SELECTING GOOD EXPONENTIAL POPULATIONS COMPARED WITH A CONTROL: A NONPARAMETRIC EMPIRICAL BAYES APPROACH*

by
Shanti S. Gupta and Ta Chen Liang
Purdue University and Wayne State University

Technical Report #96-18C

Department of Statistics
Purdue University
West Lafayette, IN USA

July 1996

*This research was supported in part by US Army Research Office, Grant DAAH04-95-1-0165 at Purdue University.
SELECTING GOOD EXPONENTIAL POPULATIONS COMPARED WITH
A CONTROL: A NONPARAMETRIC EMPIRICAL BAYES APPROACH*

by
Shanti S. Gupta
Department of Statistics
Purdue University
West Lafayette, IN 47907-1399

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

Abstract

This paper deals with empirical Bayes selection procedures for selecting good exponential populations compared with a control. Based on the accumulated historical data, an empirical Bayes selection procedure $\hat{\delta}_n$ is constructed by mimicking the behavior of a Bayes selection procedure. The empirical Bayes selection procedure $\hat{\delta}_n$ is proved to be asymptotically optimal. The analysis shows that the rate of convergence of $\hat{\delta}_n$ is influenced by the tail probabilities of the underlying distributions. It is shown that under certain regularity conditions on the moments of the prior distribution, the empirical Bayes selection procedure $\hat{\delta}_n$ is asymptotically optimal of order $O(n^{-\lambda/2})$ for some $0 < \lambda \leq 2$. A lower bound with rate of convergence of order $O(n^{-1})$ is also established for the regret Bayes risk of the empirical Bayes selection procedure $\hat{\delta}_n$. This result suggests that a rate of order $O(n^{-1})$ might be the best possible rate of convergence for this empirical Bayes selection problem.

Short Title: Empirical Bayes Selection for Exponential Populations

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

Keywords and phrases: Asymptotic optimality, comparison with a control, empirical Bayes, good populations, rate of convergence, regret Bayes risk, selection procedure.

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

The exponential distribution has played an important role for modeling the life time distribution of a variety of random phenomena. This distribution arises in many areas of applications, including reliability, life-testing and survival analysis. An overall introduction and more applications of the exponential distribution model can be seen, for example, in the contents of Johnson, Kotz and Balakrishnan (1994) and Balakrishnan and Basu (1995).

Consider $k$ independent exponential populations $\pi_1, \ldots, \pi_k$, with associated population means $\theta_i, i = 1, \ldots, k$, respectively. The $\theta_i$'s are unknown. Let $\theta_0$ be a specified standard. Population $\pi_i$ is said to be good if $\theta_i \geq \theta_0$, and bad otherwise. In many practical situations, an experimenter is often confronted with the problem of comparing the $k$ alternatives with the specified standard $\theta_0$, and selecting the more promising subset of the $k$ populations for further experimentation. The problem is known as comparison with a control problem. A review of subset selection procedures in this context is contained in Gupta and Panchapakesan (1985).

Now, consider a situation in which one is repeatedly dealing with the same selection problem independently. In such instances, it is reasonable to formulate the component problem in the sequence as a Bayes decision problem with respect to an unknown prior distribution on the parameter space. One then draws useful information from the accumulated historical data to improve the decision at each stage. This is the empirical Bayes approach due to Robbins (1956, 1964). Empirical Bayes procedures have been derived for subset selection goals by Deely (1965). Recently, Gupta and Liang (1988, 1994) and Gupta, Liang and Rau (1994a, 1994b) have studied certain selection problems using the empirical Bayes approach. Many such empirical Bayes procedures have been shown to be asymptotically optimal in the sense that the Bayes risk for the $(n+1)-st$ decision problem converges to the optimal Bayes risk which would have been obtained if the prior distributions were fully known and the Bayes procedure with respect to this prior distribution were used.

In this paper, we are dealing with the problem of selecting good exponential populations compared with a control using the empirical Bayes approach. In Section 2, the selection problem is formulated and a Bayes selection procedure is derived. In Section
3, an empirical Bayes selection procedure $\delta_n^*$ is constructed by mimicking the behavior of the Bayes selection procedure. The asymptotic optimality of the empirical Bayes selection procedure is established in Section 4. The associated rate of convergence of the regret Bayes risk of $\delta_n^*$ is also investigated. The analysis shows that the rate of convergence is influenced by the tail probabilities of the underlying distributions. It is shown that under certain regularity conditions about the moments of the prior distribution, the empirical Bayes selection procedure $\delta_n^*$ is asymptotically optimal of order $O(n^{-\lambda/2})$ for some $0 < \lambda \leq 2$. A lower bound with rate of convergence of order $O(n^{-1})$ is also established for the regret Bayes risk of the empirical Bayes selection procedure $\delta_n^*$. This result suggests that a rate of order $O(n^{-1})$ might be the best possible rate of convergence for the empirical Bayes selection problem.

2. Formulation of the Selection Problem

Consider $k$ independent exponential populations $\pi_1, \ldots, \pi_k$, with probability density function $h_i(x_i|\theta_i) = \frac{1}{\theta_i} e^{-x_i/\theta_i}, x_i > 0, \theta_i > 0$, respectively. The $\theta_i's$, $i = 1, \ldots, k$, are unknown. For a specified standard $\theta_0 > 0$, population $\pi_i$ is said to be good if $\theta_i \geq \theta_0$, and bad otherwise. The selection goal is to select all good populations and to exclude all bad populations.

Let $\Omega = \{\theta = (\theta_1, \ldots, \theta_k)|\theta_i > 0, i = 1, \ldots, k\}$ be the parameter space and let $A = \{a = (a_1, \ldots, a_k)|a_i = 0, 1; i = 1, \ldots, k\}$ be the action space. When an action $a$ is taken, it means that population $\pi_i$ is selected as good if $a_i = 1$, and excluded as bad if $a_i = 0$. For parameter $\theta$ and action $a$, the loss function $L(\theta, a)$ is defined as:

$$L(\theta, a) = \sum_{i=1}^{k} \ell(\theta_i, a_i)$$

(2.1)

where

$$\ell(\theta_i, a_i) = \begin{cases} a_i \theta_i (\theta_0 - \theta_i) I(\theta_0 > \theta_i) + (1 - a_i) \theta_i (\theta_i - \theta_0) I(\theta_i \geq \theta_0) \\ \text{and } I(A) \text{ is the indicator function of the event } A. 
\end{cases}$$

For each $i = 1, \ldots, k$, let $X_{i1}, \ldots, X_{im}$ be a sample of size $m$ arising from population $\pi_i$ with probability density $h_i(x|\theta_i)$. Let $Y_i = \sum_{j=1}^{m} X_{ij}$. Then $Y_i$ follows a gamma distribution with probability density $f_i(y|\theta_i) = \frac{y^{m-1}}{\Gamma(m)\theta_i^m} e^{-y/\theta_i}$. Note that $Y_i$ is a sufficient statistic.
for the parameter $\theta_i$. Since the Bayes and empirical Bayes approaches will be employed, it suffices to deal with the sufficient statistics $Y_1, \ldots, Y_k$. It is assumed that for each $i = 1, \ldots, k$, the parameter $\theta_i$ is a realization of a random parameter $\Theta_i$ with an unknown prior distribution on $G_i$ on $\theta_i$ over $(0, \infty)$, and $\Theta_1, \ldots, \Theta_k$ are mutually independent. It is also assumed that $G_i, i = 1, \ldots, k$, are non-degenerate and satisfy that $\int \theta^2 dG_i(\theta) < \infty$ to insure the Bayes risk to be finite and this selection problem to be meaningful. We let $G(\theta) = \prod_{i=1}^{k} G_i(\theta_i)$.

Let $Y = (Y_1, \ldots, Y_k)$ and $\mathcal{Y}$ denote the sample space of $Y$. A selection procedure $\hat{\delta} = (\delta_1, \ldots, \delta_k)$ is defined to be a measurable mapping from the sample space $\mathcal{Y}$ into the product space $[0, 1]^k$, so that for each $y \in \mathcal{Y}$, $\hat{\delta}(y) = (\delta_1(y), \ldots, \delta_k(y))$ and $\delta_i(y)$ is the probability of selecting population $\pi_i$ as good. Let $\mathcal{C}$ be the class of all selection procedures. For each $\hat{\delta} \in \mathcal{C}$, let $R(G, \hat{\delta})$ denote its associated Bayes risk. Then, $R(G) = \inf_{\hat{\delta} \in \mathcal{C}} R(G, \hat{\delta})$ is the minimum Bayes risk among the class $\mathcal{C}$. A selection procedure $\hat{\delta}_G$ satisfying $R(G, \hat{\delta}_G) = R(G)$ is called a Bayes selection procedure. Note that $R(G) < \infty$ under the assumption that $\int \theta^2 dG_i(\theta) < \infty$, $i = 1, \ldots, k$.

Let $c(\theta) = (\Gamma(m)\theta^m)^{-1}, u(y) = y^{m-1}$. Then $f_i(y|\theta_i) = c(\theta_i)u(y)e^{-y/\theta_i}$. Let $f(y|\theta) = \prod_{i=1}^{k} f_i(y_i|\theta_i)$. Also, for each $i = 1, \ldots, k, y > 0$ and nonnegative integer $a$, define

$$
\psi_{ia}(y) = \int_{\theta=0}^{\infty} \theta^a c(\theta)e^{-y/\theta} dG_i(\theta). \tag{2.3}
$$

Then, $f_i(y) = \int f_i(y|\theta) dG_i(\theta) = u(y)\psi_{io}(y)$ is the marginal probability density of $Y_i$. Let $f(y) = \prod_{i=1}^{k} f_i(y_i)$.

From the preceding statistical model and the loss function $L(\theta, a)$, the Bayes risk associated with the selection procedure $\hat{\delta}$ is:

$$
R(G, \hat{\delta}) = \sum_{i=1}^{k} R_i(G, \delta_i) \tag{2.4}
$$

and

$$
R_i(G, \delta_i) = \int \left[ \prod_{j=1}^{k} f_j(y_j)|u(y_i)\delta_i(y)\psi_{i1}(y_i) - \psi_{i2}(y_i) \right] dy + C_i \tag{2.5}
$$

$$
= \int \left[ \prod_{j=1}^{k} f_j(y_j)|u(y_i)\delta_i(y)\psi_{i1}(y_i)(\theta_0 - \varphi_i(y_i)) \right] dy + C_i,
$$
where \( C_i = \int_{\Omega} \theta_i(\theta_i - \theta_0)I(\theta_i > \theta_0)dG(\theta) \) which is independent of the selection procedure \( \hat{\delta} \), and \( \varphi_i(y_i) = \frac{\psi_{i2}(y_i)}{\psi_{i1}(y_i)} \).

Let \( H_i(y_i) = \theta_0\psi_{i1}(y_i) - \psi_{i2}(y_i) = \psi_{i1}(y_i)[\theta_0 - \varphi_i(y_i)] \). Note that \( \psi_{i1}(y_i) > 0 \). From (2.5), a Bayes selection procedure \( \delta_G = (\delta_{G1}, \ldots, \delta_{Gk}) \) can be obtained as follows: For each \( y \in \mathcal{Y} \) and \( i = 1, \ldots, k \),

\[
\delta_{Gi}(y) = \begin{cases} 
1 & \text{if } H_i(y_i) \leq 0, \\
0 & \text{otherwise},
\end{cases}
\]

\[
= \begin{cases} 
1 & \text{if } \varphi_i(y_i) \geq \theta_0, \\
0 & \text{otherwise}.
\end{cases}
\] (2.6)

Note that for each component \( i, \delta_{Gi} \) depends on \( y \) only through \( y_i \). Therefore \( \delta_{Gi}(y) \) can be written as \( \delta_{Gi}(y_i) \). The minimum Bayes risk is

\[
R(G, \delta_G) = \sum_{i=1}^{k} R_i(G, \delta_{Gi})
\] (2.7)

and

\[
R_i(G, \delta_{Gi}) = \int_{y=0}^{\infty} u(y)\delta_{Gi}(y)\psi_{i1}(y)[\theta_0 - \varphi_i(y)]dy + C_i
\]

\[
= \int_{y=0}^{\infty} u(y)\delta_{Gi}(y)H_i(y)dy + C_i.
\] (2.8)

Note that \( \varphi_i(y) = \frac{\int_{\theta_i}^{\theta_0} \frac{\theta^2 e^y}{e^{\theta y}dG_i(\theta)}}{\int_{\theta_i}^{\theta_0} e^{-y/\theta}dG_i(\theta)} \) is continuous and strictly increasing in \( y \) since the prior distribution \( G_i \) is non-degenerate. We assume

**Assumption A** \( \lim_{y \to 0} \varphi_i(y) < \theta_0 < \lim_{y \to \infty} \varphi_i(y), i = 1, \ldots, k. \)

Under Assumption A, for each \( i = 1, \ldots, k \), there exists a unique value \( a_i \equiv a_i(\theta_0) \) such that \( \varphi_i(a_i) = \theta_0, \varphi_i(y) < \theta_0 \) if \( y < a_i, \varphi_i(y) > \theta_0 \) if \( y > a_i \). Hence, the Bayes selection procedure \( \delta_G = (\delta_{G1}, \ldots, \delta_{Gk}) \) can be represented as: For each \( i = 1, \ldots, k \) and \( y > 0 \)

\[
\delta_{Gi}(y) = \begin{cases} 
1 & \text{if } y \geq a_i, \\
0 & \text{otherwise}.
\end{cases}
\] (2.9)
Finally, for each $i = 1, \ldots, k$, let $dG_i^*(\theta) = \frac{\theta dG_i(\theta)}{m_1(G_i)}$, where $m_1(G_i) = \int \theta dG_i(\theta)$. Then, $G_i^*$ is also a distribution on $\theta_i$ over $(0, \infty)$. Define
\[
f_i^*(y) = \int f_i(y|\theta)dG_i^*(\theta) = u(y)\psi_{i1}(y)/m_1(G_i).
\] (2.10)
Then, $\varphi_i(y) = \frac{\psi_{i2}(y)}{\psi_{i1}(y)} = E_{G_i^*} [\Theta_i|Y_i = y]$; the "posterior mean" of $\Theta_i$ given $Y_i = y$ and $G_i^*$ is the prior distribution of $\Theta_i$. Hence, $R_i(G, \delta_{G_i})$ can be represented as:
\[
R_i(G, \delta_{G_i}) = \int_{y=0}^{\infty} m_1(G_i)f_i^*(y)\delta_{G_i}(y)\{\theta_0 - \varphi_i(y)\}dy + C_i.
\] (2.11)
Note that the Bayes selection procedure $\delta_G$ depends on the prior distribution $G$. Since $G$ is unknown, it is not possible to implement the Bayes selection procedure $\delta_G$ for the selection problem at hand. In the following, the empirical Bayes approach is employed.

3. Construction of An Empirical Bayes Selection Procedure

3.1 Empirical Bayes Framework

The empirical Bayes framework of the selection problem is given as follows.

For each $i = 1, \ldots, k$, at stage $\ell$, let $(Y_{i\ell}, \Theta_{i\ell})$ denote a pair of random vector so that $Y_{i\ell}$ is observable, but $\Theta_{i\ell}$ is not observable. Also, given $\Theta_{i\ell} = \theta_{i\ell}, Y_{i\ell}$ follows as a gamma distribution with probability density $f_1(y|\theta_{i\ell})$ and $\Theta_{i\ell}$ has a prior distribution $G_1$. It is assumed that $(Y_{i\ell}, \Theta_{i\ell}), i = 1, \ldots, k; \ell = 1, 2, \ldots$ are mutually independent.

At the present stage $n + 1$, we let $Y_i(n) = (Y_{i1}, \ldots, Y_{in})$ denote the historical data and $Y_i = Y_{i,n+1}$ the present random observation associated with population $\pi_i, i = 1, \ldots, k$. Let $Y(n) = (Y_1(n), \ldots, Y_k(n))$ and $Y = (Y_1, \ldots, Y_k)$. At the present stage $n + 1$, we consider the problem of selecting all good from among $(\theta_{1,n+1}, \ldots, \theta_{k,n+1})$ compared with the standard value $\theta_0$ using the loss function $L(\theta_{n+1}, q)$, where $\theta_{n+1} = (\theta_{1,n+1}, \ldots, \theta_{k,n+1})$.

At stage $n + 1$, an empirical Bayes selection procedure, say $\delta_n = (\delta_{n1}, \ldots, \delta_{nk})$, is a measurable function defined on the sample space of $Y \times Y(n)$, into the product space $[0, 1]^k$, so that $\delta_n(y, Y(n)) \equiv (\delta_{n1}(y, Y(n)), \ldots, \delta_{nk}(y, Y(n))) \equiv (\delta_{n1}(y), \ldots, \delta_{nk}(y))$. $\delta_n(y)$ is the probability of selecting $\pi_i$ as good.

Let $R(G, \delta_n|Y(n))$ denote the conditional Bayes risk of the empirical Bayes selection procedure $\delta_n$ conditioning on $Y(n)$, and let $R(G, \delta_n)$ denote the overall Bayes risk of the selection procedure $\delta_n$. Then,
\[
\begin{aligned}
R(G, \delta_n|Y(n)) &= \sum_{i=1}^{k} R_i(G, \delta_{ni}|Y(n)) \\
R_i(G, \delta_{ni}|Y(n)) &= \int_Y \left[ \prod_{j=1}^{k} f_j(y_j) \right] u(y_i) \delta_{ni}(y) H_i(y_i) dy + C_i
\end{aligned}
\] (3.1)

and

\[
\begin{aligned}
R(G, \hat{\delta}_n) &= \sum_{i=1}^{k} R_i(G, \delta_{ni}) \\
R_i(G, \delta_{ni}) &= \int_Y \left[ \prod_{j=1}^{k} f_j(y_j) \right] u(y_i) E_{Y(n)}[\delta_{ni}(y)] H_i(y_i) dy + C_i
\end{aligned}
\] (3.2)

where the expectation \( E_{Y(n)} \) is taken with respect to the probability measure generated by \( Y(n) \).

Since \( \hat{\delta}_G \) is the Bayes selection procedure, \( R_i(G, \delta_{ni}|Y(n)) - R_i(G, \delta_{Gi}) \geq 0 \) for all \( Y(n), n \) and each \( i = 1, \ldots, k \). Thus \( R(G, \hat{\delta}_n) - R(G, \hat{\delta}_G) = \sum_{i=1}^{k} [R_i(G, \delta_{ni}) - R_i(G, \delta_{Gi})] \geq 0 \) for all \( n \). This nonnegative regret Bayes risk \( D(G, \hat{\delta}_n) = R(G, \hat{\delta}_n) - R(G, \hat{\delta}_G) \) is used as a measure of performance of the empirical Bayes selection procedure \( \hat{\delta}_n \).

**Definition 3.1.** A sequence of empirical Bayes selection procedures \( \{ \hat{\delta}_n \}_{n=1}^{\infty} \) is said to be asymptotically optimal relative to the prior distribution \( G \) if \( R(G, \hat{\delta}_n) - R(G, \delta_G) = O(1) \). \( \{ \hat{\delta}_n \}_{n=1}^{\infty} \) is said to be asymptotically optimal of order \( \{ \alpha_n \}_{n=1}^{\infty} \) relative to the prior distribution \( G \) if \( R(G, \hat{\delta}_n) - R(G, \delta_G) = O(\alpha_n) \) where \( \{ \alpha_n \}_{n=1}^{\infty} \) is a sequence of positive numbers such that \( \lim_{n \to \infty} \alpha_n = 0 \).

3.2. The Proposed Empirical Bayes Selection Procedure

We construct an empirical Bayes selection procedure by mimicking the behavior of the Bayes selection procedure \( \hat{\delta}_G \) of (2.6).

Note that the functions \( \psi_{ia}(y), a = 1, 2 \), can be written as:

\[
\begin{cases}
\psi_{i1}(y) = \int_{t=y}^{\infty} \psi_{io}(t) dt, \\
\psi_{i2}(y) = \int_{t=y}^{\infty} t \psi_{io}(t) dt - y \psi_{i1}(y).
\end{cases}
\] (3.3)

For each \( i = 1, \ldots, k, \ell = 1, 2, \ldots, \) and \( y > 0 \), define
\[ V_{i	heta}(y) = I(Y_{i	heta} \geq y)/u(Y_{i	heta}). \]  

(3.4)

Then, \( E_{Y_{(n)}}[V_{i\theta}(y)] = \psi_{i1}(y), \) \( E_{Y_{(n)}}[(Y_{i\theta} - y)V_{i\theta}(y)] = \psi_{i2}(y). \) Thus, for \( W_{i\theta}(y) = (\theta_0 + y - Y_{i\theta})V_{i\theta}(y), \) \( E_{Y_{(n)}}[W_{i\theta}(y)] = \theta_0\psi_{i1}(y) - \psi_{i2}(y) = H_i(y). \) Now, for each \( i = 1, \ldots, k, \) and \( y > 0, \) define

\[ H_{in}(y) = \frac{1}{n} \sum_{i=1}^{n} W_{i\theta}(y) \]  

(3.5)

\( H_{in}(y) \) is an unbiased and consistent estimator of \( H_i(y). \) By mimicking the form \( (2.6) \), we propose an empirical Bayes selection procedure \( \delta_n^* = (\delta_{ni1}^*, \ldots, \delta_{nik}^*) \) as follows: For each \( i = 1, \ldots, k, \) and \( y_i > 0, \)

\[ \delta_{ni}^*(y_i) = \begin{cases} 1 & \text{if } H_{in}(y_i) \leq 0, \\ 0 & \text{otherwise}. \end{cases} \]  

(3.6)

The Bayes risk of the empirical Bayes selection procedure \( \delta_n^* \) is:

\[ R(G, \delta_n^*) = \sum_{i=1}^{k} R_i(G, \delta_{ni}^*), \]

\[ R_i(G, \delta_{ni}^*) = \int_{y=0}^{\infty} u(y) E_{Y_{(n)}}[\delta_{ni}^*(y)]H_i(y)dy + C_i \]  

(3.7)

\[ = \int_{y=0}^{\infty} u(y)E_{Y_{(n)}}[\delta_{ni}^*(y)]\psi_{i1}(y)[\theta_0 - \varphi_i(y)]dy + C_i. \]

4. Asymptotic Optimality and Rate of Convergence

4.1. Asymptotic Optimality of \( \delta_n^* \)

From \( (2.8) \) and \( (3.7) \), the regret Bayes risk of the selection procedure \( \delta_n^* \) is:

\[ R(G, \delta_n^*) - R(G, \delta_G) = \sum_{i=1}^{k} [R_i(G, \delta_{ni}^*) - R_i(G, \delta_{Gi})] \]  

(4.1)

and from \( (2.9) \), for each \( i = 1, \ldots, k. \)

\[ R_i(G, \delta_{ni}^*) - R_i(G, \delta_{Gi}) \]

\[ = \int_{y=0}^{\infty} u(y)\psi_{i1}(y)[\theta_0 - \varphi_i(y)]E_{Y_{(n)}}[\delta_{ni}^*(y) - \delta_{Gi}(y)]dy \]

\[ = \int_{y=0}^{a_i} u(y)\psi_{i1}(y)[\theta_0 - \varphi_i(y)]P\{\delta_{ni}^*(y) = 1, \delta_{Gi}(y) = 0\}dy + \int_{y=a_i}^{\infty} u(y)\psi_{i1}(y)[\varphi_i(y) - \theta_0]P\{\delta_{ni}^*(y) = 0, \delta_{Gi}(y) = 1\}dy. \]  

(4.2)
Note that under the assumption that \( \int \theta^2 dG_i(\theta) < \infty \), we have

\[
\int_0^\infty u(y)\psi_{i1}(y) \theta_0 - \varphi_i(y) \, dy \\
\leq \int_0^\infty \theta_0 u(y)\psi_{i1}(y) \, dy + \int_0^\infty u(y)\psi_{i1}(y) \varphi_i(y) \, dy \\
= \theta_0 \int \theta dG_i(\theta) + \int \theta^2 dG_i(\theta) < \infty.
\]

Therefore, to establish the asymptotic optimality of the selection procedure \( \hat{\delta}_n^* \), it suffices to show that for each \( i = 1, \ldots, k \), \( P\{\delta_{ni}^*(y) = 1, \delta_{Gi}(y) = 0\} \to 0 \) as \( n \to \infty \) for each \( 0 < y < a_i \), and \( P\{\delta_{ni}^*(y) = 0, \delta_{Gi}(y) = 1\} \to 0 \) as \( n \to \infty \) for each \( y > a_i \).

In either case, for \( \epsilon = 0, 1 \), by (2.6), (3.6) and the definition of \( H_{in}(y) \),

\[
P\{\delta_{ni}^*(y) = \epsilon, \delta_{Gi}(y) = 1 - \epsilon\} \\
\leq P\{|H_{in}(y) - H_i(y)| > |H_i(y)|\} \\
\leq \frac{E[Y_{in}(y)] H_{in}(y) - H_i(y))^2}{|H_i(y)|^2} = \frac{1}{n|H_i(y)|^2} \text{ Var}(W_{it}(y)),
\]

which tends to 0 as \( n \to \infty \) if \( \text{Var}(W_{it}(y)) \) is finite.

Note that \( W_{it}(y) = (\theta_0 + y - Y_{it})I(Y_{it} \geq y)/Y_{it}^{m-1} \). If \( m \geq 2 \), \( W_{it}(y) \) is a bounded random variable, and therefore, \( \text{Var}(W_{it}(y)) < \infty \). If \( m = 1 \), \( V_{it}(y) = I(Y_{it} \geq y) \) and \( \text{Var}(W_{it}(y)) \leq 2(\theta_0 + y)^2 \text{Var}(Y_{it}(y)) + 2\text{Var}(Y_{it}V_{it}(y)) \) since \( \text{Var}(V_{it}(y)) \leq 1 \) and \( \text{Var}(Y_{it}V_{it}(y)) \leq E[Y_{it}^2] = E[E[Y_{it}^2|\Theta_i] = E[2\Theta_i^2] < \infty \).

We summarize the result of the preceding discussion as a theorem as follows:

**Theorem 4.1** Let \( \hat{\delta}_n^* \) be the empirical Bayes selection procedure constructed in Section 3. Suppose that \( \int \theta^2 dG_i(\theta) < \infty \) for each \( i = 1, \ldots, k \). Then, \( \hat{\delta}_n^* \) is asymptotically optimal in the sense that \( R(G, \hat{\delta}_n^*) - R(G, \hat{\delta}_G) = o(1) \).

### 4.2. Rate of Convergence

In this subsection, we investigate the rate of convergence of the empirical Bayes selection procedure \( \hat{\delta}_n^* \) by establishing an upper bound on the regret Bayes risk \( R_i(G, \delta_{ni}^*) - R_i(G, \delta_{Gi}) \) for each \( i = 1, \ldots, k \). In the following, we consider the case where \( m \geq 2 \).
Therefore, $W_{it}(y)$ is a bounded random variable, with
\[
\begin{align*}
\text{Var}(W_{it}(y)) & \leq E[W_{it}(y)]^2 \\
& \leq \int_{y}^{\infty} \frac{(t - y - \theta_0)^2}{u(t)} \psi_{i0}(t) dt \\
& \leq \frac{1}{u(y)} \int_{y}^{\infty} (t - y - \theta_0)^2 \psi_{i0}(t) dt \\
& \leq \frac{1}{u(y)}[\theta_0^2 \psi_{i1}(y) + 2\psi_{i3}(y)].
\end{align*}
\]
(4.3)

From (4.2) and by the definitions of $\delta_{n_1}^*$ and $\delta_G$,
\[
R_i(G, \delta_{n_1}^*) - R_i(G, \delta_G) = \int_{0}^{a_i/2} u(y)\psi_{i1}(y)[\theta_0 - \varphi_i(y)]P\{H_{in}(y) - H_i(y) \leq -H_i(y)\} dy \\
+ \int_{a_i/2}^{a_i} u(y)\psi_{i1}(y)[\theta_0 - \varphi_i(y)]P\{H_{in}(y) - H_i(y) \leq -H_i(y)\} dy \\
+ \int_{a_i}^{a_i+1} u(y)\psi_{i1}(y)[\varphi_i(y) - \theta_0]P\{H_{in}(y) - H_i(y) \geq -H_i(y)\} dy \\
+ \int_{a_i+1}^{\infty} u(y)\psi_{i1}(y)[\varphi_i(y) - \theta_0]P\{H_{in}(y) - H_i(y) \geq -H_i(y)\} dy \\
\equiv D_{i1} + D_{i2} + D_{i3} + D_{i4}(\text{say}).
\]
(4.4)

Note that for each $y \in [\frac{a_i}{2}, a_i], 0 \leq H_i(y) = \theta_0\psi_{i1}(y) - \psi_{i2}(y) \leq \theta_0\psi_{i1}(\frac{a_i}{2})$ since $\psi_{i1}(y)$ is decreasing in $y$ for $y > 0$, and $W_{it}(y), t = 1, 2, \ldots, n$, are iid, bounded random variables with
\[
|W_{it}(y)| = \left| \frac{(\theta_0 + y - Y_{it})I(Y_{it} \geq y)}{Y_{it}^{m-1}} \right| \leq \frac{\theta_0 I(Y_{it} \geq y)}{Y_{it}^{m-1}} + \frac{(Y_{it} - y)I(Y_{it} > y)}{Y_{it}^{m-1}} \\
\leq \frac{\theta_0}{(\frac{a_i}{2})^{m-1}} + \frac{1}{(\frac{a_i}{2})^{m-2}}.
\]
So, there exists a number, say $Q_i(\theta_0) > 0$, such that $|W_{it}(y) - H_i(y)| \leq \frac{Q_i(\theta_0)}{2}$ for all $y \in [a_i/2, a_i]$. Hence, by Hoeffding's inequality and the definition of $H_{in}(y)$, for each $y \in [a_i/2, a_i]$,
\[
P\{H_{in}(y) - H_i(y) \leq -H_i(y)\} \leq \exp \left\{ \frac{-2nH_i^2(y)}{Q_i^2(\theta_0)} \right\} \\
= \exp \left\{ \frac{-2n\psi_{i1}^2(y)[\theta_0 - \varphi_i(y)]^2}{Q_i^2(\theta_0)} \right\} \\
\leq \exp\{-nb_i[\theta_0 - \varphi_i(y)]^2\},
\]
(4.5)
where \( b_i = 2\psi_{i1}^2 \left( \frac{a_i}{2} \right) / Q_i^2(\theta_0) \).

Let \( c_i = u(a_i)\psi_{i1}(\frac{a_i}{2}) \). Then \( 0 < u(y)\psi_{i1}(y) \leq u(a_i)\psi_{i1}(\frac{a_i}{2}) = c_i \) for all \( y \in [\frac{a_i}{2}, a_i] \). Next for each \( y \in [\frac{a_i}{2}, a_i] \), since \( G_i \) is non-degenerate.

\[
\frac{-d}{dy} [\theta_0 - \varphi_i(y)] = \frac{d}{dy} \frac{\psi_{i2}(y)}{\psi_{i1}(y)} = \frac{\int \theta^2 c(\theta)e^{-y/\theta} dG_i(\theta) \cdot \int c(\theta)e^{-y/\theta} dG_i(\theta) - [\int \theta c(\theta)e^{-y/\theta} dG_i(\theta)]^2}{\int \theta c(\theta)e^{-y/\theta} dG_i(\theta)} = \frac{\psi_{i2}(y)\psi_{i0}(y) - [\psi_{i1}(y)]^2}{[\psi_{i1}(y)]^2} \geq d_i(a_i) > 0.
\]

Therefore, combining the preceding inequalities and replacing them into \( D_{i2} \), we obtain

\[
D_{i2} \leq \int_{a_i/2}^{a_i} c_i[\theta_0 - \varphi_i(y)] \exp\{-nb_i[\theta_0 - \varphi_i(y)]^2\} dy
\]

\[
= \int_{a_i/2}^{a_i} \frac{c_i[-2nb_i(\theta_0 - \varphi_i(y))(\theta_0 - \varphi_i(y))']}{-2nb_i(\theta_0 - \varphi_i(y))^1} \exp\{-nb_i[\theta_0 - \varphi_i(y)]^2\} dy
\]

\[
\leq \int_{a_i/2}^{a_i} \frac{c_i[-2nb_i(\theta_0 - \varphi_i(y))(\theta_0 - \varphi_i(y))']}{2nb_i d_i(a_i)} \exp\{-nb_i[\theta_0 - \varphi_i(y)]^2\} dy.
\]

\[
= \frac{c_i}{2nb_i d_i(a_i)} \exp\{-nb_i[\theta_0 - \varphi_i(y)]^2\} \bigg|_{y = \frac{a_i}{2}}^{a_i}
\]

\[
\leq \frac{c_i}{2nb_i d_i(a_i)} = O(n^{-1}),
\]

where \( (\theta_0 - \varphi_i(y))' = \frac{d}{dy}[\theta_0 - \varphi_i(y)] \).

Following an analogous argument, we obtain

\[
D_{i3} = O(n^{-1}).
\]
Next, for $0 < y \leq a_i/2$, by Markov's inequality,

$$P\{H_{in}(y) - H_i(y) \leq -H_i(y)\} \leq \frac{E[|H_{in}(y) - H_i(y)|^2]}{[H_i(y)]^2}$$

$$\leq \frac{\text{Var}(W_{i1}(y))}{n[H_i(y)]^2}$$

$$\leq \frac{\theta_0^2 \psi_{i1}(y) + 2\psi_{i3}(y)}{n[H_i(y)]^2 u(y)}$$

where the last inequality follows from (4.3).

Also, for $y \in (0, \frac{a_i}{2}]$, $\theta_0 - \varphi_i(y) \geq \theta_0 - \varphi_i(\frac{a_i}{2}) > 0$. Therefore,

$$D_{i1} \leq \int_0^{\frac{a_i}{2}} \frac{1}{n[\theta_0 - \varphi_i(y)]\psi_{i1}(y)} \left[\theta_0^2 \psi_{i1}(y) + 2\psi_{i3}(y)\right] dy$$

$$\leq \int_0^{\frac{a_i}{2}} \frac{\theta_0^2}{n[\theta_0 - \varphi_i(\frac{a_i}{2})]} dy + \int_0^{\frac{a_i}{2}} \frac{2}{n[\theta_0 - \varphi_i(\frac{a_i}{2})]} \times \psi_{i3}(y) dy. \tag{4.8}$$

Here, we note that since $m \geq 2$,

$$0 < \psi_{i3}(y) = \int \theta^3 c(\theta)e^{-y/\theta} dG_i(\theta)$$

$$= \int_{\theta=0}^{1} \theta^3 c(\theta)e^{-y/\theta} dG_i(\theta) + \int_{1}^{\infty} \theta^3 c(\theta)e^{-y/\theta} dG_i(\theta)$$

$$\leq \int_{0}^{1} \theta c(\theta)e^{-y/\theta} dG_i(\theta) + \int_{1}^{\infty} \frac{\theta^2}{\Gamma(m)} e^{-y/\theta} dG_i(\theta)$$

$$< \infty,$$

and $\frac{\psi_{i3}(y)}{\psi_{i1}(y)}$ is increasing in $y$ for $y > 0$. Hence, $\int_0^c \frac{\psi_{i3}(y)}{\psi_{i1}(y)} dy < \infty$ for every $c > 0$. Combining these results into (4.8), we obtain:

$$D_{i1} = O(n^{-1}). \tag{4.9}$$

For $y > a_i + 1$, by Markov's inequality, for $0 < \lambda \leq 2$,

$$P\{H_{in}(y) - H_i(y) \geq -H_i(y)\} \leq \frac{E[|H_{in}(y) - H_i(y)|^\lambda]}{|H_i(y)|^\lambda}.$$
Therefore,

$$D_{i4} \leq \int_{\alpha_{i+1}}^{\infty} u(y)|H_i(y)|^{1-\lambda} E[|H_{in}(y) - H_i(y)|^\lambda] dy$$

$$\leq \int_{\alpha_{i+1}}^{\infty} u(y)|H_i(y)|^{1-\lambda}[E[|H_{in}(y) - H_i(y)|^2]^{\lambda/2}] dy$$

$$= \int_{\alpha_{i+1}}^{\infty} \frac{u(y)|H_i(y)|^{1-\lambda}}{n^{\lambda/2}} \text{Var}^{\lambda/2}(W_{i1}(y)) dy.$$  \hspace{1cm} (4.10)

So far, the rates of order $O(n^{-1})$ regarding the three terms $D_{i1}, D_{i2}$ and $D_{i3}$ are obtained only based on the conditions that $\int \theta^2 dG_i(\theta) < \infty$ and Assumption A holds. Therefore $D_{i4}$ is an essential part for the rate of convergence of $R_i(G, \delta_{n_i}^*) - R_i(G, \delta_{Gi})$. We summarize the preceding result as a theorem as follows.

**Theorem 4.2.** Suppose that Assumption A holds, $\int \theta^2 dG_i(\theta) < \infty$ for each $i = 1, 2, \ldots, k$ and $\int_{\alpha_{i+1}}^{\infty} u(y)|H_i(y)|^{1-\lambda}\text{Var}^{\lambda/2}(W_{i1}(y)) dy < \infty$, $i = 1, \ldots, k$. Then, $R_i(G, \delta_{n_i}^*) - R_i(G, \delta_{Gi}) = O(n^{-\lambda/2})$ for each $i = 1, \ldots, k$, and the empirical Bayes selection procedure $\delta_{n}^*$ is asymptotically optimal of order $O(n^{-\lambda/2})$ for some $0 < \lambda \leq 2$. That is, $R(G, \delta_{n}^*) - R(G, \delta_G) = O(n^{-\lambda/2})$.

To see how the asymptotic behavior of $D_{i4}$ is influenced by the tail probability of the probability density $f_i(y)$, we introduce the following lemma.

**Lemma 4.1** Let $X$ be a nonnegative random variable with probability density function $h(x)$. Then for $0 < t < 1$ and $p > 1$, $\int_1^\infty h^t(x) dx \leq (pt - 1)^{t-1}[E_h[X^{p(1-t)}]]^t$. 

13
Proof: Note that \( q(x) = (pt - 1)x^{-pt} \) is a probability density on \((1, \infty)\). Now,

\[
\int_1^\infty h^t(x)dx = \frac{1}{pt-1} \int_1^\infty (pt - 1)x^{-pt} [x^p h(x)]^t dx \\
= \frac{1}{pt-1} E_q[X^p h(X)]^t \text{ where } E_q \text{ is taken wrt } q(x) \\
\leq \frac{1}{pt-1} (E_q[X^p h(X)])^t \text{ by Hölder inequality} \\
= \frac{1}{pt-1} \left( \int_1^\infty (pt - 1)x^{-pt} x^p h(x) dx \right)^t \\
= (pt - 1)^{t-1} \left( \int_1^\infty x^{p(1-t)} h(x) dx \right)^t \\
\leq (pt - 1)^{t-1} (E_h[X^{p(1-t)}])^t.
\]

\[\square\]

For \( y \geq a_i + 1, \varphi_i(y) - \theta_0 \geq \varphi_i(a_i + 1) - \theta_0 \equiv e_i > 0 \). Also, note that \( u(y)\psi_{i1}(y) = m_i(G_i)f_i^*(y) \) where \( f_i^*(y) \) is a probability density defined in (2.10). Hence, by (4.3),

\[
\int_{a_i+1}^\infty u(y)|H_i(y)|^{\lambda-1} \text{Var}^{\lambda/2}(W_{i1}(y))dy.
\]

\[
\leq \int_{a_i+1}^\infty \frac{u(y)\psi_{i1}^{1-\lambda}(y)}{|\varphi_i(y) - \theta_0|^{\lambda-1}} \times \frac{1}{u^{\lambda/2}(y)}[\theta_0^2 \psi_{i1}(y) + 2\psi_{i3}(y)]^{\lambda/2} dy \\
\leq \int_{a_i+1}^\infty \frac{u^{1-\lambda/2}(y)\psi_{i1}^{1-\lambda}(y)}{e_i^{\lambda-1}}[(2\theta_0^2 \psi_{i1}(y))^{\lambda/2} + (4\psi_{i3}(y))^{\lambda/2}] dy \\
= \frac{2^{\lambda/2} \theta_0^\lambda}{e_i^{\lambda-1}} \int_{a_i+1}^\infty [u(y)\psi_{i1}(y)]^{1-\lambda/2} dy \\
+ \frac{2^\lambda}{e_i^{1-1}} \int_{a_i+1}^\infty [u(y)\psi_{i1}(y)]^{1-\lambda/2} (\frac{\psi_{i3}(y)}{\psi_{i1}(y)})^{\lambda/2} dy
\]
\[
\frac{2^{\lambda/2} \theta_0^2 (m_1(G_i))^{1-\lambda/2}}{\epsilon_i^{\lambda-1}} \int_{a_i+1}^{\infty} f_i^{*1-\lambda/2}(y)dy \\
+ \frac{2\lambda (m_1(G_i))^{1-\lambda/2}}{\epsilon_i^{\lambda-1}} \int_{a_i+1}^{\infty} f_i^{*1-\lambda/2}(y) \left[ \frac{\psi_{i3}(y)}{\psi_{i1}(y)} \right]^{\lambda/2} dy.
\] (4.11)

By Lemma 4.1,
\[
\int_{a_i+1}^{\infty} f_i^{*1-\lambda/2}(y)dy \leq \int_{1}^{\infty} f_i^{*1-\lambda/2}(y)dy \\
\leq (p(1 - \lambda/2) - 1)^{-\lambda/2} (E f_i^* [Y_i^{p \lambda/2}])^{1-\lambda/2},
\] (4.12)

for \( p > \frac{1}{1-\lambda/2} = \frac{2}{2-\lambda}, \) and where
\[
E f_i^* [Y_i^{p \lambda/2}] = \frac{E(G_i,f_i)[\Theta_i Y_i^{p \lambda/2}]}{m_1(G_i)} \\
= \frac{E_G[i][\Theta_i Y_i^{p \lambda/2} | \Theta_i]}{m_1(G_i)} \\
= \frac{\Gamma(m+p/2)}{\Gamma(m)m_1(G_i)} E_G[i][\Theta_i^{p \lambda/2+1}].
\] (4.13)

Therefore,
\[
\int_{a_i+1}^{\infty} f_i^{*1-\lambda/2}(y)dy.
\leq (p(1 - \lambda/2) - 1)^{-\lambda/2} \times \left[ \frac{\Gamma(m+p/2)}{\Gamma(m)m_1(G_i)} \right]^{1-\lambda/2} \times [E_G[i](\Theta_i^{p \lambda/2+1})]^{1-\lambda/2}.
\] (4.14)

Also, by Hölder inequality, for \( s > 1 \) such that \( \lambda s < 2 \) and \( \frac{1}{2} \times \frac{s}{s-1} \geq 1, \)
\[
\int_{a_i+1}^{\infty} f_i^{*1-\lambda/2}(y) \left[ \frac{\psi_{i3}(y)}{\psi_{i1}(y)} \right]^{\lambda/2} dy \\
= E f_i^* \left[ f_i^{* -\lambda/2}(Y)I(Y \geq a_i + 1) \left[ \frac{\psi_{i3}(Y)}{\psi_{i1}(Y)} \right]^{\lambda/2} \right] \\
\leq [E f_i^* [f_i^{*}(Y)I(Y \geq a_i + 1)]^{-\lambda s/2}]^{1/s} 	imes \left[ E f_i^* \left[ \frac{\psi_{i3}(Y)}{\psi_{i1}(Y)} \right]^{\lambda/2} \right]^{s-1}.
\] (4.15)
where
\[ E_{f_i}^*[f_i^*(Y)I(Y \geq a_i + 1)]^{-\lambda s/2}. \]
\[ = \int_{a_i+1}^{\infty} f_i^{*1-\lambda s/2}(y)dy \]
\[ \leq (p(1 - \lambda s/2) - 1)^{-\lambda s/2} \left[ \frac{\Gamma(m + p\lambda s/2)}{\Gamma(m)m_1(G_i)} \right]^{1-\lambda s/2} \times \left[ E_{G_i}^*([\Theta_i^{p\lambda s/2+1}]) \right]^{1-\lambda s/2}, \]
for \( p > \frac{1}{1 - \frac{\lambda s}{s-1}} = \frac{2}{2 - \lambda s}. \)

Since \( \frac{\psi_{13}(y)}{\psi_{11}(y)} = E_{G_i}^*[\Theta_i^2|Y_i = y] \) and \( \frac{1}{2} \times \frac{s}{s-1} \geq 1 \)
\[ E_{f_i}^* \left[ \frac{\psi_{13}(y)}{\psi_{11}(y)} \right]^{1-\lambda s/2} \leq E_{f_i}^* E_{G_i}^*[\Theta_i^{p\lambda s/2+1}|Y_i] \]
\[ = E_{f_i}^* E_{G_i}^*[\Theta_i^{\lambda s/2+1}] \]
\[ = E_{G_i}^*[\Theta_i^{p\lambda s/2+1}]/m_1(G_i). \]

For the two moments \( E_{G_i}^*[\Theta_i^{p\lambda s/2+1}] \) and \( E_{G_i}^*[\Theta_i^{\lambda s/2+1}] \) with parameters \( p, s \) and \( \lambda \) such that \( 1 < \lambda < 2, s > 1, \lambda s < 2, \frac{\lambda s}{2(s-1)} \geq 1 \) and \( p > \frac{1}{1 - \frac{s}{s-1}} = \frac{2}{2 - \lambda s} \), they will be equal if we let \( \frac{p\lambda s}{2} = \frac{\lambda s}{s-1} \), which implies that \( p = \frac{2}{s-1} \). Also, \( p = \frac{2}{s-1} > \frac{2}{2 - \lambda s} \) \( \Rightarrow 1 < s < \frac{3}{1+\lambda}. \)

We summarize the results of the preceding discussion as a Corollary of Theorem 4.2 as follows.

**Corollary 4.1.** Suppose that Assumption A holds and for each \( i = 1, \ldots, k \), \( \int \theta^{\frac{\lambda s}{2}+1} dG_i(\theta) < \infty \) for some \( 1 < \lambda < 2, s > 1 \), such that \( \lambda s < 2, \frac{\lambda s}{2(s-1)} \geq 1 \) and \( s < \frac{3}{1+\lambda}. \) Then,

(a) \( \int_{a_i+1}^{\infty} u(y)|H_i(y)|^{1-\lambda} \text{Var}^{\lambda/2}(W_{11}(y))dy < \infty \) for each \( i = 1, \ldots, k. \)

(b) The empirical Bayes selection procedure \( \delta_n^* \) is asymptotically optimal of order \( O(n^{-\lambda/2}). \)

5. A Lower Bound for \( R(G, \delta_n^*) - R(G, \delta_G) \) and the Best Possible Rate of Convergence

In this section, we will establish a lower bound with its rate of convergence for the regret Bayes risk \( R(G, \delta_n^*) - R(G, \delta_G) \). In fact, it suffices to consider a lower bound, say, for \( R_i(G, \delta_n^*) - R_i(G, \delta_G) \) since \( R(G, \delta_n^*) - R(G, \delta_G) \geq R_i(G, \delta_n^*) - R_i(G, \delta_G) \).
Theorem 5.1. Let Assumption A hold. Then,

\[ R_i(G, \delta_{ni}^*) - R_i(G, \delta_{Gi}) \geq O(n^{-1}). \]

Proof: From (4.4)

\[ R_i(G, \delta_{ni}^*) - R_i(G, \delta_{Gi}) \]
\[ \geq \int_{\frac{a_i}{2}}^{a_i} u(y)\psi_{i1}(y)[\theta_0 - \varphi_i(y)]P\{H_{in}(y) - H_i(y) \leq -H_i(y)\}dy. \]  \hspace{1cm} (5.1)

It follows from Lemma 3, on page 47, of Lamperti (1966) that for all \( \xi > 0 \), and for \( n \) being sufficiently large,

\[ P\{H_{in}(y) - H_i(y) \leq -H_i(y)\} \]
\[ = P\left\{ \frac{\sqrt{n}(H_{in}(y) - H_i(y))}{\sqrt{\text{Var} \left( W_{i1}(y) \right)}} \leq \frac{-\sqrt{n}H_i(y)}{\sqrt{\text{Var} \left( W_{i1}(y) \right)}} \right\} \]
\[ \geq \exp \left\{ -\frac{nH_i^2(y)(1 + \xi)}{2 \text{Var} \left( W_{i1}(y) \right)} \right\} \]
\[ = \exp \left\{ -\frac{n\psi_{i1}^2(y)(1 + \xi)(\theta_0 - \varphi_i(y))^2}{2 \text{Var} \left( W_{i1}(y) \right)} \right\} \]
\[ \geq \exp \left\{ -\frac{n\psi_{i1}^2(\frac{a_i}{2})(1 + \xi)(\theta_0 - \varphi_i(y))^2}{\gamma_i} \right\} \]
\[ = \exp \{-n\tau_{i1}(\theta_0 - \varphi_i(y))^2\}, \] \hspace{1cm} (5.2)

where \( \gamma_i = 2 \min_{\frac{a_i}{2} \leq y \leq a_i} \text{Var}(W_{i1}(y)) > 0 \) Since \( G_i \) is non-degenerate under Assumption A and \([\frac{a_i}{2}, a_i]\) is a compact interval, and \( \tau_{i1} = \psi_{i1}^2(\frac{a_i}{2})(1 + \xi)/\gamma_i > 0 \). Let \( \tau_{i2} = \min_{\frac{a_i}{2} \leq y \leq a_i} \frac{u(y)\psi_{i1}(y)}{\varphi_i(y)}. \)

Then \( \tau_{i2} > 0. \) Therefore, from (5.1) and (5.2),

\[ R_i(G, \delta_{ni}^*) - R_i(G, \delta_{Gi}) \]
\[ \geq \int_{\frac{a_i}{2}}^{a_i} u(y)\psi_{i1}(y)[\theta_0 - \varphi_i(y)]\exp \{-n\tau_{i1}(\theta_0 - \varphi_i(y))^2\}dy \]
\[ = \int_{\frac{a_i}{2}}^{a_i} \frac{u(y)\psi_{i1}(y)2n\tau_{i1}(\theta_0 - \varphi_i(y))\varphi_i'(y)}{2n\tau_{i1}\varphi_i(y)} \exp \{-n\tau_{i1}(\theta_0 - \varphi_i(y))^2\}dy \]
\[ \geq \int_{a_i/2}^{a_i} \frac{\tau_{i2}}{2\tau_{i1}} 2n\tau_{i1}(\theta_0 - \phi_i(y))\phi'_i(y) \exp\{-n\tau_{i1}(\theta_0 - \phi_i(y))^2\} dy \]

\[ = \frac{\tau_{i2}}{2n\tau_{i1}} \exp\{-n\tau_{i1}(\theta_0 - \phi_i(y))^2\} \bigg|_{y=a_i/2}^{a_i} \]

\[ = \frac{\tau_{i2}}{2n\tau_{i1}} [1 - \exp\{-n\tau_{i1}(\theta_0 - \phi_i(a_i/2))^2\}] \\
= O(n^{-1}). \] (5.3)

Therefore, the proof is complete. \qed

Theorem 5.1 provides a lower bound with a rate of convergence of order $O(n^{-1})$ for the regret Bayes risk $R(G, \hat{\delta}_{n}^*) - R(G, \hat{\delta}_G)$, while Corollary 4.1 gives an upper bound with a rate of convergence of order $O(n^{-\lambda/2})$ for the regret Bayes risk. When $\lambda$ is close to 2, a rate of convergence of order $O(n^{-1})$ will be the best possible rate of convergence for the empirical Bayes selection procedure $\hat{\delta}_{n}^*$. Suppose that the unknown prior distribution $G$ is such that $G_i(\theta^*) = 1$ for some $0 < \theta^* < \infty$, $i = 1, \ldots, k$, and Assumption A holds. Then $E_{G_i}[\Theta_i] < \infty$ for all $t > 0$. Thus, the quantity $\lambda$ in Corollary 4.1 can be chosen to be very close to 2. For example, if we let $\lambda = \lambda_n = 2 - 2 \ln \ln n / \ln n$ for $n \geq 16$, then the rate of convergence is of order $O(n^{-1} \ln \ln n)$, which is close to $O(n^{-1})$.

6. Concluding Remarks

We have presented a method to construct an empirical Bayes procedures $\hat{\delta}_{n}^*$ for selecting good exponential populations compared with a control. Through the analysis developed in Section 4, we can see that the part $D_{i4}$ plays an essential role in determining the rate of convergence of the selection procedure $\hat{\delta}_{n}^*$. We have demonstrated that the rate of convergence of $\hat{\delta}_{n}^*$ is influenced by the tail probability of the marginal probability densities $f_i(y), i = 1, \ldots, k$, or the moments of the prior distribution $G$. This empirical Bayes selection procedure $\hat{\delta}_{n}^*$ is asymptotically optimal and achieves a rate of convergence with order $O(n^{-\lambda/2}), 0 < \lambda \leq 2$, which may be close to the best possible rate of order $O(n^{-1})$ under some regularity conditions about the moments of $G$ according to Corollary 4.1. When the random parameters $\Theta_N$ are bounded and Assumption A holds, we have exhibited that $\hat{\delta}_{n}^*$ may achieve a rate of order $O(n^{-1} \ln \ln n)$, which is close to $O(n^{-1})$. However, it is not known whether $\hat{\delta}_{n}^*$ achieves the rate of order $O(n^{-1})$ or not. Singh (1979) and
Singh and Wei (1992) have commented that "a rate of the order $O(n^{-1})$ has not been achieved for any empirical Bayes procedures, whatever may be the component problem, in any Lebesgue-exponential, non-exponential regular or irregular family".
References


This paper deals with empirical Bayes selection procedures for selecting good exponential populations compared with a control. Based on the accumulated historical data, an empirical Bayes selection procedure $\delta_n^*$ is constructed by mimicking the behavior of a Bayes selection procedure. The empirical Bayes selection procedure $\delta_n^*$ is proved to be asymptotically optimal. The analysis shows that the rate of convergence of $\delta_n^*$ is influenced by the tail probabilities of the underlying distributions. It is shown that under certain regularity conditions on the moments of the prior distribution, the empirical Bayes selection procedure $\delta_n^*$ is asymptotically optimal of order $O(n^{-1/2})$ for some $0 < \lambda \leq 2$. A lower bound with rate of convergence of order $O(n^{-1})$ is also established for the regret Bayes risk of the empirical Bayes selection procedure $\delta_n^*$. This result suggests that a rate of order $O(n^{-1})$ might be the best possible rate of convergence for this empirical Bayes selection problem.
### GENERAL INSTRUCTIONS FOR COMPLETING SF 298

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

| Block 1. | Agency Use Only (Leave blank) |
| Block 2. | Report Date. Full publication date including day, month, and year, if available (e.g. 1 Jan 88). Must cite at least year. |
| Block 3. | Type of Report and Dates Covered. State whether report is interim, final, etc. If applicable, enter inclusive report dates (e.g. 10 Jun 87 - 30 Jun 88). |
| Block 4. | Title and Subtitle. A title is taken from the part of the report that provides the most meaningful and complete information. When a report is prepared in more than one volume, repeat the primary title, add volume number, and include subtitle for the specific volume. On classified documents enter the title classification in parentheses. |
| Block 5. | Funding Numbers. To include contract and grant numbers; may include program element number(s), project number(s), task number(s), and work unit number(s). Use the following labels: |
| | C - Contract  PR - Project |
| | G - Grant  TA - Task |
| | PE - Program  WU - Work Unit |
| | Element  Accession No. |
| Block 6. | Author(s), Name(s) of person(s) responsible for writing the report, performing the research, or credited with the content of the report. If editor or compiler, this should follow the name(s). |
| Block 7. | Performing Organization Name(s) and Address(es). Self-explanatory. |
| Block 8. | Performing Organization Report Number. Enter the unique alphanumeric report number(s) assigned by the organization performing the report. |
| Block 9. | Sponsoring/Monitoring Agency Name(s) and Address(es). Self-explanatory. |
| Block 10. | Sponsoring/Monitoring Agency Report Number. (If known) |
| Block 11. | Supplementary Notes. Enter information not included elsewhere such as; prepared in cooperation with...; Trans. of...; To be published in...; When a report is revised, include a statement whether the new report supersedes or supplements the older report. |

| Block 12a. | Distribution/Availability Statement. Denotes public availability or limitations. Cite any availability to the public. Enter additional limitations or special markings in all capitals (e.g. NORFON, REL, ITAR). |
| DOD | See DoDD 4230.25, "Distribution Statements on Technical Documents." |
| DOE | See authorities. |
| NASA | See Handbook NHB 2200.2. |
| NTIS | Leave blank. |

| Block 12b. | Distribution Code. |
| DOD | Leave blank |
| DOE | Enter DOE distribution categories from the Standard Distribution for Unclassified Scientific and Technical Reports |
| NASA | Leave blank. |
| NTIS | Leave blank. |

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

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

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

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


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