ON EMPIRICAL BAYES SELECTION RULES FOR SAMPLING INSPECTION *

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

Technical Report #91-35C

Department of Statistics
Purdue University

July 1991
Revised November 1992

* This research was supported in part by NSF Grants DMS–8923071 and DMS–8717799 at Purdue University.

⋆

ON EMPIRICAL BAYES SELECTION RULES FOR SAMPLING INSPECTION *

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

Abstract

We study the problem of acceptance sampling for k independent hypergeometric populations, say $\pi_i = \pi(M_i, m_i, d_i)$, $i = 1, \ldots, k$, where M_i is the number of items in population π_i, d_i is the number of defectives in π_i and m_i is the number of items taken from π_i for inspection. Let d_{i0} be an integer such that $0 < d_{i0} < M_i$, which is used to evaluate the quality of population π_i. Population π_i is said to be good and is acceptable if $d_i < d_{i0}$ and to be bad otherwise. Our goal is to select all good populations and to exclude all bad populations.

This selection problem is formulated through the empirical Bayes approach. Two empirical Bayes selection rules, called as parametric empirical Bayes and hierarchical empirical Bayes, are studied. The asymptotic optimality of the proposed empirical Bayes selection rules are investigated. It is shown that for each empirical Bayes selection rule, the corresponding Bayes risk tends to the minimum Bayes risk with a rate of convergence of order $O(\exp(-ck + \ln k))$ for some positive constant c, where the value of c varies depending on the rule. A simulation study is also carried out to investigate the performance of the empirical Bayes selection rules for small to moderate k case.

Short Title: Empirical Bayes Sampling Inspection

AMS 1980 Subject Classification: 62C12

Keywords and Phrases: Asymptotic optimality; empirical Bayes selection rule; hierarchical empirical Bayes; parametric empirical Bayes; rate of convergence; sampling inspection.

---

* This research was supported in part by NSF Grants DMS–8923071 and DMS–8717799 at Purdue University.

1. Introduction

The hypergeometric distribution arises in sampling without replacement from a finite population. Consider a finite population, say, a batch of M items, which is inspected for defectives. One takes a sample of size m without replacement from the population. Let X denote the random number of defectives in the sample. Also, let d denote the number of defectives in the population. Then, the random variable X has a probability function

$$f(x|d) = \binom{d}{x}\binom{M-d}{m-x} \bigg/ \binom{M}{m}, \tag{1.1}$$

where $\max(0, d+m-M) \leq x \leq \min(m, d)$. Such a finite population with the probability model (1.1) is denoted by $\pi(M, m, d)$.

The problem of acceptance sampling has been studied by several authors for a long time, e.g., see Booth and Smith (1976), and Craig and Bland (1981), among the many authors. In the area of ranking and selection, Bartlett and Govindarajula (1970) have studied the problem of selecting the best population from among $k(\geq 2)$ independent hypergeometric populations via the subset selection approach. The reader is referred to Gupta and Panchapakesan (1979) for detailed discussions on the research area of ranking and selection. Also, see Gupta and Liang (1991) for recent developments in this research field.

In this paper, we consider the problem of acceptance sampling for k independent hypergeometric populations, say, $\pi_i = \pi(M_i, m_i, d_i)$, $i = 1, \ldots, k$. Let d_{i0} be a positive integer such that $0 < d_{i0} < M_i$, $i = 1, \ldots, k$. In general, d_{i0} is a given number used as a standard to evaluate the quality of the population π_i. Population π_i is said to be a good population if $d_i < d_{i0}$, and said to be bad otherwise. Our goal is to select all good populations and to exclude all bad populations.

In this paper, it is assumed that there is a prior distribution depending on some unknown hyperparameter(s) on the parameter space of the interest. The empirical Bayes approach is employed here. We combine information from each of the k populations and use them to estimate the unknown hyperparameter(s) of the prior distribution. Then, we do the typical Bayesian analysis based on the estimated prior distribution, which is obtained by substituting the unknown hyperparameter(s) by the corresponding estimator(s). Two empirical Bayes selection rules, called parametric empirical Bayes and hierarchical empirical Bayes selection rules, are studied, respectively, according to the corresponding statistical models which will be described later. When k, the number of the populations involved in the selection problem, is large, the asymptotic optimality of the proposed empirical Bayes selection rules is investigated. It is shown that in each case, the Bayes risk of the concerned empirical Bayes selection rule tends to the minimum Bayes risk with a rate of convergence of order $O(\exp(-ck + \ln k))$ for some positive constant c, where the value of c varies depending on the rules. Finally, a simulation study is carried out to investigate the performance of the empirical Bayes selection rules for small to moderate values of k.

2. Formulation of the Selection Problem

Let X_i denote a random variable arising from the population $\pi_i = \pi(M_i, m_i, d_i)$. Conditional on d_i, X_i has a hypergeometric distribution with probability function

$$f_i(x|d_i) = \binom{d_i}{x}\binom{M_i - d_i}{m_i - x} \Big/ \binom{M_i}{m_i}, \tag{2.1}$$

where $\max(0, d_i + m_i - M_i) \le x \le \min(m_i, d_i)$. It is assumed that $X_1, \ldots, X_k$ are mutually independent so that $(X_1, \ldots, X_k)$ has a joint probability function

$$f(\underset{\sim}{x}|\underset{\sim}{d}) = \prod_{i=1}^{k} f_i(x_i|d_i), \;\; \underset{\sim}{x} = (x_1, \ldots, x_k), \;\; \underset{\sim}{d} = (d_1, \ldots, d_k).$$

It is also assumed that for each i, d_i follows a binomial $B(M_i, \theta_i)$ distribution and that $d_1, \ldots, d_k$ are mutually independent. That is, d_i has the (prior) probability function

$$g_i(d_i|\theta_i) = \binom{M_i}{d_i}\theta_i^{d_i}(1-\theta_i)^{M_i - d_i}, \;\; d_i = 0, 1, \ldots, M_i,$$

and θ_i is the probability that any item in the population π_i will be defective. It follows that given θ_i, X_i has a marginal probability function

$$\begin{aligned} f_i(x|\theta_i) &= \sum_{d=x}^{M_i - m_i + x} f_i(x|d) g_i(d|\theta_i) \\ &= \binom{m_i}{x}\theta_i^x (1-\theta_i)^{m_i - x}, \;\; x = 0, 1, \ldots, m_i. \end{aligned}$$

Also, $X_1, \ldots, X_k$ are marginally mutually independent.

Let d_{i0} denote a positive integer such that $0 < d_{i0} < M_i$. The d_{i0} will be used to assess whether population π_i should be accepted or not. Population π_i is said to be good if $d_i < d_{i0}$, and to be bad otherwise. Our goal is to select all good populations and to exclude all bad populations.

Let $\Omega = \{\underset{\sim}{d} = (d_1, \ldots, d_k) | 0 \leq d_i \leq M_i,\ i = 1, \ldots, k\}$ denote the parameter space and let $\mathcal{A} = \{\underset{\sim}{a} = (a_1, \ldots, a_k) | a_i = 0, 1;\ i = 1, \ldots, k\}$ denote the action space. When action $\underset{\sim}{a}$ is taken, it means that population π_i is selected as a good population if $a_i = 1$, and excluded as a bad one if $a_i = 0$. For the parameter $\underset{\sim}{d}$ and the action $\underset{\sim}{a}$, the loss function is defined as follows:

$$\begin{aligned} L(\underset{\sim}{d}, \underset{\sim}{a}) &= \sum_{i=1}^{k} (1 - a_i) L_{i0}(d_i) + \sum_{i=1}^{k} a_i L_{i1}(d_i), \\ &= \sum_{i=1}^{k} L_{i0}(d_i) + \sum_{i=1}^{k} a_i [L_{i1}(d_i) - L_{i0}(d_i)] \end{aligned} \tag{2.2}$$

where $L_{i0}(d_i)$ and $L_{i1}(d_i)$ are bounded functions and satisfy

$$L_{i0}(d_i) \begin{cases} = 0 & \text{if } d_i \geq d_{i0}, \\ > 0, & \text{and nonincreasing in } d_i \text{ for } d_i < d_{i0}; \end{cases}$$

$$L_{i1}(d_i) \begin{cases} = 0 & \text{if } d_i \leq d_{i0}, \\ > 0, & \text{and nondecreasing in } d_i \text{ for } d_i > d_{i0}. \end{cases}$$

In (2.2), the first summation is the loss due to not selecting some good populations, while the second summation is the loss due to including some bad populations in the selected set.

Let $\mathcal{X}$ be the sample space generated by $(X_1, \ldots, X_k)$. A selection rule $\underset{\sim}{\delta}_k = (\delta_{k1}, \ldots, \delta_{kk})$ is defined to be a mapping from the sample space $\mathcal{X}$ to the product space $[0, 1]^k$, such that for each $\underset{\sim}{x} \in \mathcal{X}$, $\delta_{ki}(\underset{\sim}{x})$ is the probability of selecting population π_i as a good population, $i = 1, \ldots, k$.

In this paper, the following two cases will be considered.

Case 1. It is assumed that $\theta_1 = \theta_2 = \ldots = \theta_k = \theta$, and the value of the common parameter θ is fixed, but unknown.

Case 2. It is assumed that for each $i = 1, \ldots, k$, the parameter θ_i is a realization of a random variable Θ_i; and $\Theta_1, \ldots, \Theta_k$ are iid, having a beta distribution Beta $(\alpha\mu, \alpha(1-\mu))$ with probability density function $h(\theta|\alpha, \mu)$

$$h(\theta|\alpha,\mu) = \frac{\Gamma(\alpha)}{\Gamma(\alpha\mu)\Gamma(\alpha(1-\mu))}\,\theta^{\alpha\mu-1}\,(1-\theta)^{\alpha(1-\mu)-1},\ 0<\theta<1,$$

where $0 < \mu < 1$, $\alpha > 0$, and the values of both the parameters α and μ are fixed but unknown.

3. Bayes Selection Rules

In order to construct the empirical Bayes selection rules, as a first step, we derive the Bayes selection rule for each case. Also, in the following, we assume that $M_i < M^*$ for all $i = 1, \ldots, k$, where the upper bound M^* is independent of k for all k, and $\max(2, d_{i0}) \le m_i$ for all $i = 1, \ldots, k$.

3.1. Case 1. A Bayes Selection Rule Relative to θ

Consider the unknown hyperparameter θ as fixed, we define, for each $i = 1, \ldots, k$,

$$\begin{cases} C_i(\theta) = E[L_{i0}(d_i)|\theta] = \sum\limits_{d_i=0}^{d_{i0}-1} L_{i0}(d_i)g_i(d_i|\theta), \\ H_i(x_i,\theta) = E[L_{i1}(d_i) - L_{i0}(d_i)|x_i,\theta] \\ \qquad = \sum\limits_{d_i=d_{i0}+1}^{M_i-m_i+x_i} L_{i1}(d_i)g_i(d_i|x_i,\theta) - \sum\limits_{d_i=x_i}^{d_{i0}-1} L_{i0}(d_i)g_i(d_i|x_i,\theta), \end{cases} \tag{3.1}$$

where

$$\begin{aligned} g_i(d_i|x_i,\theta) &= f_i(x_i|d_i)g_i(d_i|\theta)/f_i(x_i|\theta) \\ &= \frac{\binom{d_i}{x_i}\binom{M_i-d_i}{m_i-x_i}}{\binom{M_i}{m_i}\binom{m_i}{x_i}}\theta^{d_i-x_i}(1-\theta)^{M_i-d_i-m_i+x_i}, x_i \le d_i \le m_i, \end{aligned}$$

is the posterior probability function of d_i given $X_i = x_i$, and the summation $\sum\limits_{d=s}^{t} \equiv 0$ if $t < s$. It can be seen that given $X_i = x_i$, $d_i - x_i \sim B(M_i - m_i, \theta)$. Let $r_{ki}(\theta, \delta_{ki})$ denote the i–th component Bayes risk, $i = 1, \ldots, k$, and let $r_k(\theta, \underset{\sim}{\delta}_k)$ be the overall Bayes risk of the selection rule $\underset{\sim}{\delta}_k = (\delta_{k1}, \ldots, \delta_{kk})$. From the statistical model described in the preceding and the loss function $L(\underset{\sim}{d}, \underset{\sim}{a})$ given in (2.2), it follows that

$$\begin{cases} r_{ki}(\theta,\delta_{ki}) = \sum\limits_{\mathcal{X}} \delta_{ki}(\underset{\sim}{x})H_i(x_i,\theta)f(\underset{\sim}{x}|\theta) + C_i(\theta), \\ r_k(\theta,\underset{\sim}{\delta}_k) = \sum\limits_{i=1}^{k} r_{ki}(\theta,\delta_{ki}); \end{cases} \tag{3.2}$$

where $f(\underset{\sim}{x}|\theta) = \prod_{i=1}^{k} f_i(x_i|\theta)$ is the marginal joint probability function of $(X_1, \ldots, X_k)$.

Therefore, relative to the fixed parameter θ, a Bayes rule, denoted by $\underset{\sim}{\delta}_k^\theta = (\delta_{k1}^\theta, \ldots, \delta_{kk}^\theta)$, can be obtained as follows:

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

$$\delta_{ki}^\theta(\underset{\sim}{x}) = \begin{cases} 1 & \text{if } H_i(x_i, \theta) < 0, \\ 0 & \text{otherwise.} \end{cases} \tag{3.3}$$

3.2. Case 2. A Hierarchical Bayes Selection Rule wrt Beta $(\alpha\mu,\ \alpha(1-\mu))$ Prior

First, we introduce some notation. For each $i = 1, \ldots, k$, let

$$\begin{aligned} f_i(x_i|\alpha,\mu) &= \int_0^1 f_i(x_i|\theta)h(\theta|\alpha,\mu)d\theta, \\ f_i(x_i, d_i|\alpha,\mu) &= \int_0^1 f_i(x_i|d_i)g_i(d_i|\theta)h(\theta|\alpha,\mu)d\theta, \\ g_i(d_i|x_i,\alpha,\mu) &= f_i(x_i,d_i|\alpha,\mu)/f_i(x_i|\alpha,\mu) \\ C_i &= \int_0^1 C_i(\theta)h(\theta|\alpha,\mu)d\theta \\ Q_i(x_i,\alpha,\mu) &= \int H_i(x_i,\theta)h(\theta|\alpha,\mu)d\theta \\ &= \sum_{d_i=d_{i0}+1}^{M_i-m_i+x_i} L_{i1}(d_i)g_i(d_i|x_i,\alpha,\mu) - \sum_{d_i=x_i}^{d_{i0}-1} L_{i0}(d_i)g_i(d_i|x_i,\alpha,\mu). \end{aligned} \tag{3.4}$$

Straightforward computations show that

$$g_i(d_i|x_i,\alpha,\mu) = \binom{M_i - m_i}{d_i - x_i} \times \frac{\Gamma(d_i+\alpha\mu)\Gamma(M_i+\alpha(1-\mu)-d_i)\Gamma(m_i+\alpha)}{\Gamma(M_i+\alpha)\Gamma(x_i+\alpha\mu)\Gamma(m_i+\alpha(1-\mu)-x_i)}. \tag{3.5}$$

For fixed values of the parameters α and μ, we denote the i–th component Bayes risk and the overall Bayes risk of the selection rule $\underset{\sim}{\delta}_k = (\delta_{k1}, \ldots, \delta_{kk})$ by $R_{ki}(\alpha,\mu,\delta_{ki})$ and $R_k(\alpha,\mu,\underset{\sim}{\delta}_k)$, respectively. Then, under the corresponding statistical model,

$$\begin{cases} R_{ki}(\alpha,\mu,\delta_{ki}) = \sum_{\mathcal{X}} \delta_{ki}(\underset{\sim}{x})Q_i(x_i,\alpha,\mu)f(\underset{\sim}{x}|\alpha,\mu) + C_i, \\ R_k(\alpha,\mu,\underset{\sim}{\delta}_k) = \sum_{i=1}^{k} R_{ki}(\alpha,\mu,\delta_{ki}), \end{cases} \tag{3.6}$$

where $f(\underset{\sim}{x}|\alpha,\mu) = \prod_{i=1}^{k} f_i(x_i|\alpha,\mu)$.

Under this hierarchical statistical model, a Bayes selection rule, called the hierarchical Bayes selection rule, is $\underset{\sim}{\delta}_k^{\mathrm{HB}} = (\delta_{k1}^{\mathrm{HB}}, \ldots, \delta_{kk}^{\mathrm{HB}})$, where for each $\underset{\sim}{x} \in \mathcal{X}$,

$$\delta_{ki}^{\mathrm{HB}}(\underset{\sim}{x}) = \begin{cases} 1 & \text{if } Q_i(x_i,\alpha,\mu) < 0, \\ 0 & \text{otherwise.} \end{cases} \tag{3.7}$$

Remark 3.1.

(a) Note that for each i, both functions $H_i(x_i,\theta)$ and $Q_i(x_i,\alpha,\mu)$ are independent of all x_j for which $j \neq i$. Hence, both the i–th component selection rules δ_{ki}^{θ} and $\delta_{ki}^{\mathrm{HB}}$ depend on $\underset{\sim}{x}$ only through x_i only. Also, as $x_i \geq d_{i0}$, both $H_i(x_i,\theta) \geq 0$ and $Q_i(x_i,\alpha,\mu) \geq 0$ and hence, $\delta_{ki}^{\theta}(\underset{\sim}{x}) = 0$ and $\delta_{ki}^{\mathrm{HB}}(\underset{\sim}{x}) = 0$.

(b) **Definition 3.1**. A selection rule $\underset{\sim}{\delta}_k = (\delta_{k1}, \ldots, \delta_{kk})$ is said to be monotone if for each $i = 1, \ldots, k$, $\delta_{ki}(\underset{\sim}{x})$ is nonincreasing in x_j while all the other variables are kept fixed for each $j = 1, \ldots, k$.

Note that $L_{i1}(d_i) - L_{i0}(d_i)$ is nondecreasing in d_i. Also, the posterior probability function $g_i(d_i|x_i,\theta)$ has monotone likelihood ratio. Therefore, $H_i(x_i,\theta) = E[L_{i1}(d_i) - L_{i0}(d_i)|x_i,\theta]$ and $Q_i(x_i,d,\mu) = \int H_i(x_i,\theta)h(\theta|\alpha,\mu)d\theta$ are increasing functions of x_i for $0 \leq x_i \leq d_{i0} - 1$. This fact together with Remark 3.1(a), and the definitions of $\underset{\sim}{\delta}_k^{\theta}$ and $\underset{\sim}{\delta}_k^{\mathrm{HB}}$ imply the monotonicity of the selection rules $\underset{\sim}{\delta}_k^{\theta}$ and $\underset{\sim}{\delta}_k^{\mathrm{HB}}$.

(c) For each $i = 1, \ldots, k$, let

$$B_i(\theta) = \{x|H_i(x,\theta) \geq 0,\ x = 0, 1, \ldots, m_i\},$$

$$B_i = \{x|Q_i(x,\alpha,\mu) \geq 0,\ x = 0, 1, \ldots, m_i\}.$$

Note that $B_i(\theta) \neq \phi$ and $B_i \neq \phi$ since $H_i(d_{i0},\theta) \geq 0$ and $Q_i(d_{i0},\alpha,\mu) \geq 0$. Let $b_i(\theta) = \min B_i(\theta)$ and $b_i^{\mathrm{HB}} = \min B_i$. By the increasing property of $H_i(x_i,\theta)$ and $Q_i(x_i,\alpha,\mu)$, $H_i(x_i,\theta) \geq 0$ ($Q_i(x_i,\alpha,\mu) \geq 0$) iff $x_i \geq b_i(\theta)$($x_i \geq b_i^{\mathrm{HB}}$). Therefore, the Bayes rule $\underset{\sim}{\delta}_k^{\theta}$ and the hierarchical Bayes rule $\underset{\sim}{\delta}_k^{\mathrm{HB}}$ can be written as follows: For each $\underset{\sim}{x} \in \mathcal{X}$ and $i = 1, \ldots, k$,

$$\delta_{ki}^{\theta}(\underset{\sim}{x}) = \begin{cases} 1 & \text{if } x_i < b_i(\theta); \\ 0 & \text{otherwise.} \end{cases} \tag{3.8}$$

$$\delta_{ki}^{\mathrm{HB}}(\underset{\sim}{x}) = \begin{cases} 1 & \text{if } x_i < b_i^{\mathrm{HB}}; \\ 0 & \text{otherwise.} \end{cases} \tag{3.9}$$

Accordingly, the minimum Bayes risks in the two cases are:

$$\begin{aligned} r_k(\theta, \underset{\sim}{\delta}_k^{\theta}) &= \sum_{i=1}^{k} r_{ki}(\theta, \delta_{ki}^{\theta}) \\ &= \sum_{i=1}^{k} \left[\sum_{x_i=0}^{b_i(\theta)-1} H_i(x_i, \theta) f_i(x_i|\theta) + C_i(\theta) \right]; \end{aligned} \tag{3.10}$$

$$\begin{aligned} R_k(\alpha, \mu, \underset{\sim}{\delta}_k^{\mathrm{HB}}) &= \sum_{i=1}^{k} R_{ki}(\alpha, \mu, \delta_{ki}^{\mathrm{HB}}) \\ &= \sum_{i=1}^{k} \left[\sum_{x_i=0}^{b_i^{\mathrm{HB}}-1} Q_i(x_i, \alpha, \mu) f_i(x_i|\alpha, \mu) + C_i \right]. \end{aligned} \tag{3.11}$$

(d) Consider the following linear loss: For each $i = 1, \ldots, k$, let

$$\begin{cases} L_{i1}(d_i) = (d_i - d_{i0}) I(d_i > d_{i0}), \\ L_{i0}(d_i) = (d_{i0} - d_i) I(d_{i0} > d_i); \end{cases} \tag{3.12}$$

where $I(A)$ denotes the indicator function of the set A. Then,

$$H_i(x_i, \theta) = (M_i - m_i)\theta + x_i - d_{i0};$$

$$Q_i(x_i, \alpha, \mu) = (M_i - m_i)(x_i + \alpha\mu)/(m_i + \alpha) + x_i - d_{i0}.$$

Therefore, we have:

$$\delta_{ki}^{\theta}(\underset{\sim}{x}) = \begin{cases} 1 & \text{if } (M_i - m_i)\theta + x_i < d_{i0}; \\ 0 & \text{otherwise;} \end{cases} \tag{3.13}$$

and

$$\delta_{ki}^{\mathrm{HB}}(\underset{\sim}{x}) = \begin{cases} 1 & \text{if } \frac{(M_i - m_i)(x_i + \alpha\mu)}{m_i + \alpha} + x_i < d_{i0}; \\ 0 & \text{otherwise.} \end{cases} \tag{3.14}$$

4. Construction of Empirical Bayes Selection Rules

Note that the Bayes selection rule $\underset{\sim}{\delta}_k^\theta$ and the hierarchical Bayes selection rule $\underset{\sim}{\delta}_k^{\mathrm{HB}}$ are strongly sensitive to the values of the parameters θ and (α, μ), respectively. In many practical problems, the values of these parameters may not be known. In such situations, it is not possible to employ the Bayes selection rule $\underset{\sim}{\delta}_k^\theta$ and/or the hierarchical Bayes selection rule $\underset{\sim}{\delta}_k^{\mathrm{HB}}$. Hence, the empirical Bayes approach is employed in the following. We combine information from each of the k populations and use them to make a decision for each of the k component problems. Two empirical Bayes selection rules, called parametric empirical Bayes and hierarchical empirical Bayes, are studied, respectively, according to the appropriate statistical model.

Case 1. A Parametric Empirical Bayes Selection Rule

Under the statistical model of Case 1, for the common parameter θ being kept fixed, $X_1, \ldots, X_k$ are mutually independent with $X_i \sim B(m_i, \theta)$, $i = 1, \ldots, k$. Therefore, $X_1 + \ldots + X_k \sim B\left(\sum_{i=1}^k m_i, \theta\right)$. Hence $\hat{\theta} = (X_1 + \ldots + X_k)/\sum_{i=1}^k m_i$ is an unbiased estimator of and sufficient statistic for the parameter θ. It is suggested to estimate the function $H_i(x_i, \theta)$ by $H_i(x_i, \hat{\theta})$. We then propose a parametric empirical Bayes selection rule $\underset{\sim}{\delta}_k^* = (\delta_{k1}^*, \ldots, \delta_{kk}^*)$ as follows: For each $i = 1, \ldots, k$, and $\underset{\sim}{x} \in \mathcal{X}$,

$$\delta_{ki}^*(\underset{\sim}{x}) = \begin{cases} 1 & \text{if } H_i(x_i; \hat{\theta}) < 0, \\ 0 & \text{otherwise.} \end{cases} \tag{4.1}$$

For the linear loss of (3.12), the empirical Bayes selection rule $\underset{\sim}{\delta}_k^*$ turns out to be:

$$\delta_{ki}^*(\underset{\sim}{x}) = \begin{cases} 1 & \text{if } (M_i - m_i)\hat{\theta} + x_i < d_{i0}; \\ 0 & \text{otherwise.} \end{cases} \tag{4.2}$$

Case 2. A Hierarchical Empirical Bayes Selection Rule

Under the statistical model of Case 2, for each $i = 1, \ldots, k$,

$$E[X_i] = m_i\mu, \text{ and}$$
$$E[X_i^2] = \frac{m_i\alpha(1-\mu)\mu}{\alpha+1} + \frac{m_i^2(\alpha\mu+1)\mu}{\alpha+1} \equiv \mu_{i2},$$

where μ_{i2} is decreasing in α and tends to $m_i\mu(1-\mu)+m_i^2\mu^2$ as α tends to infinity. Therefore,

$$E\left[\sum_{i=1}^k X_i\right]=\mu\sum_{i=1}^k m_i, \text{ and}$$

$$E\left[\sum_{i=1}^k X_i^2\right]=\frac{\alpha\mu(1-\mu)\sum_{i=1}^k m_i}{\alpha+1}+\frac{(\alpha\mu+1)\mu\sum_{i=1}^k m_i^2}{\alpha+1}.$$

Hence,

$$\alpha=\frac{\mu\sum_{i=1}^k m_i^2-\sum_{i=1}^k \mu_{i2}}{\sum_{i=1}^k \mu_{i2}-\mu^2\sum_{i=1}^k m_i^2-\mu(1-\mu)\sum_{i=1}^k m_i}. \tag{4.3}$$

We may use $\hat{\mu}=\sum_{i=1}^k X_i/\sum_{i=1}^k m_i$ to estimate the parameter μ and $\sum_{i=1}^k X_i^2$ to estimate $\sum_{i=1}^k \mu_{i2}$. Note that, since $m_i\geq 2$, $\mu_{i2}-m_i^2\mu^2-m_i\mu(1-\mu)>0$ for each $i=1,\ldots,k$, and hence, $\sum_{i=1}^k \mu_{i2}-\mu^2\sum_{i=1}^k m_i^2-\mu(1-\mu)\sum_{i=1}^k m_i>0$. However, it is possible that $D(X_1,\ldots,X_k)\equiv\sum_{i=1}^k X_i^2-\hat{\mu}^2\sum_{i=1}^k m_i^2-\hat{\mu}(1-\hat{\mu})\left(\sum_{i=1}^k m_i\right)\leq 0$. Also, it is possible that $\hat{\mu}\sum_{i=1}^k m_i^2-\sum_{i=1}^k X_i^2<0$ though $\mu\sum_{i=1}^k m_i^2-\sum_{i=1}^k \mu_{i2}$ is always nonnegative. Motivated by the form of (4.3) and the decreasing property of μ_{i2} with respect to α, we define

$$\hat{\alpha}=\begin{cases}\frac{\max(\hat{\mu}\sum_{i=1}^k m_i^2-\sum_{i=1}^k X_i^2,0)}{D(X_1,\ldots,X_k)} & \text{if } D(X_1,\ldots,X_k)>0,\\ \infty & \text{otherwise.}\end{cases} \tag{4.4}$$

Note that when x_i and μ are kept fixed, $\lim_{\alpha\to\infty} Q_i(x_i,\alpha,\mu)$ exists. We denote this limit value by $Q_i(x_i,\infty,\mu)$. Then,

$$Q_i(x_i,\infty,\mu)=\sum_{d_i=d_{i0}+1}^{M_i-m_i+x_i} L_{i1}(d_i)g_i(d_i|x_i,\infty,\mu)-\sum_{d_i=x_i}^{d_{i0}-1} L_{i0}(d_i)g_i(d_i|x_i,\infty,\mu),$$

where

$$g_i(d_i|x_i,\infty,\mu)=\binom{M_i-m_i}{d_i-x_i}\mu^{d_i-x_i}(1-\mu)^{M_i-m_i+x_i-d_i}.$$

We then estimate $Q_i(x_i, \alpha, \mu)$ by $Q_i(x_i, \hat{\alpha}, \hat{\mu})$, and propose a hierarchical empirical Bayes selection rule $\hat{\underset{\sim}{\delta}}_k = (\hat{\delta}_{k1}, \ldots, \hat{\delta}_{kk})$ defined as follows: For each $i = 1, \ldots, k$, and $\underset{\sim}{x} \in \mathcal{X}$,

$$\hat{\delta}_{ki}(\underset{\sim}{x}) = \begin{cases} 1 & \text{if } Q_i(x_i, \hat{\alpha}, \hat{\mu}) < 0, \\ 0 & \text{otherwise.} \end{cases} \tag{4.5}$$

For the linear loss of (3.12), the hierarchical empirical Bayes rule $\hat{\underset{\sim}{\delta}}_k$ turns out to be:

$$\hat{\delta}_{k_i}(\underset{\sim}{x}) = \begin{cases} 1 & \text{if } \frac{(M_i - m_i)(x_i + \hat{\alpha}\hat{\mu})}{m_i + \hat{\alpha}} + x_i < d_{i0}; \\ 0 & \text{otherwise.} \end{cases} \tag{4.6}$$

Remark 4.1

(a) It has already been mentioned in Remark 3.1(b) that $H_i(x_i, \theta)$ is an increasing function of x_i for $0 \le x_i \le d_{i0} - 1$. Also, an algebraic computation shows that for each fixed x_i, $0 \le x_i \le d_{i0} - 1$, $H_i(x_i, \theta)$ is an increasing function of θ, $0 \le \theta \le 1$. Since the estimator $\hat{\theta}$ is increasing in x_j for each $j = 1, \ldots, k$, it can be seen that $H_i(x_i, \hat{\theta})$ is increasing in x_j for $j \ne i$, and increasing in x_i for $0 \le x_i \le d_{i0} - 1$. By the definition of $\underset{\sim}{\delta}_k^*$, we can see that $\underset{\sim}{\delta}_k^*$ has the monotonicity property.

However, though the hierarchical Bayes selection rule $\underset{\sim}{\delta}_k^{\mathrm{HB}}$ has the monotonicity property, the hierarchical empirical Bayes selection rule $\hat{\underset{\sim}{\delta}}_k$ may not possess the same property.

(b) For the statistical model of Case 1, consider a Beta$(\alpha\mu, \alpha(1-\mu))$ prior having density $h(\theta|\alpha, \mu)$ on the common parameter θ, $0 < \theta < 1$. Under the linear loss of (3.12), a hierarchical Bayes selection rule, denoted by $\underset{\sim}{\delta}_k^B = (\delta_{k1}^B, \ldots, \delta_{kk}^B)$, is:

$$\delta_{ki}^B(\underset{\sim}{x}) = \begin{cases} 1 & \text{if } E[d_i|\underset{\sim}{x}, \alpha, \mu] < d_{i0}, \\ 0 & \text{otherwise;} \end{cases} \tag{4.7}$$

where $E[d_i|\underset{\sim}{x}, \alpha, \mu]$ is the posterior mean of d_i given $\underset{\sim}{X} = \underset{\sim}{x}$. Note that for the $M_i - m_i$ unsampled units in population π_i, conditional on θ, $d_i - x_i$ has a $B(M_i - m_i, \theta)$ distribution, and the posterior on θ is Beta$(\alpha\mu + \sum_{j=1}^{k} x_j, \alpha(1-\mu) + \sum_{j=1}^{k}(m_j - x_j))$. Then,

$$E[d_i|\underset{\sim}{x}, \alpha, \mu] = (M_i - m_i)(\alpha\mu + \sum_{j=1}^{k} x_j)/(\alpha + \sum_{j=1}^{k} m_j) + x_i. \tag{4.8}$$

The Bayes risk of the selection rule $\underset{\sim}{\delta}_k^B$ is:

$$r_k(\alpha,\mu,\underset{\sim}{\delta}_k^B)=\sum_{i=1}^{k} r_{ki}(\alpha,\mu,\delta_{ki}^B)$$

where

$$r_{ki}(\alpha,\mu,\delta_{ki}^B)=\sum_{\underset{\sim}{x}\in\mathcal{X}}\delta_{ki}^B(\underset{\sim}{x})[E[d_i|\underset{\sim}{x},\alpha,\mu]-d_{i0}]f(\underset{\sim}{x})+C_i'$$

$$C_i'=\sum_{d_i=0}^{d_{i0}-1}(d_{i0}-d_i)\int_0^1 g_i(d_i|\theta)h(\theta|\alpha,\mu)d\theta,$$

$$f(\underset{\sim}{x})=\int_0^1\prod_{i=1}^{k} f_i(x_i|\theta)h(\theta|\alpha,\mu)d\theta.$$

From (4.8) we see that $E[d_i|\underset{\sim}{x},\alpha,\mu]$ is a continuous function of α and $\lim_{\alpha\to 0}E[d_i|\underset{\sim}{x},\alpha,\mu]=(M_i-m_i)\sum_{j=1}^{k}x_j/\sum_{j=1}^{k}m_j+x_i$. Since the sample space $\mathcal{X}$ is finite, for $\alpha>0$ being very close to zero, we have $E[d_i|\underset{\sim}{x},\alpha,\mu]<d_{i0}$ iff $(M_i-m_i)\sum_{j=1}^{k}x_j/\sum_{j=1}^{k}m_j+x_i<d_{i0}$ for all $\underset{\sim}{x}$ and for $i=1,\dots,k$. This fact implies that the empirical Bayes selection rule $\underset{\sim}{\delta}_k^*$ is a hierarchical Bayes selection rule relative to the hierarchical priors: $d_i\sim B(M_i,\theta),d_1,\dots,d_k$ are independently distributed and $\theta\sim$ Beta$(\alpha\mu,\alpha(1-\mu))$ for some $\alpha>0$, α being very close to zero, (see (4.2)). Hence, the empirical Bayes selection rule $\underset{\sim}{\delta}_k^*$ is admissible in the sense that there exists no other selection rule $\underset{\sim}{\delta}_k^0$ such that

$$r_k(\theta,\underset{\sim}{\delta}_k^0)\le r_k(\theta,\underset{\sim}{\delta}_k^*)\text{ for all }\theta\in(0,1),$$

and $r_k(\theta,\underset{\sim}{\delta}_k^0)<r_k(\theta,\underset{\sim}{\delta}_k^*)$ for all $\theta\in A\subset(0,1)$ for which $\int_A d\theta>0$.

5. Asymptotic Optimality of Empirical Bayes Selection Rules

5.1. Case 1. Asymptotic Optimality of $\underset{\sim}{\delta}_k^*$

Under the statistical model of Case 1, $r_{ki}(\theta,\delta_{ki}^*)\ge r_{ki}(\theta,\delta_{ki}^\theta)$ since $r_{ki}(\theta,\delta_{ki}^\theta)$ is the minimum Bayes risk for the i–th component selection problem. Hence, $r_k(\theta,\underset{\sim}{\delta}_k^*)-r_k(\theta,\underset{\sim}{\delta}_k^\theta)\ge 0$ for all k. This nonnegative difference of the regret risk $r_k(\theta,\underset{\sim}{\delta}_k^*)-r_k(\theta,\underset{\sim}{\delta}_k^\theta)$ is a suitable measure of the performance of the empirical Bayes selection rule $\underset{\sim}{\delta}_k^*$. We evaluate the performance of $\underset{\sim}{\delta}_k^*$ for large k case.

Definition 5.1. A selection rule $\underset{\sim}{\delta}_k$ is said to be asymptotically optimal of order $\beta_k(\theta)$, if

$$r_k(\theta, \underset{\sim}{\delta}_k) - r_k(\theta, \underset{\sim}{\delta}_k^\theta) = O(\beta_k(\theta)) \text{ for each } \theta \in (0,1) \text{ as } k \to \infty$$

where $\beta_k(\theta) > 0$ for all k and $\lim_{k\to\infty} \beta_k(\theta) = 0$ for each $\theta \in (0,1)$.

For the empirical Bayes selection rule $\underset{\sim}{\delta}_k^*$, the associated Bayes risk is:

$$r_k(\theta, \underset{\sim}{\delta}_k^*) = \sum_{i=1}^{k} r_{ki}(\theta, \delta_{ki}^*), \tag{5.1}$$

where from (3.2),

$$\begin{aligned} r_{ki}(\theta, \delta_{ki}^*) &= \sum_{\underset{\sim}{x} \in \mathcal{X}} \delta_{ki}^*(\underset{\sim}{x}) H_i(x_i, \theta) f(\underset{\sim}{x}|\theta) + C_i(\theta) \\ &= \sum_{x_i} E_i[\delta_{ki}^*(\underset{\sim}{X})|X_i = x_i] H_i(x_i,\theta) f_i(x_i|\theta) + C_i(\theta), \end{aligned} \tag{5.2}$$

where the expectation E_i is taken with respect to the probability measure generated by $\underset{\sim}{X}(i) = (X_1, \ldots, X_{i-1}, X_{i+1}, \ldots, X_k)$.

From (3.2), (3.10), (5.1) and (5.2),

$$0 \le r_k(\theta, \underset{\sim}{\delta}_k^*) - r_k(\theta, \underset{\sim}{\delta}_k^\theta) = \sum_{i=1}^{k} [r_{ki}(\theta, \delta_{ki}^*) - r_{ki}(\theta, \delta_{ki}^\theta)] \tag{5.3}$$

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

$$\begin{aligned} 0 &\le r_{ki}(\theta, \delta_{ki}^*) - r_{ki}(\theta, \delta_{ki}^\theta) \\ &= \sum_{x_i} E_i[\delta_{ki}^*(\underset{\sim}{X}) - \delta_{ki}^\theta(\underset{\sim}{X})|X_i = x_i] H_i(x_i,\theta) f_i(x_i|\theta) \\ &= \sum_{x_i=0}^{b_i(\theta)-1} [-H_i(x_i,\theta)] P_i\{\delta_{ki}^*(\underset{\sim}{X}) = 0 | X_i = x_i\} f_i(x_i|\theta) \\ &\quad + \sum_{x_i=b_i(\theta)}^{d_{i0}-1} H_i(x_i,\theta) P_i\{\delta_{ki}^*(\underset{\sim}{X}) = 1 | X_i = x_i\} f_i(x_i|\theta), \end{aligned} \tag{5.4}$$

where P_i is the probability measure generated by $\underset{\sim}{X}(i)$.

For each fixed x_i, let $H_i(\theta|x_i) = H(x_i, \theta)$. Then $H_i(\theta|x_i)$ is an increasing function of the parameter θ.

For each $0 \le x_i \le b_i(\theta) - 1$, $H_i(\theta|x_i) < 0$, and therefore, $\theta < H^{-1}(0|x_i)$. Hence, by (4.1),

$$\begin{aligned}
&P_i\{\delta^*_{ki}(\underset{\sim}{X}) = 0|X_i = x_i\} \\
=&P_i\{H_i(\hat{\theta}|x_i) \ge 0|X_i = x_i\} \\
=&P_i\{\hat{\theta} \ge H_i^{-1}(0|x_i)|X_i = x_i\} \\
=&P_i\left\{\sum_{\substack{j=1\\j\neq i}}^{k} X_j \Big/ \sum_{\substack{j=1\\j\neq i}}^{k} m_j - \theta \ge H_i^{-1}(0|x_i) - \theta + \Psi_i(x_i)\right\}
\end{aligned}$$

where $H_i^{-1}(0|x_i) - \theta > 0$ and $\Psi_i(x_i) \equiv [H_i^{-1}(0|x_i)m_i - x_i]/\sum_{\substack{j=1\\j\neq i}}^{k} m_j$ tends to 0 as k tends to infinity. Hence, for sufficiently large k, $H_i^{-1}(0|x_i) - \theta + \Psi_i(x_i) \ge [H_i^{-1}(0|x_i) - \theta]/2$, and therefore,

$$\begin{aligned}
&P_i\{\delta^*_{ki}(\underset{\sim}{X}) = 0|X_i = x_i\} \\
\le&P_i\left\{\sum_{\substack{j=1\\j\neq i}}^{k} X_j \Big/ \sum_{\substack{j=1\\j\neq i}}^{k} m_j - \theta \ge [H_i^{-1}(0|x_i) - \theta]/2\right\} \\
\le&\exp\left\{-[H_i^{-1}(0|x_i) - \theta]^2 \sum_{\substack{j=1\\j\neq i}}^{k} m_j/2\right\}.
\end{aligned} \tag{5.6}$$

In (5.6), the last inequality follows from Theorem 1 of Hoeffding (1963) and the fact that marginally, $\sum_{\substack{j=1\\j\neq i}}^{k} X_j \sim B\left(\sum_{\substack{j=1\\j\neq i}}^{k} m_j, \theta\right)$.

For each $b_i(\theta) \le x_i \le d_{i0} - 1$ such that $H_i(\theta|x_i) > 0$, following a similar discussion as in the preceding, we can obtain: $H_i^{-1}(0|x_i) - \theta < 0$ and

$$P_i\{\delta^*_{ki}(\underset{\sim}{X}) = 1|X_i = x_i\} \le \exp\left\{-[H_i^{-1}(0|x_i) - \theta]^2 \sum_{\substack{j=1\\j\neq i}}^{k} m_j/2\right\}. \tag{5.7}$$

We summarize the preceding results in the following theorem.

Theorem 5.1. Suppose that $m_* \le m_i \le M_i \le M^*$ for all $i = 1, \ldots, k$, and $L_{ij}(d_i) \le L^*$ for all $j = 0, 1$ and $i = 1, \ldots, k$, where $L_{ij}(\cdot), j = 0, 1$, depend on i only through the values of m_i and M_i; and where the bound values L^*, m_* and M^* are independent of k. Then, under the statistical model of Case 1, for each $\theta \in (0,1)$, the empirical Bayes selection rule $\underset{\sim}{\delta}_k^*$ is asymptotically optimal, its rate of convergence being $O(\exp(-e_1(\theta,k)k + \ln k))$, where

$$e_1(\theta,k) = \min_{1\le i\le k} \min_{x_i}\{m_*[H_i^{-1}(0|x_i) - \theta]^2/2 : \ H_i^{-1}(0|x_i) - \theta \neq 0\}.$$

Note that $e_1(\theta,k) > 0$ since $M_i \le M^*$ for all $i = 1,\dots,k$.

Proof: By the boundness of the loss $L_{ij}(d_i) \le L^*$ for all $j = 0,1,\ i = 1,\dots,k$, we have

$$|H_i(x_i,\theta)| \le L^* \text{ for all } x_i = 0,1,\dots,d_{i0}-1 \text{ and } i = 1,\dots,k.$$

By the definition of $e_1(\theta,k)$,

$$P_i\{\delta_{ki}^*(\underset{\sim}{X}) = 0|X_i = x_i\} \le \exp\{-e_1(\theta,k)k\},\ 0 \le x_i \le b_i(\theta) - 1,$$

and
$$P_i\{\delta_{ki}^*(\underset{\sim}{X}) = 1|X_i = x_i\} \le \exp\{-e_1(\theta,k)k\},\ b_i(\theta) \le x_i \le d_{i0} - 1.$$

Substituting the preceding inequalities into (5.4), we obtain:

$$0 \le r_{ki}(\theta,\delta_{ki}^*) - r_{ki}(\theta,\delta_{ki}^{\theta}) \le L^* \exp\{-e_1(\theta,k)k\},\ i = 1,\dots,k,$$

and hence

$$0 \le r_k(\theta,\underset{\sim}{\delta}_k^*) - r_k(\theta,\underset{\sim}{\delta}_k^{\theta}) \le L^* \exp\{-e_1(\theta,k)k + \ln k\}.$$

Note that since the loss functions $L_{ij}(\cdot), j = 0,1$, depend on i only through the values of m_i and M_i and $M_i \le M^*$ for all $i = 1,\dots,k$, it follows by the definition of $e_1(\theta,k)$ that $e_1(\theta,k) \ge e_2(\theta)$ for all k for some positive value $e_2(\theta)$. This completes the proof. □

5.2. Case 2. Asymptotic Optimality of $\hat{\underset{\sim}{\delta}}_k$

Under the statistical model of Case 2, for any selection rule $\underset{\sim}{\delta}_k$, $R_{ki}(\alpha,\mu,\delta_{ki}) - R_{ki}(\alpha,\mu,\delta_{ki}^{\mathrm{HB}}) \ge 0$ since $R_{ki}(\alpha,\mu,\delta_{ki}^{\mathrm{HB}})$ is the minimum Bayes risk for the i–th component selection problem. Hence $R_k(\alpha,\mu,\underset{\sim}{\delta}_k) - R_k(\alpha,\mu,\underset{\sim}{\delta}_k^{\mathrm{HB}}) \ge 0$. This nonnegative difference is used as a measure of the performance of the selection rule $\underset{\sim}{\delta}_k$.

Definition 5.2. Under the statistical model of Case 2, a selection rule $\underset{\sim}{\delta}_k$ is said to be asymptotically optimal of order $\beta_k(\alpha,\mu)$ if

$$R_k(\alpha,\mu,\underset{\sim}{\delta}_k) - R_k(\alpha,\mu,\underset{\sim}{\delta}_k^{\text{HB}}) = O(\beta_k(\alpha,\mu)),$$

for each (α,μ) as $k \to \infty$, where $\beta_k(\alpha,\mu) > 0$ for all k, and $\lim_{k\to\infty} \beta_k(\alpha,\mu) = 0$ for each (α,μ).

For the empirical Bayes selection rule $\hat{\underset{\sim}{\delta}}_k$, the associated Bayes risk is:

$$R_k(\alpha,\mu,\hat{\underset{\sim}{\delta}}_k) = \sum_{i=1}^{k} R_{ki}(\alpha,\mu,\hat{\delta}_{ki}), \tag{5.8}$$

where

$$\begin{aligned} R_{ki}(\alpha,\mu,\hat{\delta}_{ki}) &= \sum_{\underset{\sim}{x}\in\mathcal{X}} \hat{\delta}_{ki}(\underset{\sim}{x}) Q_i(x_i,\alpha,\mu) f(\underset{\sim}{x}|\alpha,\mu) + C_i \\ &= \sum_{x_i} E_i[\hat{\delta}_{ki}(\underset{\sim}{X})|X_i = x_i] Q_i(x_i,\alpha,\mu) f_i(x_i|\alpha,\mu) + C_i. \end{aligned} \tag{5.9}$$

From (3.6), (3.7), (3.11), (5.8) and (5.9),

$$\begin{aligned} 0 &\le R_k(\alpha,\mu,\hat{\underset{\sim}{\delta}}_k) - R_k(\alpha,\mu,\underset{\sim}{\delta}_k^{\text{HB}}) \\ &= \sum_{i=1}^{k} [R_{ki}(\alpha,\mu,\hat{\delta}_{ki}) - R_{ki}(\alpha,\mu,\delta_{ki}^{\text{HB}})] \end{aligned} \tag{5.10}$$

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

$$\begin{aligned} 0 &\le R_{ki}(\alpha,\mu,\hat{\delta}_{ki}) - R_{ki}(\alpha,\mu,\delta_{ki}^{\text{HB}}) \\ &= \sum_{x_i} E_i[\hat{\delta}_{ki}(\underset{\sim}{X}) - \delta_{ki}^{\text{HB}}(\underset{\sim}{X})|X_i = x_i] Q_i(x_i,\alpha,\mu) f_i(x_i|\alpha,\mu) \\ &= \sum_{x_i=0}^{b_i^{\text{HB}}-1} [-Q_i(x_i,\alpha,\mu)] P_i\{\hat{\delta}_{ki}(\underset{\sim}{X}) = 0|X_i = x_i\} f_i(x_i|\alpha,\mu) \\ &+ \sum_{x_i=b_i^{\text{HB}}}^{d_{i0}-1} Q_i(x_i,\alpha,\mu) P_i\{\hat{\delta}_{ki}(\underset{\sim}{X}) = 1|X_i = x_i\} f_i(x_i|\alpha,\mu). \end{aligned} \tag{5.11}$$

Note that for each x_i, $Q_i(x_i,\alpha,\mu)$ is a continuous function of the variables (α,μ) (see (3.4) and (3.5)).

For each $0 \le x_i \le b_i^{\rm HB} - 1$, $Q_i(x_i, \alpha, \mu) < 0$. Then,

$$
\begin{aligned}
& P_i\{\hat{\delta}_{ki}(\underset{\sim}{X}) = 0 | X_i = x_i\} \\
&= P_i\{Q_i(x_i, \hat{\alpha}, \hat{\mu}) \ge 0 | X_i = x_i\} \\
&= P_i\{Q_i(x_i, \hat{\alpha}, \hat{\mu}) - Q_i(x_i, \alpha, \mu) > -Q_i(x_i, \alpha, \mu) | X_i = x_i\} \\
&\le P_i\{|\hat{\alpha} - \alpha| > q_i(x_i, \alpha, \mu) \text{ or } |\hat{\mu} - \mu| > q_i(x_i, \alpha, \mu) | X_i = x_i\} \\
&\le P_i\{|\hat{\alpha} - \alpha| > q_i(x_i, \alpha, \mu) | X_i = x_i\} + P_i\{|\hat{\mu} - \mu| > q_i(x_i, \alpha, \mu) | X_i = x_i\}
\end{aligned}
$$

for some $q_i(x_i, \alpha, \mu) > 0$ and the first inequality is obtained due to the continuity property of $Q_i(x_i, \alpha, \mu)$ wrt (α, μ).

Similarly, for each x_i, $b_i^{\rm HB} \le x_i \le d_{i0} - 1$ such that $Q_i(x_i, \alpha, \mu) > 0$,

$$
\begin{aligned}
& P_i\{\hat{\delta}_{ki}(\underset{\sim}{X}) = 1 | X_i = x_i\} \\
&= P_i\{Q_i(x_i, \hat{\alpha}, \hat{\mu}) < 0 | X_i = x_i\} \\
&\le P_i\{|\hat{\alpha} - \alpha| > q_i(x_i, \alpha, \mu) | X_i = x_i\} + P_i\{|\hat{\mu} - \mu| > q_i(x_i, \alpha, \mu) | X_i = x_i\}
\end{aligned}
$$

for some $q_i(x_i, \alpha, \mu) > 0$.

Therefore, it suffices to investigate the behavior of

$$
P_i\{|\hat{\alpha} - \alpha| > q_i(x, \alpha, \mu) | X_i = x_i\} \text{ and } P_i\{|\hat{\mu} - \mu| > q_i(x_i, \alpha, \mu) | X_i = x_i\}
$$

for each x_i and $i = 1, \ldots, k$. For this, we first present several useful preliminary lemmas.

Lemma 5.1

(a) For positive c such that $\frac{c}{2} > |x_i^2 - \mu_{i2}|$,

$$
P_i\left\{\sum_{j=1}^{k}(X_j^2 - \mu_{j2}) < -c | X_i = x_i\right\} \le \exp\left\{-c^2 \Big/ \left(2\sum_{\substack{j=1\\ j\ne i}}^{k} m_j^4\right)\right\},
$$

$$
P_i\left\{\sum_{j=1}^{k}(X_j^2 - \mu_{j2}) > c | X_i = x_i\right\} \le \exp\left\{-c^2 \Big/ \left(2\sum_{\substack{j=1\\ j\ne i}}^{k} m_j^4\right)\right\}.
$$

(b) For positive c such that $c\sum_{j=1}^{k} m_j/2 > |m_i\mu - x_i|$,

$$P_i\left\{\hat{\mu}-\mu < -c|X_i = x_i\right\} \le \exp\left\{-\left(c\sum_{j=1}^{k} m_j\right)^2 \Big/ \left(2\sum_{\substack{j=1\\ j\neq i}}^{k} m_j^2\right)\right\},$$

$$P_i\left\{\hat{\mu}-\mu > c|X_i = x_i\right\} \le \exp\left\{-\left(c\sum_{j=1}^{k} m_j\right)^2 \Big/ \left(2\sum_{\substack{j=1\\ j\neq i}}^{k} m_j^2\right)\right\}.$$

(c) For positive c such that $c\sum_{j=1}^{k} m_j > 4|m_i\mu - x_i|$,

$$P_i\left\{\hat{\mu}^2-\mu^2 < -c|X_i = x_i\right\} \le \exp\left\{-\left(c\sum_{j=1}^{k} m_j\right)^2 \Big/ \left(8\sum_{\substack{j=1\\ j\neq i}}^{k} m_j^2\right)\right\},$$

$$P_i\left\{\hat{\mu}^2-\mu^2 > c|X_i = x_i\right\} \le \exp\left\{-\left(c\sum_{j=1}^{k} m_j\right)^2 \Big/ \left(8\sum_{\substack{j=1\\ j\neq i}}^{k} m_j^2\right)\right\}.$$

Proof: Straightforward computation and application of Theorem 2 of Hoeffding (1963) will yield the results. Details are omitted here.

The following lemma is a direct consequence of Lemma 5.1(b).

Lemma 5.2. Suppose that $m_i \ge m_* > 1$ for all $i = 1, \ldots, k$, where the value of the bound m_* is independent of k. Then, for sufficiently large k,

$$P_i\{|\hat{\mu}-\mu| > q_i(x_i,\alpha,\mu)|X_i = x_i\} \le \exp\{-k\tau_1\},$$

where $\tau_1 = \frac{m_*^2}{2M^{*2}} \min_{1\le j\le k} \min_{A_j} q_j^2(x_j,\alpha,\mu) > 0$, and

$$A_j = \{x_j | 0 \le x_j \le d_{j0} - 1,\ q_j(x_j,\alpha,\mu) > 0\}.$$

Lemma 5.3. Suppose that $1 < m_* \le m_i \le M_i \le M^*$ for all $i = 1, \ldots, k$, where the values of the bounds m_* and M^* are independent of k. Then for sufficiently large k,

$$P_i\{\hat{\alpha} = \infty | X_i = x_i\} = O(\exp(-\tau_2 k))$$

for some positive value τ_2 defined below.

Proof: By the definition of $\hat{\alpha}$ and an application of Bonferroni's inequality,

$$P_i\{\hat{\alpha} = \infty | X_i = x_i\} \leq P_i\left\{\sum_{j=1}^{k}(X_j^2 - \mu_{j2}) \leq -c(k) | X_i = x_i\right\}$$
$$+ P_i\left\{(\hat{\mu} - \mu)\sum_{j=1}^{k} m_j \geq c(k) | X_i = x_i\right\}$$
$$+ P_i\left\{(\hat{\mu}^2 - \mu^2)\sum_{j=1}^{k}(m_j^2 - m_j) \geq c(k) | X_i = x_i\right\},$$

where

$$\begin{aligned} c(k) &= \sum_{j=1}^{k}[\mu_{j2} - \mu^2(m_j^2 - m_j) - \mu m_j]/3 \\ &= \sum_{j=1}^{k}[\mu m_j^2 - \mu_{j2}]/(3\alpha), \end{aligned} \tag{5.11}$$

which is obtained from (4.3). Also,

$$\begin{aligned} \sum_{j=1}^{k}[\mu m_j^2 - \mu_{j2}] &= \sum_{j=1}^{k} E[X_j(m_j - X_j)] \\ &\geq \sum_{j=1}^{k}(m_j - 1)[1 - f_j(0|\alpha, \mu) - f_j(m_j|\alpha, \mu)] \\ &\geq (m_* - 1)k\varepsilon(\alpha, \mu), \end{aligned} \tag{5.12}$$

since $\min_{1 \leq j \leq k}[1 - f_j(0|\alpha, \mu) - f_j(m_j|\alpha, \mu)] \geq \varepsilon(\alpha, \mu)$ for some positive value $\varepsilon(\alpha, \mu)$ for all k which is guaranteed by our assumption that the bound m_* and M^* are independent of k.

By (5.11) and (5.12), $c(k)$ tends to infinity as k tends to infinity. Therefore, for sufficiently large k, by Lemma 5.1, we have,

$$P_i\left\{\sum_{j=1}^{k}(X_j^2 - \mu_{j2}) \leq -c(k) | X_i = x_i\right\} \leq \exp\left\{-c^2(k) \Big/ \left(2\sum_{\substack{j=1 \\ j \neq i}}^{k} m_j^4\right)\right\}, \tag{5.13}$$

$$P_i\left\{(\hat{\mu} - \mu)\sum_{j=1}^{k} m_j \geq c(k) | X_i = x_i\right\} \leq \exp\left\{-c^2(k) \Big/ \left(2\sum_{\substack{j=1 \\ j \neq i}}^{k} m_j^2\right)\right\}, \tag{5.14}$$

$$P_i\left\{(\hat{\mu}^2-\mu^2)\sum_{j=1}^{k}(m_j^2-m_j)\geq c(k)|X_i=x_i\right\}$$

$$\leq \exp\left\{-c^2(k)\Big/\left[8M^{*^2}\sum_{\substack{j=1\\ j\neq i}}^{k}m_j^2\right]\right\}. \tag{5.15}$$

Let $\tau_2=(m_*-1)^2\varepsilon^2(\alpha,\mu)/[72\alpha^2M^{*4}]$. Then, by (5.11)–(5.15) and the definition of τ_2, for sufficiently large k,

$$P_i\{\hat{\alpha}=\infty|X_i=x_i\}=O(\exp(-\tau_2 k)).$$ □

Note that this rate is independent of x_i and i for all $i=1,\ldots,k$.

Lemma 5.4. Under the assumption of Lemma 5.3,

$$P\{\hat{\alpha}-\alpha>q_i(x_i,\alpha,\mu),\hat{\alpha}<\infty|X_i=x_i\}\leq\exp\{-\tau_3 k\},\text{ and}$$
$$P\{\hat{\alpha}-\alpha<-q_i(x_i,\alpha,\mu),\hat{\alpha}<\infty|X_i=x_i\}\leq\exp\{-\tau_3 k\}$$

for some positive number τ_3 defined below.

Proof: In the following, we let $q_i=q_i(x_i,\alpha,\mu)$. By the definition of $\hat{\alpha}$ and (4.3),

$$P_i\{\hat{\alpha}-\alpha>q_i,\hat{\alpha}<\infty|X_i=x_i\}$$
$$\leq P_i\left\{(\hat{\mu}-\mu)\left[\sum_{j=1}^{k}m_j^2+(\alpha+q_i)\sum_{j=1}^{k}m_j\right]+(\alpha+q_i)(\hat{\mu}^2-\mu^2)\sum_{j=1}^{k}(m_j^2-m_j)\right.$$
$$\left.-(\alpha+q_i+1)\sum_{j=1}^{k}(X_j^2-\mu_{j2})>3q_i\ c(k)|X_i=x_i\right\}$$
$$\leq P_i\left\{(\hat{\mu}-\mu)\sum_{j=1}^{k}(m_j^2+(\alpha+q_i)m_j)>q_i\ c(k)|X_i=x_i\right\} \tag{5.16}$$
$$+P_i\left\{(\hat{\mu}^2-\mu^2)(\alpha+q_i)\sum_{j=1}^{k}(m_j^2-m_j)>q_i\ c(k)|X_i=x_i\right\}$$
$$+P_i\left\{(\alpha+q_i+1)\sum_{j=1}^{k}(X_j^2-\mu_{j2})<-q_i\ c(k)|X_i=x_i\right\}.$$

By Lemma 5.1, for sufficiently large k, we have

$$P_i\left\{(\hat{\mu}-\mu)\sum_{j=1}^{k}[m_j^2+(\alpha+q_i)m_j] > q_i\ c(k)|X_i=x_i\right\}$$

$$\leq \exp\left\{-\left[q_i\ c(k)\sum_{j=1}^{k}m_j\right]^2 \Bigg/ \left(2\left[\sum_{j=1}^{k}(m_j^2+(\alpha+q_i)m_j)\right]^2\sum_{\substack{j=1\\ j\neq i}}^{k}m_j^2\right)\right\}, \quad (5.17)$$

$$P_i\left\{(\hat{\mu}^2-\mu^2)(\alpha+q_i)\sum_{j=1}^{k}(m_j^2-m_j) > q_i\ c(k)|X_i=x_i\right\}$$

$$\leq \exp\left\{-\left[q_i\ c(k)\sum_{j=1}^{k}m_j\right]^2 \Bigg/ \left(8\left[(\alpha+q_i)\sum_{j=1}^{k}(m_j^2-m_j)\right]^2\sum_{\substack{j=1\\ j\neq i}}^{k}m_j^2\right)\right\}, \quad (5.18)$$

and

$$P_i\left\{(\alpha+q_i+1)\sum_{j=1}^{k}(X_j^2-\mu_{j2}) < -q_i\ c(k)|X_i=x_i\right\}$$

$$\leq \exp\left\{-[q_i\ c(k)]^2 \Bigg/ \left(2(\alpha+q_i+1)^2\sum_{\substack{j=1\\ j\neq i}}^{k}m_j^4\right)\right\}. \quad (5.19)$$

Let $\tau_3 = \frac{(m_*-1)^2\varepsilon^2(\alpha,\mu)m_*^2}{72\alpha^2M^{*6}}\min_{1\leq i\leq k}\min_{x\in A_i}\left[\frac{q_i(x,\alpha,\mu)}{\alpha+q_i(x,\alpha,\mu)+1}\right]^2$, where $A_i=\{x|0\leq x\leq d_{i0}-1,\ q_i(x,\alpha,\mu)>0\}$. Hence $\tau_3>0$. From the definition of τ_3 and (5.16)–(5.19), we conclude that

$$P_i\{\hat{\alpha}-\alpha>q_i,\ \hat{\alpha}<\infty|X_i=x_i\}\leq\exp\{-\tau_3 k\}.$$

Note that the rate is independent of x_i and i.

Part (b) can be proved in a similar way and hence the proof is omitted. □

Note that when the loss functions $L_{ij}(\cdot), j=0,1$, which are bounded above by L^*, depend on i only through the values of m_i and M_i, and the values of the bounds L^*, m_* and M^* are independent of k, τ_1 and τ_3 are always positive and min $(\tau_1,\tau_3)\geq\tau(\alpha,\mu)$ for all k for some positive value $\tau(\alpha,\mu)$.

We summarize the results in this subsection in the following theorem:

Theorem 5.2. Suppose that $1 < m_* \le m_i \le M_i \le M^*$ for all $i = 1, \ldots, k$, and $L_{ij}(d_i) \le L^*$ for all $j = 0, 1$ and $i = 1, \ldots, k$, where $L_{ij}(\cdot), j = 0, 1$, depend on i only through the values of m_i and M_i; and where the values of the bounds L^*, m_* and M^* are independent of k. Then, under the statistical model of Case 2, for each pair of the values (α, μ), $0 < \alpha < \infty$, $0 < \mu < 1$, the empirical Bayes selection rule $\hat{\underset{\sim}{\delta}}_k$ is asymptotically optimal with rate of convergence $O(\exp(-\tau k + \ln k))$, where $\tau = \min(\tau_1, \tau_2, \tau_3) > 0$.

6. Small Sample Performance: Simulation Studies

Monte Carlo studies have been carried out to investigate the small sample performance of the proposed empirical Bayes selection rules $\underset{\sim}{\delta}_k^*$ and $\hat{\underset{\sim}{\delta}}_k$, respectively. In these Monte Carlo studies, we have assumed that

$$\begin{cases} M_1 = \ldots = M_k = M, \\ m_1 = \ldots = m_k = m, \\ d_{10} = \ldots = d_{k0} = d_0. \\ L_{1j}(\cdot) = L_{2j}(\cdot) = \ldots = L_{kj}(\cdot), j = 0, 1. \end{cases} \tag{6.1}$$

Under the preceding assumption and the statistical model of Case 1, for the Bayes selection rule $\underset{\sim}{\delta}_k^\theta = (\delta_{k1}^\theta, \ldots, \delta_{kk}^\theta)$

$$r_{k1}(\theta, \delta_{k1}^\theta) = \ldots = r_{kk}(\theta, \delta_{kk}^\theta) \text{ and } r_k(\theta, \underset{\sim}{\delta}_k^\theta) = k r_{kk}(\theta, \delta_{kk}^\theta).$$

Also, for the parametric empirical Bayes selection rule $\underset{\sim}{\delta}_k^* = (\delta_{k1}^*, \ldots, \delta_{kk}^*)$, it can be seen that

$$r_{k1}(\theta, \delta_{k1}^*) = \ldots = r_{kk}(\theta, \delta_{kk}^*) \text{ and therefore } r_k(\theta, \underset{\sim}{\delta}_k^*) = k r_{kk}(\theta, \delta_{kk}^*).$$

Similarly, under the assumption (6.1), and the statistical model of Case 2, for the hierarchical Bayes selection rule $\underset{\sim}{\delta}_k^{\text{HB}} = (\delta_{k1}^{\text{HB}}, \ldots, \delta_{kk}^{\text{HB}})$,

$$R_{k1}(\alpha, \mu, \delta_{k1}^{\text{HB}}) = \ldots = R_{kk}(\alpha, \mu, \delta_{kk}^{\text{HB}}) \text{ and } R_k(\alpha, \mu, \underset{\sim}{\delta}_k^{\text{HB}}) = k R_{kk}(\alpha, \mu, \delta_{kk}^{\text{HB}}).$$

Also, for the hierarchical empirical Bayes selection rule $\hat{\underset{\sim}{\delta}}_k = (\hat{\delta}_{k1}, \ldots, \hat{\delta}_{kk})$, we have:

$$R_{k1}(\alpha, \mu, \hat{\delta}_{k1}) = \ldots = R_{kk}(\alpha, \mu, \hat{\delta}_{kk}) \text{ and } R_k(\alpha, \mu, \hat{\underset{\sim}{\delta}}_k) = k R_{kk}(\alpha, \mu, \hat{\delta}_{kk}).$$

Therefore, in the following, we simulated the differences

$$r_{kk}(\theta, \delta^*_{kk}) - r_{kk}(\theta, \delta^\theta_{kk}) \text{ and } R_{kk}(\alpha, \mu, \hat{\delta}_{kk}) - R_{kk}(\alpha, \mu, \delta^{\text{HB}}_{kk}),$$

and used $k[r_{kk}(\theta, \delta^*_{kk}) - r_{kk}(\theta, \delta^\theta_{kk})]$ to estimate $r_k(\theta, \underset{\sim}{\delta}^*_k) - r_k(\theta, \underset{\sim}{\delta}^\theta_k)$ and used $k[R_{kk}(\alpha, \mu, \hat{\delta}_{kk}) - R_{kk}(\alpha, \mu, \delta^{\text{HB}}_{kk})]$ to estimate $R_k(\alpha, \mu, \hat{\underset{\sim}{\delta}}_k) - R_k(\alpha, \mu, \underset{\sim}{\delta}^{\text{HB}}_k)$, respectively.

We used the linear loss of (3.12) as the loss function. The simulation scheme used in this paper is described as follows:

Case 1. The Parameter θ Being Fixed

(1) For any fixed value of the parameter θ and a given value of m, generate $k-1$ independent random numbers, say, $X_1, \ldots, X_{k-1}$ from a $B(m, \theta)$ distribution.

(2) Let x_k be an observed value from a $B(m, \theta)$ distribution. Use $X_1, \ldots, X_{k-1}$ and x_k to estimate θ and construct the parametric empirical Bayes selection rule δ^*_{kk} according to (4.2). Then, we computed the conditional regret Bayes risk of δ^*_{kk} (conditional on $X_1, \ldots, X_{k-1}$) by

$$D^\theta_k(X_1, \ldots, X_{k-1}) = r_{kk}(\theta, \delta^*_{kk} | X_1, \ldots, X_{k-1}) - r_{kk}(\theta, \delta^\theta_{kk}).$$

(3) The above process was repeated 500 times. The average of the regret Bayes risk based on the 500 repetitions denoted by $\overline{D}^\theta_k$ was used as an estimator of the regret Bayes risk $r_{kk}(\theta, \delta^*_{kk}) - r_{kk}(\theta, \delta^\theta_{kk})$. Then, we used $k\overline{D}^\theta_k$ as as estimator of the total regret risk $r_k(\theta, \underset{\sim}{\delta}^*_k) - r_k(\theta, \underset{\sim}{\delta}^\theta_k)$.

A summary of the simulated results is given in Table 1.

Case 2. The Parameter θ with a Prior Distribution Beta $(\alpha\mu, \alpha(1-\mu))$

(1) For given values of α and μ, we generated $k-1$ random numbers from a distribution having the probability function $f(x|\alpha, \mu)$,

$$f(x|\alpha, \mu) = \int_0^1 f(x|\theta) h(\theta|\alpha, \mu) d\theta,$$

where $f(x|\theta) = \binom{m}{x}\theta^x(1-\theta)^{m-x},\ 0 < \theta < 1,\ x = 0, 1, \ldots, m$, and where $h(\theta|\alpha, \mu)$ is defined in Section 2.

Then we followed steps (2) and (3) analogous to steps (2) and (3) of Case 1, by just replacing the Bayes risks by the corresponding Bayes risks of the hierarchical empirical Bayes selection rule $\hat{\delta}_{kk}$ and the hierarchical Bayes selection rule δ_{kk}^{HB}, respectively. We denote the average of the conditional regret Bayes risk based on 500 repetitions by $\overline{D}_k^{\alpha\mu}$. $\overline{D}_k^{\alpha\mu}$ was used to estimate the regret Bayes risk $R_{kk}(\alpha, \mu, \hat{\delta}_{kk}) - R_{kk}(\alpha, \mu, \delta_{kk}^{\text{HB}})$. Then, we used $k\overline{D}_k^{\alpha\mu}$ to estimate the total regret Bayes risk $R_k(\alpha, \mu, \underset{\sim}{\hat{\delta}}_k) - R_k(\alpha, \mu, \underset{\sim}{\delta}_k^{\text{HB}})$.

The results of these simulations are reported in Table 2 and Table 3, respectively.

The simulated results indicate that in each of the two models, the total regret risks ($k\overline{D}_k^{\theta}$ in Case 1 model and $k\overline{D}_k^{\alpha\mu}$ in Case 2 model) converge to zero as the value of k becomes large. Of course, the rates of convergence vary according to the models. In Case 1 model, for small values of k, $k\overline{D}_k^{\theta}$ is small and $k\overline{D}_k^{\theta}$ tends to zero gradually. While in Case 2 model, though for small values of k, the $k\overline{D}_k^{\alpha\mu}$ values are larger (compared with $k\overline{D}_k^{\theta}$), however, the rate of convergence of $k\overline{D}_k^{\alpha\mu}$ to zero is very fast, as for example, in Table 3, $k\overline{D}_k^{\alpha\mu} = 0$ for $k \geq 180$.

Acknowledgement: We are thankful to two referees and the Associate Editor for very useful comments and suggestions which improved the presentation of results in this paper.

Table 1. The Small Sample Performance of $\underset{\sim}{\delta}_k^*$

$M = 100,\ m = 20,\ d_0 = 6$, and $\theta = 0.02$.

k	$\overline{D}_k^\theta$	$k\overline{D}_k^\theta$	$SE(\overline{D}_k^\theta)$
10	2.9878×10^{-3}	29.8785×10^{-3}	0.6169×10^{-3}
20	0.5254×10^{-3}	10.5073×10^{-3}	0.0822×10^{-3}
30	0.1811×10^{-3}	5.4324×10^{-3}	0.0369×10^{-3}
40	0.1097×10^{-3}	4.3891×10^{-3}	0.0190×10^{-3}
50	0.0719×10^{-3}	3.5932×10^{-3}	0.0047×10^{-3}
60	0.0774×10^{-3}	4.6418×10^{-3}	0.0190×10^{-3}
70	0.0539×10^{-3}	3.7705×10^{-3}	0.0043×10^{-3}
80	0.0386×10^{-3}	3.0882×10^{-3}	0.0038×10^{-3}
90	0.0359×10^{-3}	3.2319×10^{-3}	0.0037×10^{-3}
100	0.0310×10^{-3}	3.0972×10^{-3}	0.0035×10^{-3}
120	0.0269×10^{-3}	3.2319×10^{-3}	0.0033×10^{-3}
140	0.0162×10^{-3}	2.2623×10^{-3}	0.0026×10^{-3}
160	0.0148×10^{-3}	2.3700×10^{-3}	0.0025×10^{-3}
180	0.0076×10^{-3}	1.3755×10^{-3}	0.0018×10^{-3}
200	0.0108×10^{-3}	2.1546×10^{-3}	0.0021×10^{-3}
250	0.0022×10^{-3}	0.5611×10^{-3}	0.0010×10^{-3}
300	0.0031×10^{-3}	0.9426×10^{-3}	0.0012×10^{-3}
350	0.0004×10^{-3}	0.1571×10^{-3}	0.0004×10^{-3}
400	0.0004×10^{-3}	0.1795×10^{-3}	0.0004×10^{-3}

Table 2. The Small Sample Performance of $\underset{\sim}{\delta}_k^{\mathrm{HB}}$

$M = 100,\ m = 20,\ d_0 = 6,\ \alpha = 10$ and $\mu = 0.02.$

k	$\overline{D}_k^{\alpha\mu}$	$k\overline{D}_k^{\alpha\mu}$	$SE(\overline{D}_k^{\alpha\mu})$
10	36.7345×10^{-3}	367.3446×10^{-3}	1.8780×10^{-3}
20	30.8095×10^{-3}	616.1903×10^{-3}	1.8231×10^{-3}
30	26.0696×10^{-3}	782.0877×10^{-3}	1.7493×10^{-3}
40	17.4362×10^{-3}	697.4465×10^{-3}	1.5324×10^{-3}
50	17.0976×10^{-3}	854.8797×10^{-3}	1.5213×10^{-3}
60	11.6805×10^{-3}	700.8327×10^{-3}	1.3069×10^{-3}
70	9.3106×10^{-3}	651.7408×10^{-3}	1.1856×10^{-3}
80	9.8963×10^{-3}	791.7044×10^{-3}	1.2592×10^{-3}
90	6.7713×10^{-3}	609.4202×10^{-3}	1.0280×10^{-3}
100	8.6335×10^{-3}	863.3451×10^{-3}	1.1468×10^{-3}
120	3.8935×10^{-3}	467.2222×10^{-3}	0.7938×10^{-3}
140	3.5550×10^{-3}	497.6932×10^{-3}	0.7600×10^{-3}
160	2.2009×10^{-3}	352.1094×10^{-3}	0.6030×10^{-3}
180	0.8464×10^{-3}	152.3550×10^{-3}	0.3770×10^{-3}
200	0.5079×10^{-3}	101.5700×10^{-3}	0.2926×10^{-3}
250	0.3386×10^{-3}	84.6417×10^{-3}	0.2392×10^{-3}
300	0.1693×10^{-3}	50.7850×10^{-3}	0.1693×10^{-3}
350	0.	0.	0.
400	0.	0.	0.

the entry 0 means that the exact value is less than 10^{-7}.

Table 3. The Small Sample Performance of $\underset{\sim}{\delta}_k^{\mathrm{HB}}$

$M = 100,\ m = 20,\ d_0 = 6,\ \alpha = 1$ and $\mu = 0.02$.

k	$\overline{D}_k^{\alpha\mu}$	$k\overline{D}_k^{\alpha\mu}$	$SE(\overline{D}_k^{\alpha\mu})$
10	4.4973×10^{-3}	44.9733×10^{-3}	0.5255×10^{-3}
20	4.6379×10^{-3}	92.7573×10^{-3}	0.5324×10^{-3}
30	3.8649×10^{-3}	115.9467×10^{-3}	0.4921×10^{-3}
40	3.3027×10^{-3}	132.1090×10^{-3}	0.4590×10^{-3}
50	2.1784×10^{-3}	108.9196×10^{-3}	0.3793×10^{-3}
60	1.9676×10^{-3}	118.0547×10^{-3}	0.3616×10^{-3}
70	1.1946×10^{-3}	83.6221×10^{-3}	0.2851×10^{-3}
80	0.5622×10^{-3}	44.9732×10^{-3}	0.1974×10^{-3}
90	0.7730×10^{-3}	69.5680×10^{-3}	0.2307×10^{-3}
100	0.3514×10^{-3}	35.1353×10^{-3}	0.1565×10^{-3}
120	0.2108×10^{-3}	25.2975×10^{-3}	0.1215×10^{-3}
140	0.2811×10^{-3}	39.3516×10^{-3}	0.1401×10^{-3}
160	0.0703×10^{-3}	11.2433×10^{-3}	0.0703×10^{-3}
180	0.	0.	0.
200	0.	0.	0.
250	0.	0.	0.
300	0.	0.	0.
350	0.	0.	0.
400	0.	0.	0.

the entry 0 means that the exact value is less that 10^{-7}.

References

Bartlett, N.S. and Govindarajulu, Z. (1970). Selecting a subset containing the best hypergeometric population. *Sankhya B* **32**, 341–352.

Booth, N.B. and Smith, A.F.M. (1976). Batch acceptance schemes based on an autoregressive prior. *Biometrika* **63**, 133–136.

Craig, J.A. and Bland, R.P. (1981). An empirical Bayes approach to variables acceptance sampling plan. *Commun. Statist.* **A(10)**, 2399–2410.

Gupta, S.S. and Liang, T. (1991). On the asymptotic optimality of certain empirical Bayes simultaneous testing procedures for Poisson populations. *Statistics & Decisions*, vol. **9**, 263–283.

Gupta, S.S. and Panchapakesan, S. (1979). *Multiple Decision Procedures.* Wiley, New York.

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