (S_{1i} \cap S_{1j})} P_n\{i^* = i, i_n^* = j\}(\varphi_i(x_i) - \varphi_j(x_j)) f(\underset{\sim}{x}) \end{aligned} \tag{5.19}$$

$$+\sum_{i=1}^{k}\sum_{j=1}^{k}\sum_{\underset{\sim}{x}\in(S_{1i}\cap S_{1j}^c)} P_n\{i^*=i, i_n^*=j\}(\varphi_i(x_i)-\varphi_j(x_j))f(\underset{\sim}{x})$$

$$\leq \sum_{i\neq j}\sum_{\underset{\sim}{x}\in(S_{1i}\cap S_{1j})} (P_n\{|\varphi_{in}(x_i)-\varphi_i(x_i)|\geq b_3\}$$
$$+P_n\{|\varphi_{jn}(x_j)-\varphi_j(x_j)|\geq b_3\})f(\underset{\sim}{x})$$

$$\leq \sum_{i\neq j}\sum_{\underset{\sim}{x}\in(S_{1i}\cap S_{1j}^c)} (P_n\{|\varphi_{in}(x_i)-\varphi_i(x_i)|\geq b_4\}$$
$$+P_n\{|\varphi_{jn}(x_j)-\varphi_j(x_j)|\geq b_4\})f(\underset{\sim}{x})$$

Let $b_5=\min_{1\leq k}\min_{\substack{(\underset{\sim}{x},\underset{\sim}{y})\epsilon S_{2i}\\ S_{2i}\neq\phi}}\{\psi_i(x_i+y_i)-\theta_0\}$, $b_{6i}=\min_{(\underset{\sim}{x},\underset{\sim}{y})\epsilon S_{2i}}\{\theta_0-\psi_i(x_i+y_i)\}$, and $b_6=\min_{1\leq i\leq k}\{b_{6i}\}$.

Since for $i=1,\ldots,k, S_{2i}$ has only finite elements, $b_5>0$. As for D_{2i}, since $\psi_i(x_i+y_i)$ is strictly decreasing as (x_i+y_i) increases,

we have

$$b_{6i}=\begin{cases}\theta_0-\psi_i(L_{2i}+1), & \text{if}\quad \theta_0>\psi_i(L_{2i}+1);\\ \theta_0-\psi_i(L_{2i}+2), & \text{if}\quad \theta_0=\psi_i(L_{2i}+1).\end{cases}$$

Therefore, $b_{6i}>0$ for $i=1,\ldots,k$. and $b_6=\min_{1\leq i\leq k} b_{6i}>0$.

$$E_n(I_{n,2})=\sum_{\underset{\sim}{x}\in\mathcal{X}} E_n(1-\tau^{*n}(\underset{\sim}{x}))\bigg(\sum_{i=1}^{k}\delta_i^{*n}(\underset{\sim}{x})(\sum_{y_i=0}^{+\infty}(d_{2i}^{*n}(\underset{\sim}{x},y_i)-d_{2i}^{B}(\underset{\sim}{x},y_i)\bigg)$$
$$\times(\theta_0-\psi_i(x_i+y_i))f_i(y_i|\underset{\sim}{x},\alpha_i,\mu_i)f(\underset{\sim}{x})$$
$$\leq\sum_{\underset{\sim}{x}\in\mathcal{X}}\sum_{i=1}^{k}\sum_{y_i=0}^{+\infty}E_n(d_{2i}^{*n}(\underset{\sim}{x},y_i)-d_{2i}^{B}(\underset{\sim}{x},y_i))(\theta_0-\psi_i(x_i+y_i))$$
$$\times f_i(y_i|\underset{\sim}{x},\alpha_i,\mu_i)f(\underset{\sim}{x})$$
$$\leq\sum_{i=1}^{k}\sum_{(\underset{\sim}{x},\underset{\sim}{y})\epsilon S_{2i}} P_n\{\psi_i(x_i+y_i)\geq\theta_0 \text{ and } \psi_{in}(x_i+y_i)\leq\theta_0\}$$
$$\times f_i(y_i|\underset{\sim}{x},\alpha_i,\mu_i)f(\underset{\sim}{x})$$

$$
\begin{aligned}
&+\sum_{i=1}^{k} \sum_{(\underset{\sim}{x},\underset{\sim}{y}) \epsilon D_{2i}} P_n\{\psi_i(x_i+y_i) \geq \theta_0 \text{ and } \psi_{in}(x_i+y_i) \leq \theta_0\} \\
&\qquad \times f_i(y_i|\underset{\sim}{x}, \alpha_i, \mu_i) f(\underset{\sim}{x}) \\
\leq &\sum_{i=1}^{k} \sum_{(\underset{\sim}{x},\underset{\sim}{y}) \epsilon S_{2i}} P_n\{\psi_i(x_i+y_i) - \psi_{in}(x_i+y_i) > b_5\} \\
&\qquad \times f_i(y_i|\underset{\sim}{x}, \alpha_i, \mu_i) f(\underset{\sim}{x}) \\
&+\sum_{i=1}^{k} \sum_{(\underset{\sim}{x},\underset{\sim}{y}) \epsilon D_{2i}} P_n\{\psi_{in}(x_i+y_i) - \psi_i(x_i+y_i) > b_6\} \\
&\qquad \times f_i(y_i|\underset{\sim}{x}, \alpha_i, \mu_i) f(\underset{\sim}{x})
\end{aligned} \tag{5.20}
$$

(II) Analysis for II_n.

From (5.11) and (5.12), for $i = 1, \ldots, k$, we have

$$
S_{3i} = \{\underset{\sim}{x} \in \mathcal{X} | T_i(\underset{\sim}{x}) < 0\} = \{\underset{\sim}{x} \in X | 0 \leq x_i \leq L_{2i} \text{ and } L_{2i} \geq 0\} \tag{5.21}
$$

$$
\text{Let} \quad b_7 = \frac{1}{2} \min_{\substack{1 \leq i \leq k \\ 1 \leq j \leq k \\ i \neq j}} \min_{\substack{0 \leq x_i \leq L_{2i} \\ 0 \leq x_j \leq L_{2j}}} \{T_j(\underset{\sim}{x}) - T_i(\underset{\sim}{x}) | T_j(\underset{\sim}{x}) \neq T_i(\underset{\sim}{x})\}
$$

and

$$
b_8 = \frac{1}{2} \min_{\substack{1 \leq i \leq k \\ 1 \leq j \leq k \\ i \neq j}} \min_{\substack{0 \leq x_i \leq L_{2i} \\ x_j > L_{2j}}} \{T_j(\underset{\sim}{x}) - T_i(\underset{\sim}{x})\}
$$

We can see that $b_7 > 0$ since the minimum is taken on a finite set. As for b_8, from (5.12), we have $T_j(\underset{\sim}{x}) = 0$ while $T_i(\underset{\sim}{x}) < 0$, therefore,

$$
b_8 = \frac{1}{2} \min_{1 \leq i \leq k} \min_{\substack{0 \leq x_i \leq L_{2i} \\ L_{2i} \geq 0}} \{|T_i(\underset{\sim}{x})|\} > 0
$$

Since $|T_j(\underset{\sim}{x}) - T_i(\underset{\sim}{x})| \leq 1$, we have

$$
\begin{aligned}
E_n(II_n) &\leq \sum_{i=1}^{k}\sum_{j=1}^{k}\sum_{\underset{\sim}{x}\in\mathcal{X}} P_n\{j^* = i, j_n^* = j\}(T_j(\underset{\sim}{x}) - T_i(\underset{\sim}{x}))f(\underset{\sim}{x}) \\
&= \sum_{i=1}^{k}\sum_{j=1}^{k}\sum_{\underset{\sim}{x}\in S_{3i}} P_n\{j^* = i, j_n^* = j\}(T_j(\underset{\sim}{x}) - T_i(\underset{\sim}{x}))f(\underset{\sim}{x}) \\
&= \sum_{i=1}^{k}\sum_{j\neq i}\sum_{\underset{\sim}{x}\in(S_{3i}\cap S_{3j})} P_n\{j^* = i, j_n^* = j\}(T_j(\underset{\sim}{x}) - T_i(\underset{\sim}{x}))f(\underset{\sim}{x}) \\
&\quad + \sum_{i=1}^{k}\sum_{j\neq i}\sum_{\underset{\sim}{x}\epsilon(S_{3i}\cap S_{3j})^c} P_n\{j^* = i, j_n^* = j\}(T_j(\underset{\sim}{x}) - T_i(\underset{\sim}{x}))f(\underset{\sim}{x}) \\
&\leq \sum_{i=1}^{k}\sum_{j\neq i}\sum_{\substack{0\leq x_i\leq L_{2i}\\ 0\leq x_j\leq L_{2i}}} P_n\{T_i(\underset{\sim}{x}) < T_j(\underset{\sim}{x}) \quad \text{and} \quad T_{in}(\underset{\sim}{x}) \geq T_{jn}(\underset{\sim}{x})\}f(\underset{\sim}{x}) \\
&\quad + \sum_{i=1}^{k}\sum_{j\neq i}\sum_{\substack{0\leq x_i\leq L_{2i}\\ x_j> L_{2j}}} P_n\{T_i(\underset{\sim}{x}) < T_j(\underset{\sim}{x}) \quad \text{and} \quad T_{in}(\underset{\sim}{x}) \geq T_{jn}(\underset{\sim}{x})\}f(\underset{\sim}{x}) \\
&\leq \sum_{i=1}^{k}\sum_{j\neq i}\sum_{\substack{0\leq x_i\leq L_{2i}\\ x_j\leq L_{2i}}} P_n\{|T_{in}(\underset{\sim}{x}) - T_i(\underset{\sim}{x})| \geq b_7\} + P_n\{|T_{jn}(\underset{\sim}{x}) - T_j(\underset{\sim}{x})| \geq b_7\}f(\underset{\sim}{x}) \\
&\quad + \sum_{i=1}^{k}\sum_{j\neq i}\sum_{\substack{0\leq x_i\leq L_{2i}\\ x_j> L_{2j}}} P_n\{|T_{in}(\underset{\sim}{x}) - T_i(\underset{\sim}{x})| \geq b_7\} + P_n\{|T_{jn}(\underset{\sim}{x}) - T_j(\underset{\sim}{x})| \geq b_7\}f(\underset{\sim}{x})
\end{aligned} \tag{5.22}
$$

(III) Analysis for III_n

Denote $K(\underset{\sim}{x}) = \{i = 1, \ldots, k | 0 \leq x_i \leq L_{2i}\}$ and $S_4 = \bigcup_{i=1}^{k} S_{3i}$. Also define $b_9 = \min_{\underset{\sim}{x}\epsilon S_4}\{|Q(\underset{\sim}{x})| \neq 0\}$. From (3.5) and (5.12), we know that for $\underset{\sim}{x} \in S_4^c$, $Q(\underset{\sim}{x}) = -c_2$, and for $\underset{\sim}{x} \in S_4$,

$$
\begin{aligned}
Q(\underset{\sim}{x}) &= \sum_{i=0}^{k} d_{1i}^B(\underset{\sim}{x})(\theta_0 - \varphi_i(x_i)) - c_2 - \sum_{i=1}^{k} \delta_i^B(\underset{\sim}{x})T_i(\underset{\sim}{x}) \\
&= \sum_{i\in K(\underset{\sim}{x})} d_{1i}^B(\underset{\sim}{x})(\theta_0 - \varphi_i(x_i)) - c_2 - \sum_{i\in K(\underset{\sim}{x})} \delta_i^B(\underset{\sim}{x})T_i(\underset{\sim}{x}).
\end{aligned} \tag{5.23}
$$

Therefore,

$$b_9 = \min_{\underset{\sim}{x} \epsilon S_4} \{ |Q(\underset{\sim}{x})| \; : \; Q(\underset{\sim}{x}) \neq 0 \} > 0.$$

Since $|Q(\underset{\sim}{x})| \leq 2 + C_2$, we have

$$\begin{aligned}
E_n(III_n) &= \sum_{\underset{\sim}{x} \in \mathcal{X}} E_n(\tau^{*n}(\underset{\sim}{x}) - T^B(\underset{\sim}{x}))\; Q(\underset{\sim}{x}) f(\underset{\sim}{x}) \\
&= \sum_{\underset{\sim}{x} \in S_4^c} E_n(\tau^{*n}(\underset{\sim}{x}) - \tau^B(\underset{\sim}{x}))\; Q(\underset{\sim}{x}) f(\underset{\sim}{x}) \\
&\quad + \sum_{\underset{\sim}{x} \in S_4} E_n(\tau^{*n}(\underset{\sim}{x}) - \tau^B(\underset{\sim}{x}))\; Q(\underset{\sim}{x}) f(\underset{\sim}{x}) \\
&\leq c_2 \sum_{\underset{\sim}{x} \in S_4^c} P_n\{Q_n(\underset{\sim}{x}) > 0 \quad \text{and} \quad Q(\underset{\sim}{x}) = -c_2\} f(\underset{\sim}{x}) \qquad (5.24) \\
&\quad + (2 + c_2) \sum_{\underset{\sim}{x} \in S_4} P_n\{\tau^{*n}(\underset{\sim}{x}) \neq \tau^B(\underset{\sim}{x}), \quad \text{and} \quad Q(\underset{\sim}{x}) \neq 0\} f(\underset{\sim}{x}) \\
&\leq c_2 \sum_{\underset{\sim}{x} \in S_4^c} P_n\{|Q_n(\underset{\sim}{x}) - Q(\underset{\sim}{x})| \geq c_2\} f(\underset{\sim}{x}) \\
&\quad + (2 + c_2) \sum_{\underset{\sim}{x} \in S_4} P_n\{|Q_n(\underset{\sim}{x}) - Q(\underset{\sim}{x})| \geq b_9\} f(\underset{\sim}{x}).
\end{aligned}$$

Therefore, in order to investigate the asymptotic optimality of the empirical Bayes two-stage selection procedure $(\tau^{*n}, \underset{\sim}{\delta}^{*n}, \underset{\sim}{d}_1^{*n}, \underset{\sim}{d}_2^{*n})$, it suffices to study the asymptotic behavior of

$$P_n\{|\varphi_{in}(x_i) - \varphi_i(x_i)| \geq (b_1 \wedge b_2 \wedge b_3 \wedge b_4)\};$$

$$P_n\{|\psi_{in}(x_i + y_i) - \psi_i(x_i + y_i)| \geq (b_5 \wedge b_6)\};$$

$$P_n\{|T_{in}(\underset{\sim}{x}) - T_i(\underset{\sim}{x})| \geq (b_7 \wedge b_8)\};$$

$$P_n\{|Q_n(\underset{\sim}{x}) - Q(\underset{\sim}{x})| \geq (b_9 \wedge c_2)\}.$$

We have

$$\begin{aligned}
&\{|\varphi_{in}(x_i) - \varphi_i(x_i)| \geq (b_1 \wedge b_2 \wedge b_3 \wedge b_4)\} \\
&\subset \{|\varphi_{in}(x_i) - \varphi_i(x_i)\} \geq (b_1 \wedge b_2 \wedge b_3 \wedge b_4), C_{in} > 0\} \cup \{C_{in} \leq 0\}.
\end{aligned} \qquad (5.25)$$

Lemma 5.1 For $\epsilon > 0$, there exists a positive constant $q_{i1}(\epsilon, \alpha_i, \mu_i)$, such that for any $\underset{\sim}{x} \in \mathcal{X}$, when $C_{in} > 0, |\alpha_i - \alpha_{in}| \le q_{i1}(\epsilon, \alpha_i, \mu_i)$ and $|\mu_i - \mu_{in}| \le q_{i1}(\epsilon, \alpha_i, \mu_i)$, we have $|\varphi_{in}(x_i) - \varphi_i(x_i)| \le \epsilon$ for all $\underset{\sim}{x} \in \mathcal{X}$. This property is called equi-continuity. In other words, $\varphi_{in}(x_i)$ is an equi-continuous function of α_i and μ_i.

Proof: Since $\mu_{in} \le \mu_i + |\mu_i - \mu_{in}|$ and $\alpha_{in} \le \alpha_i + |\alpha_i - \alpha_{in}|$ we have

$$|\alpha_{in}\mu_{in} - \alpha_i\mu_i| \le |\alpha_i - \alpha_{in}| \times |\mu_i - \mu_{in}| + \alpha_i|\mu_i - \mu_{in}| + \mu_i|\alpha_i - \alpha_{in}|. \tag{5.26}$$

Therefore,

$$\begin{aligned}
&|\varphi_{in}(x_i) - \varphi_i(x_i)| \\
&= \left| \frac{\alpha_{in}\mu_{in} + m_1}{\alpha_{in} + m_1 + x_i} - \frac{\alpha_i\mu_i + m_1}{\alpha_i + m_1 + x_i} \right| \\
&= \left| \frac{\alpha_i\alpha_{in}(\mu_{in} - \mu_i) + (m_1 + x_i)(\alpha_{in}\mu_{in} - \alpha_i\mu_i) + m_1(\alpha_i - \alpha_{in})}{(\alpha_{in} + m_1 + x_i)(\alpha_i + m_1 + x_i)} \right| \\
&\le \frac{\alpha_i\alpha_{in}|\mu_{in} - \mu_i| + m_1|\alpha_{in} - \alpha_i| + (m_1 + x_i)|\alpha_{in}\mu_{in} - \alpha_i\mu_i|}{(\alpha_{in} + m_1 + x_i)(\alpha_i + m_1 + x_i)} \\
&\le \frac{\alpha_i\alpha_{in}|\mu_{in} - \mu_i| + m_1|\alpha_{in} - \alpha_i| + (m_1 + x_i)|\alpha_{in} - \alpha_i| \cdot |\mu_{in} - \mu_i|}{(\alpha_{in} + m_1 + x_i)(\alpha_i + m_1 + x_i)} \\
&\qquad + \frac{\alpha_i(m_1 + x_i)|\mu_{in} - \mu_i| + \mu_i(m_1 + x_i)|\alpha_{in} - \alpha_i|}{(\alpha_{in} + m_1 + x_i)(\alpha_i + m_1 + x_i)} \qquad (5.27) \\
&= \frac{\alpha_i|\mu_i - \mu_{in}|(\alpha_{in} + m_1 + x_i) + m_1|\alpha_i - \alpha_{in}|}{(\alpha_{in} + m_1 + x_i)(\alpha_i + m_1 + x_i)} \\
&\quad + \frac{(m_1 + x_i)|\alpha_i - \alpha_{in}| \cdot |\mu_i - \mu_{in}| + \mu_i(m_1 + x_i)|\alpha_i - \alpha_{in}|}{(\alpha_{in} + m_1 + x_i)(\alpha_i + m_1 + x_i)} \\
&= \frac{\alpha_i \cdot |\mu_i - \mu_{1i}|}{\alpha_i + m_1 + x_i} + |\alpha_i - \alpha_{in}| \cdot \frac{m_1 + (m_1 + x_i)|\mu_i - \mu_{in}| + (m_1 + x_i)\mu_i}{(\alpha_{in} + m_1 + x_i)(\alpha_i + m_1 + x_i)} \\
&\le \frac{\alpha_i \cdot |\mu_i - \mu_{in}|}{\alpha_i + m_1} + |\alpha_i - \alpha_{in}| \cdot \frac{1 + |\mu_i - \mu_{in}| + (m_1 + x_i)\mu_i}{(\alpha_i + m_1 + x_i)} \\
&\le \frac{\alpha_i \cdot |\mu_i - \mu_{in}|}{\alpha_i + m_1} + \frac{|\alpha_i - \alpha_{in}| \cdot (1 + |\mu_i - \mu_{in}| + \mu_i)}{\alpha_i + m_1}
\end{aligned}$$

From the above we see that for any $\epsilon > 0$, there exists a positive constant $q_{i1}(\epsilon, \alpha_i, \mu_i)$ such that for any $\underset{\sim}{x} \in \mathcal{X}$, when $C_{in} > 0, |\alpha_i - \alpha_{in}| \le q_{i1}(\epsilon, \alpha_i, \mu_i)$ and $|\mu_i - \mu_{in}| \le q_{i1}(\epsilon, \alpha_i, \mu_i)$, we have $|\varphi_{in}(x_i) - \varphi_i(x_i)| \le \epsilon$.

From Lemma 5.1, when $C_{in} > 0$, there exists a positive constant $q_{i1} = q_{i1}((b_1 \wedge b_2 \wedge b_3 \wedge b_4), \alpha_i, \mu_i)$ such that

$$\begin{aligned} &\{|\varphi_{in}(x_i) - \varphi_i(x_i)| \geq (b_1 \wedge b_2 \wedge b_3 \wedge b_4), C_{in} > 0\} \\ &\subset \{|\alpha_{in} - \alpha_i| \geq q_{i1}, C_{i1} > 0\} \cup \{|\mu_{in} - \mu_i| \geq q_{i1}, C_{in} > 0\}. \end{aligned} \tag{5.28}$$

Hence,

$$\begin{aligned} &P_n\{|\varphi_{in}(x_i) - \varphi(x_i)| \geq (b_1 \wedge b_2 \wedge b_3 \wedge b_4)\} \\ &\leq P_n\{|\alpha_{in} - \alpha_i| \geq q_{i1}, C_{in} > 0\} \\ &\quad + P_n\{|\mu_{in} - \mu_i| \geq q_{i1}, C_{in} > 0\} \\ &\quad + P_n\{C_{in} \leq 0\}. \end{aligned} \tag{5.29}$$

Similarly, when $C_{in} > 0$, there exists a positive constant $q_{i2} = q_{i2}((b_5 \wedge b_6), \alpha_i, \mu_i)$ such that

$$\begin{aligned} &\{|\psi_{in}(x_i + y_i) - \psi_i(x_i + y_i)| \geq (b_5 \wedge b_6), C_{in} > 0\} \\ &\subset \{|\alpha_{in} - \alpha_i| \geq q_{i2}, C_{in} > 0\} \cup \{|\mu_{in} - \mu_i| \geq q_{i2}, C_{in} > 0\}, \end{aligned} \tag{5.30}$$

and

$$\begin{aligned} &P_n\{|\psi_{in}(x_i + y_i) - \psi_i(x_i + y_i)| \geq (b_5 \wedge b_6)\} \\ &\leq P_n\{|\alpha_{in} - \alpha_i| \geq q_{i2}, C_{in} > 0\} \\ &\quad + P_n\{|\mu_{in} - \mu_i| \geq q_{i2}, C_{in} > 0\} \\ &\quad + P_n\{C_{in} \leq 0\}. \end{aligned} \tag{5.31}$$

Using the similar approach, we can obtain the following two lemmas.

Lemma 5.2 For any $\epsilon > 0$, there exists a positive constant $q_{i3}(\epsilon, \alpha_i, \mu_i)$ such that when $C_{in} > 0, |\alpha_{in} - \alpha_i| \leq q_{i3}(\epsilon, \alpha_i, \mu_i)$ and $|\mu_{in} - \mu_i| \leq q_{i3}(\epsilon, \alpha_i, \mu_i)$, we have, for all $\underset{\sim}{x} \in \mathcal{X}$,

$$|T_{in}(\underset{\sim}{x}) - T_i(\underset{\sim}{x})| \leq \epsilon. \tag{5.32}$$

Lemma 5.3 For any $\epsilon > 0$, there exists a positive constant $q_4(\epsilon, \underset{\sim}{\alpha}, \underset{\sim}{\mu})$ such that when $C_{in} > 0, |\alpha_{in} - \alpha_i| \leq q_4(\epsilon, \underset{\sim}{\alpha}, \underset{\sim}{\mu})$ and $|\mu_{in} - \mu_i| \leq q_4(\epsilon, \underset{\sim}{\alpha}, \underset{\sim}{\mu})$, we have $|Q_n(\underset{\sim}{x}) - Q(\underset{\sim}{x})| \leq \epsilon$.

From Lemma 5.2 and 5.3, if we let $q_{i3} = q_{i3}((b_7 \wedge b_8), \alpha_i, \mu_i)$ and $q_4 = q_4((b_9 \wedge c_2), \underset{\sim}{\alpha}, \underset{\sim}{\mu})$, we can immediately obtain

$$\begin{aligned} &P_n\{|T_{in}(\underset{\sim}{x}) - T_i(\underset{\sim}{x})| \geq (b_7 \wedge b_8)\} \\ &\leq P_n\{|\alpha_{in} - \alpha_i| \geq q_{i3}, C_{in} > 0\} \\ &\quad + P_n\{|\mu_{in} - \mu_i| \geq q_{i3}, C_{in} > 0\} \\ &\quad + P_n\{C_{in} \leq 0\} \end{aligned} \tag{5.33}$$

and

$$\begin{aligned} &P_n\{|Q_n(\underset{\sim}{x}) - Q(\underset{\sim}{x})| \geq (b_9 \wedge c_2)\} \\ &\leq \sum_{i=1}^{k} P_n\{|\alpha_{in} - \alpha_i| > q_4, C_{in} > 0\} \\ &\quad + \sum_{i=1}^{k} P_n\{|\mu_{in} - \mu_i| > q_4, C_{in} > 0\} \\ &\quad + \sum_{i=1}^{k} P_n\{C_{in} \leq 0\}. \end{aligned} \tag{5.34}$$

From (5.29), (5.31), (5.33) and (5.34), it suffices to investigate the asymptotic behavior of $P_n\{|\alpha_{in} - \alpha_i| > \epsilon, C_{in} > 0\}, P_n\{|\mu_{in} - \mu_i| > \epsilon,\ C_{in} > 0\}$, and $P_n\{C_{in} \leq 0\}$ for each $i = 1, \ldots, k$, where $\epsilon > 0$.

5.2 Lemmas

In this section, we introduce some lemmas which are helpful to investigate the asymptotic behavior. The following lemma is from Hoeffding.

Lemma 5.4 (Hoeffding Inequality) If random variables $Z_1, \ldots, Z_n$ are i.i.d., such that $a \leq Z_i \leq b, i = 1, \ldots, k$, then for any $\epsilon > 0$.

$$P_n\left\{\bar{Z} - \mu \geq \epsilon\right\} \leq \exp(-2n\epsilon^2/(b-a)^2), \tag{5.35}$$

where

$$\bar{Z} = \frac{1}{n}\sum_{i=1}^{n} Z_i \quad \text{and} \quad \mu = E(\overline{Z}).$$

Proof: This lemma is a special case of Theorem 1 of Hoeffding (1963).

Lemma 5.5 Let μ_i, ν_i, μ_{in} and ν_{in} be defined as in Section 4, respectively. Then for any $\epsilon > 0$,

a) $P_n\{\mu_{in} - \mu_i \leq -\epsilon\} \leq O(\exp(-2n\epsilon^2))$,

b) $P_n\{\mu_{in} - \mu_i \geq \epsilon\} \leq O(\exp(-2n\epsilon^2))$,

c) $P_n\{\nu_{in}\nu_i \leq -\epsilon\} \leq O(\exp(-2n\epsilon^2))$,

d) $P_n\{\nu_{in} - \nu_i \geq \epsilon\} \leq O(\exp(-2n\epsilon^2))$.

Proof: We only prove part b. The proof of the other inequalities are analogous.

For each $i = 1, \ldots, k$, and $j = 1, \ldots, n$, denote $Z_{ij} = \binom{m_1+X_{ij}-2}{m_1-2} / \binom{m_1+X_{ij}-1}{m_1-1}$. Then we can see that $Z_{ij} = (m_1 - 1)/(m_1 + X_{ij} - 1)$, from which we have $0 \leq Z_{ij} \leq 1$. For $i = 1, \ldots, k, (Z_{i1}, Z_{i2}, \ldots, Z_{in})$ are i.i.d. Therefore, from Lemma 5.4, we obtain the proof of part b.

Lemma 5.6 Let C_i, D_i, C_{in} and D_{in} be defined as in Section 4, respectively. Then for any $\epsilon > 0$, we have

a) $P_n\{C_{in} - C_i \leq -\epsilon\} \leq O(\exp(-n\epsilon^2/8))$,

b) $P_n\{C_{in} - C_i \geq \epsilon\} \leq O(\exp(-n\epsilon^2/8))$,

c) $P_n\{D_{in} - D_i \leq -\epsilon\} \leq O(\exp(-n\epsilon^2/2))$,

d) $P_n\{D_{in} - D_i \geq \epsilon\} \leq O(\exp(-n\epsilon^2/2))$.

Proof: The techniques used to prove the four inequalities are similar. So we only give the proof of part a. We have

$$\begin{aligned} & P_n\{C_{in} - C_i \leq -\epsilon\} \\ &= P_n\{(\nu_{in} - \nu_i) + (1 - \mu_{in})^2 - (1 - \mu_i)^2 \leq -\epsilon\} \qquad (5.36) \\ &\leq P_n\{\nu_{in} - \nu_i \leq -\epsilon/2\} + P_n\{(\mu_{in} - \mu_i)(2 - \mu_{in} - \mu_i) \leq -\epsilon/2\}. \end{aligned}$$

Since $0 \leq \mu_{in} \leq 1,\ 0 < \mu_i < 1,\ 0 < (2 - \mu_{in} - \mu_i) < 2$,

$$P_n\{(\mu_{in} - \mu_i)(2 - \mu_{in} - \mu_i) \leq -\epsilon/2\} \leq P_n\{(\mu_{in} - \mu_i) \leq -\epsilon/4\}.$$

Thus,

$$\begin{aligned}
&P_n\{C_{in} - C_i \le -\epsilon\} \\
&\le P_n\{(\mu_i - \mu_i) \le -\epsilon/4\} + P_n\{(\nu_{in} - \nu_i) \le -\epsilon/2\} \\
&\le O(\exp(-n\epsilon^2/8)).
\end{aligned} \tag{5.37}$$

Corollary 5.7 From Lemma 5.6, we have

$$P_n\{C_{in} \le 0\} = O(\exp(-nC_i^2/8)). \tag{5.38}$$

Lemma 5.8 For $\epsilon > 0$,

$$P_n\{|\alpha_{in} - \alpha_i| > \epsilon, C_{in} > 0\} = O\Big(\exp(-\frac{n\epsilon^2 C_i^2}{8}\min(1, \frac{1}{4(\alpha_i + \epsilon)^2}))\Big). \tag{5.39}$$

Proof:

$$\begin{aligned}
&P_n\{|\alpha_{in} - \alpha_i| > \epsilon,\ C_{in} > 0\} \\
&= P_n\{\alpha_{in} - \alpha_i < -\epsilon,\ C_{in} > 0\} + P_n\{\alpha_{in} - \alpha_i > \epsilon, C_{in} > 0\},
\end{aligned} \tag{5.40}$$

where $P_n\{\alpha_{in} - \alpha_i < -\epsilon, C_{in} > 0\} = 0$, if $\alpha_i - \epsilon \le 0$. When $\alpha_i - \epsilon > 0$, by the Bonferroni inequality, we have

$$\begin{aligned}
&P_n\{\alpha_{in} - \alpha_i < -\epsilon,\ C_{in} > 0\} \\
&\le P_n\{(D_{in} - D_i) - (C_{in} - C_i)(\alpha_i - \epsilon) < -\epsilon C_i\} \\
&\le P_n\{D_{in} - D_i \le -\epsilon C_i/2\} + P_n\{C_{in} - C_i > \epsilon C_i/(2(\alpha_i - \epsilon))\} \\
&= O(\exp(-n\epsilon^2 C_i^2/8)) + O(\exp(-n\epsilon^2 C_i^2/(32(\alpha_i - \epsilon)^2)) \\
&= O\Big(\exp(-\frac{n\epsilon^2 C_i^2}{8}\min(1, \frac{1}{4(\alpha_i - \epsilon)^2}))\Big).
\end{aligned} \tag{5.41}$$

Similarly, we have

$$\begin{aligned}
&P_n\{\alpha_{in} - \alpha_i > \epsilon, C_{in} > 0\} \\
&\le P_n\{(D_{in} - D_i) - (C_{in} - C_i)(\alpha_i + \epsilon) > \epsilon C_i\} \\
&\le P_n\{D_{in} - D_i > \epsilon C_i/2\} + P_n\{C_{in} - C_i < -\epsilon C_i/(2(\alpha_i + \epsilon))\} \\
&= O\Big(\exp(-\frac{n\epsilon^2 C_i^2}{8}\min(1, \frac{1}{4(\alpha_i + \epsilon)^2}))\Big).
\end{aligned} \tag{5.42}$$

This completes the proof of Lemma 5.8.

Define $e_1 = \min_{1\le i\le k} q_{i1}$, $e_2 = \min_{1\le i\le k} q_{i2}$, $e_3 = \min_{1\le i\le k} q_{i3}$, $e_4 = \frac{1}{2}\min_{1\le i\le k}\{\alpha_i\}$, and $e^* = \min_{1\le j\le 4}\{e_j\}$. We see that $e^* > 0$.

In the following theorem we prove the asymptotic optimality of the empirical Bayes two-stage procedure $(\tau^{*n}, \underset{\sim}{\delta}^{*n}, \underset{\sim}{d}_1^{*n}, \underset{\sim}{d}_2^{*})$.

Theorem 5.1 Let $\{(\tau^{*n}, \underset{\sim}{\delta}^{*n}, \underset{\sim}{d}_1^{*n}, \underset{\sim}{d}_2^{*n})\}_{n=1}^{\infty}$ be the sequence of empirical Bayes two-stage selection procedures constructed in Section 4. Then,

$$E_n R(\tau^{*n}, \underset{\sim}{\delta}^{*n}, \underset{\sim}{d}_1^{*n}, \underset{\sim}{d}_2^{*n}) - R(\tau^B, \underset{\sim}{\delta}^B, \underset{\sim}{d}_1^B, \underset{\sim}{d}_2^B) = O(\exp(-c^* n)) \tag{5.43}$$

where $c^* = \min\left(\frac{e^{*2}}{8}, \min_{1\le i\le k}\frac{C_i^2}{8}, \min_{1\le i\le k}(\frac{e^{*2}C_i^2}{8}\min(1, \frac{1}{9\alpha_i^2}))\right) > 0.$

Proof: By the definition of e^* and C^*, from Lemma (5.5) to (5.8), we have

$$\begin{aligned}
&P_n\{|\varphi_{in}(x_i) - \varphi_i(x_i)| \ge (b_1 \wedge b_2 \wedge b_3 \wedge b_4)\} = O(\exp(-c^* n)),\\
&P_n\{|\psi_{in}(x_i + y_i) - \psi_i(x_i + y_i)| \ge (b_5 \wedge b_6)\} = O(\exp(-c^* n)),\\
&P_n\{|T_{in}(\underset{\sim}{x}) - T_i(\underset{\sim}{x})| \ge (b_7 \wedge b_8)\} = O(\exp(-c^* n)),\\
&P_n\{|Q_n(\underset{\sim}{x}) - Q(\underset{\sim}{x})| \ge (b_9 \wedge c_2)\} = O(\exp(-c^* n)).
\end{aligned} \tag{5.44}$$

We obtain

$$\begin{aligned}
&E_n R(\tau^{*n}, \underset{\sim}{\delta}^{*n}, \underset{\sim}{d}_1^{*n}, \underset{\sim}{d}_2^{*n}) - R(\tau^B, \underset{\sim}{\delta}^B, \underset{\sim}{d}_1^B, \underset{\sim}{d}_2^B)\\
&\le O\exp(-c^* n))k^2 \sum_{\underset{\sim}{x}\in\mathcal{X}} f(\underset{\sim}{x})\\
&\quad + O(\exp(-c^* n)) \sum_{\underset{\sim}{x}\in\mathcal{X}} \sum_{i=1}^{k} \left[\sum_{y_1=0}^{+\infty} f_i(y_i|\underset{\sim}{x}, \alpha_i, \mu_i)\right] f(\underset{\sim}{x}),\\
&\quad + O(\exp(-c^* n))k^2 \sum_{\underset{\sim}{x}\in\mathcal{X}} f(\underset{\sim}{x})\\
&\quad + O(\exp(-c^* n)) \sum_{\underset{\sim}{x}\in\mathcal{X}} (2 + c_2) f(\underset{\sim}{x})\\
&= O(\exp(-c^* n)).
\end{aligned} \tag{5.45}$$

This completes the proof of the theorem.

6. Simulation Study

A simulation study was carried out to investigate the performance of the proposed empirical Bayes two-stage selection procedures for small to moderate values of n.

For given past observations $(\underset{\sim}{X}_j = (X_{1j}, \ldots, X_{kj}), j = 1, \ldots, n), R(\tau^n, \underset{\sim}{\delta}^n, \underset{\sim}{d}_1^n, \underset{\sim}{d}_1^n)$ is the associated conditional Bayes risk of the proposed selection procedure $(\tau^n, \underset{\sim}{\delta}_n, \underset{\sim}{d}_1^n, \underset{\sim}{d}_2^n)$. Then we use $R(\tau^n, \underset{\sim}{\delta}^n, \underset{\sim}{d}_1^n, \underset{\sim}{d}_2^n) - R(\tau^B, \underset{\sim}{\delta}^B, \underset{\sim}{d}_1^B, \underset{\sim}{d}_2^B)$ as an estimator of the difference $E_n R(\tau^n, \underset{\sim}{\delta}^n, \underset{\sim}{d}_1^n, \underset{\sim}{\delta}_2^n) - R(\tau^B, \underset{\sim}{\delta}^B, \underset{\sim}{d}_1^B, \underset{\sim}{d}_2^B)$.

We consider the following case in which $k = 3$. That is, there are 3 Bernoulli treatments π_1, π_2 and π_3, and we would like to make a two-stage selection using the inverse binomial sampling scheme proposed in Section 4.

The scheme of the simulation is described as follows:

(1) For each n and for each $i = 1, 2, 3$, generate independent random variables $X_{i1}, X_{i2}, \ldots, X_{in}$ as follows:

for $j = 1, \ldots, n$,
(a) first generate Θ_{ij} from a Beta distribution with density $h_i(\theta_i|\alpha_i, \mu_i)$
(b) then generate X_{ij} from a negative binomial $\text{NB}(m_1, \theta_{ij})$ distribution.

(2) Based on the past observations $(\underset{\sim}{X}_j, j = 1, \ldots, n)$ and the present observations $\underset{\sim}{X} = (X_1, \ldots, X_k)$ and $\underset{\sim}{Y} = (Y_1, \ldots, Y_k)$, we construct the empirical Bayes two-stage selection procedure $(\tau^{*n}, \underset{\sim}{\delta}^{*n}, \underset{\sim}{d}_1^{*n}, \underset{\sim}{d}_2^{*n})$ and compute the conditional difference

$$D(n) = R(\tau^{*n}, \underset{\sim}{\delta}^{*n}, \underset{\sim}{d}_1^{*n}, \underset{\sim}{d}_2^{*n}) - R(\tau^B, \underset{\sim}{\delta}^B, \underset{\sim}{d}_1^B, \underset{\sim}{d}_2^B).$$

(3) Repeat step (1) and (2) 5000 times. The average of the conditional difference on the 5000 repetitions which is denoted by $\overline{D}(n)$, is used as an estimator of the difference $E_n R(\tau^{*n}, \underset{\sim}{\delta}^{*n}, \underset{\sim}{d}_1^{*n}, \underset{\sim}{d}_2^{*n}) - R(\tau^B, \underset{\sim}{\delta}^B, \underset{\sim}{d}_1^B, \underset{\sim}{d}_2^B)$. The estimated standard error, denoted by $SE(\overline{D}(n))$, is also computed.

Tables (1) and (2) give the results of this simulation study on the performance of the proposed empirical Bayes two-stage selection procedures. In both cases we choose

$\theta_0 = 0.7, c_1 = c_2 = 0.05$. $m_1 = m_2 = 20$, and $\alpha_1 = \alpha_2 = \alpha_3 = 3$. Furthermore, we use $\mu_1 = \mu_2 = \mu_3 = 0.6$ in Table 1 and $\mu_1 = 0.67$ $\mu_2 = 0.69$ and $\mu_3 = 0.71$ in Table 2.

From these results, we see that $\overline{D}(n)$ decreases to zero very rapidly in both cases, which coincides with theorem 5.1 that the convergence rate is $O(\exp(-c^*n))$, $c^* > 0$.

Table 1
Performance of the selection rule
when $\mu_1 = \mu_2 = \mu_3 = 0.6$

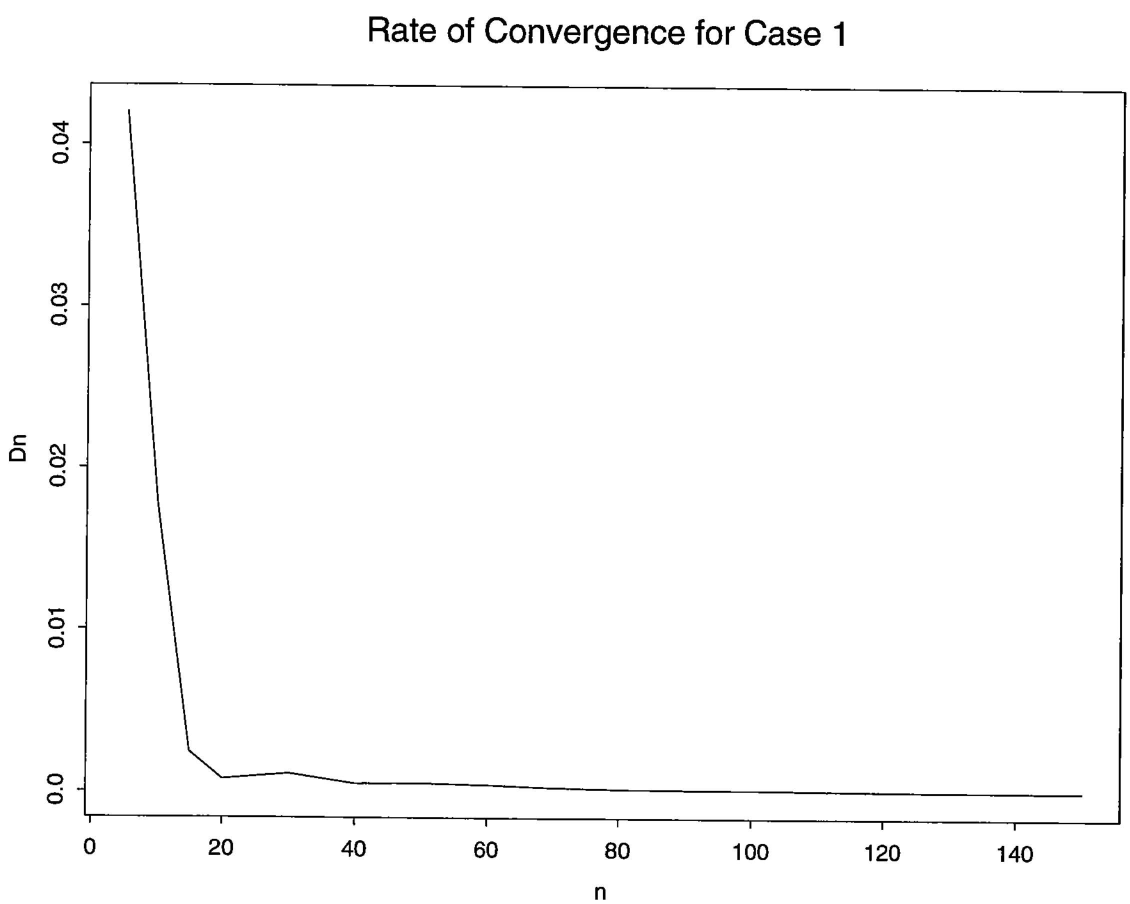
n	$\overline{D}(n)$	$SE(\overline{D}(n))$
5	0.04200000	0.000456063
10	0.01767500	0.000045031
15	0.00244500	0.000003197
20	0.00076700	0.000001487
30	0.00110000	0.000003565
40	0.00049980	0.000000801
50	0.00051410	0.000000265
60	0.00042840	0.000000943
70	0.00029078	0.000000101
80	0.00019035	0.000000127
90	0.00016365	0.000000081
100	0.00016000	0.000000037
125	0.00009243	0.000000012
150	0.00008065	0.000000009

Table 2
Performance of the selection rule
when $\mu_1 = 0.67,\ \mu_2 = 0.69,\ \mu_3 = 0.71$

n	$\overline{D}(n)$	$SE(\overline{D}(n))$
5	0.0576700	0.000332773
10	0.0124585	0.000080474
15	0.0075045	0.000082546
20	0.0024041	0.000001701
30	0.0005337	0.000000883
40	0.0008987	0.000000967
50	0.0009491	0.000003398
60	0.0004567	0.000000698
70	0.0004036	0.000001782
80	0.0003893	0.000000979
90	0.0001930	0.000000135
100	0.0001525	0.000000051
125	0.0000012	0.000000011
150	0.0000000	0.000000000

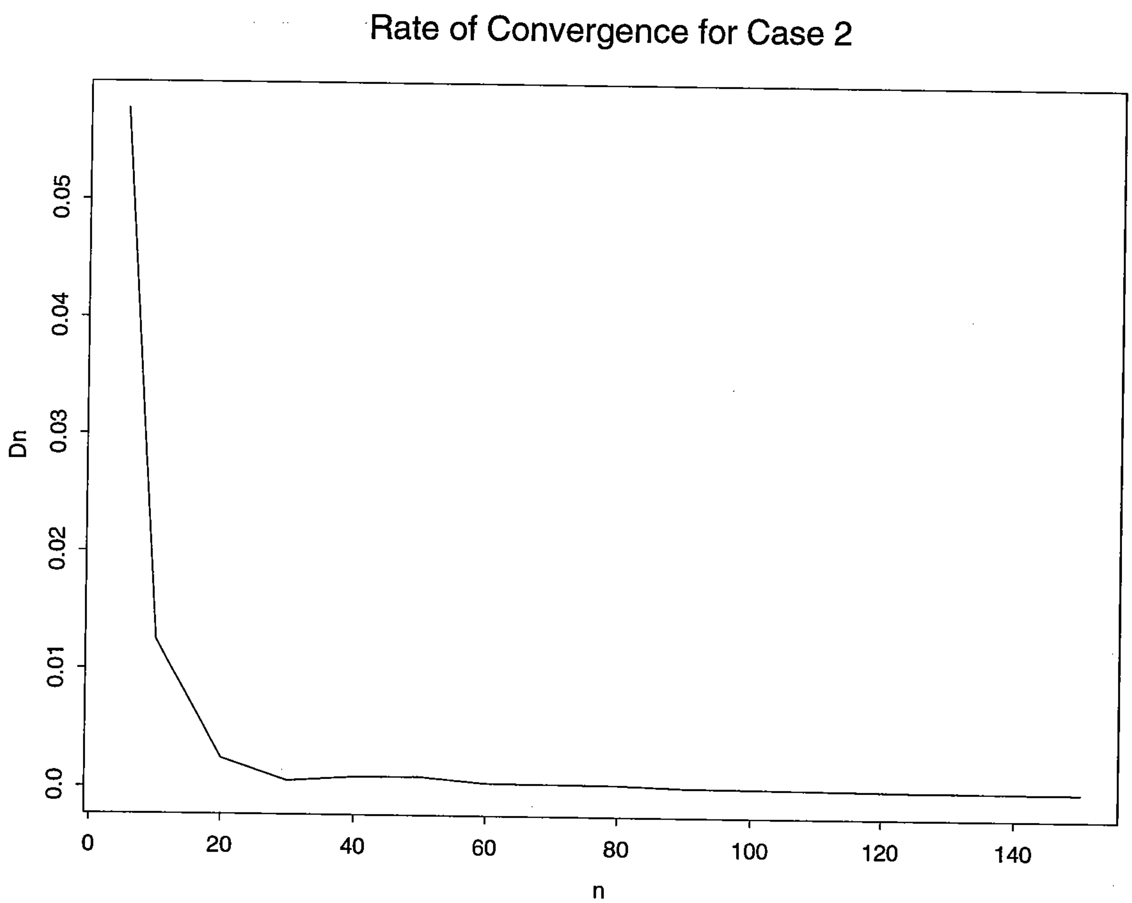
Graph for Table 1

$\mu_1 = \mu_2 = \mu_3 = 0.6$



Graph for Table 2

$\mu_1 = 0.67,\ \mu_2 = 0.69,\ \mu_3 = 0.71$



References

Bechhofer, R., Santner, T. and Goldsman, D. (1996). *Design and Analysis of Experiments for Statistical Selection, Screening, and Multiple Comparisons.* John Wiley.

Gupta, S. S. and Liang, T. (1986). Empirical Bayes rules for selecting good binomial populations. In *Adaptive Procedures and Related Topics* (Ed. J. Van Ryzin), IMS Lecture Notes-Monograph Series, Volume 8, 110-128, Hayward, California.

Gupta, S. S. and Liang, T. (1988). Empirical Bayes rules for selecting the best binomial populations. Statistical Decision Theory and Related Topics IV, Springer-Verlag, Vol. 1, 213-224.

Gupta, S. S. and Liang, T. (1989). Selecting the best binomial population: parametric empirical Bayes approach. *J. Statist. Plann. Inference*, **23**, 21-31.

Gupta, S. S., Liang, T. and Rau, R. B. (1994). Empirical Bayes two-stage procedures for selecting the best Bernoulli population compared with a control. Statistical Decision and Related Topics V. Springer-Verlag, Berlin, New York, 277-292.

Gupta, S. S., Liang, T. and Rau, R. B. (1995). Empirical Bayes two-stage procedures for selecting the best normal population compared with a control. Symposium Gaussiana: Proceedings of the 2nd Gauss Symposium Conference B; *Statistical Science*, 171-196.

Gupta, S. S. and Panchapakesan, S. (1979). *Multiple Decision Procedures: Theory and Methodology of Selecting and Ranking Populations.* New York: John Wiley.

Gupta, S. S. and Panchapakesan, S. (1996). Design of experiments with selection and ranking goals. *Handbook of Statistics*, Vol. 13, 555-586.

Hoeffding, W. (1963). Probabilities for sums of bounded random variables. *J. Amer. Statist. Assoc.*, **58**, 13-30.

Sanchez, S. M. (1987). A modified least-failures sampling procedure for Bernoulli subset selection. *Commun. Statist. Theory and Methods*, **16**, 3609-3629.

Tamhane, A. C. (1980). Selecting the better Bernoulli treatment using a matched sample design. *J. Royd Statist. Soc. B.*, **42**, 26-30.

Yang, H. M. (1989). On selecting the treatment with the largest probability of survival. *Commun. Statist. Theory and Methods*, **18**, 639-659.

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
03-19-98

3. REPORT TYPE AND DATES COVERED
Final March 1998

4. TITLE AND SUBTITLE
Empirical Bayes Two-Stage Selection Procedures for Selecting the Best Bernoulli Treatment Using the Inverse Binomial Sampling

5. FUNDING NUMBERS
DAAH04-95-01-0165

6. AUTHOR(S)
Shanti S. Gupta, Ta Chen Liang and Xun Lin

7. PERFORMING ORGANIZATION NAMES(S) AND ADDRESS(ES)
Purdue University
1399 Mathematical Sciences Building
West Lafayette, IN 47907-1399

8. PERFORMING ORGANIZATION REPORT NUMBER

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*)

In this paper we investigate the problem of selecting the treatment with the largest probability of success from $k(\geq 2)$ independent Bernoulli treatments. The selected treatment must also be better than a given control. We employ the empirical Bayes approach and develop a two-stage selection procedure. We prove that the proposed selection rule is asymptotically optimal at the rate of convergence of order $O(\exp(-c^*n))$, for some positive constant c^*, where n is the number of the historical data at hand. We also carry out a simulation study to investigate the performance of the proposed empirical Bayes selection procedure for small to moderate values of n. The simulated results are provided in the paper.

14. SUBJECT TERMS
Asymptotically optimal; empirical Bayes; two-stage selection procedure; rate of convergence; inverse binomial sampling

15. NUMBER OF PAGES
33

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 Enclosure 1 Standard Form 298 (Rev. 2-89) Prescribed by ANSI Std. 239-18