AN EMPIRICAL BAYES PROCEDURE FOR SELECTING GOOD POPULATIONS IN SOME POSITIVE EXPONENTIAL FAMILY

by

Shanti S. Gupta and Jianjun Li
Purdue University

Technical Report #99-25

Department of Statistics
Purdue University
West Lafayette, IN USA

November 1999

◇

AN EMPIRICAL BAYES PROCEDURE FOR SELECTING GOOD POPULATIONS IN SOME POSITIVE EXPONENTIAL FAMILY [1]

Shanti S. Gupta *and* Jianjun Li

Department of Statistics, Purdue University, W. Lafayette, IN 47907

Department of Statistics, Purdue University, W. Lafayette, IN 47907

Abstract:

This paper deals with the problem of selecting good ones compared with a control from $k(\geq 2)$ populations. The random variable associated with population π_i is assumed to be positive-valued and has density $f(x_i|\theta_i) = c(\theta_i)exp(-x_i/\theta_i)h(x_i)$ with unknown parameter θ_i, for each $i = 1, \cdots, k$. The distributions of parameters θ_i's are also unknown. A nonparametric empirical Bayes approach is used to construct the selection procedure. It is shown that this procedure is asymptotically optimal with a rate of order $O(n^{-1})$. The results are applicable to data arising from (most) life-test experiments.

AMS 1991 Subject classification: Primary 62F07; secondary 62C12.

Key words and phrases: Empirical Bayes; exponential family; rate of convergence; regret Bayes risk.

---

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

§ 1 Introduction and Formulation

In this paper, we are interested in the problem of simultanious inference and selection from among $k(\geq 2)$ populations in comparison with a standard or control. The populations are denoted by $\pi_1, \cdots, \pi_k$. The random variable X_i associated with π_i has the density $f(x_i|\theta_i) = c(\theta_i)e^{-x_i/\theta_i}h(x_i)$ with $h(x) > 0$ in $(0, \infty)$ or $(0, \tau_0]$ for some $\tau_0 > 0$, where the unknown parameter θ_i is the characterization of population π_i. For convenience, we write $(0, \tau]$ uniformly for $(0, \infty)$ and $(0, \tau_0]$.

Let θ_0 denote a standard or a control. In practical situations, we desire to differentiate between *good* and *bad* populations and select *good* ones and exclude *bad* ones. Here a population π_i is said to be *good* if $\theta_i \geq \theta_0$ and *bad* otherwise. This type of decision problem has been considered by many authors. For example, see early papers: Gupta and Sobel (1958) and Lehmann (1961), and later: Gupta and Hsiao (1983), and more recently: Gupta and Liang (1999).

Let $\Omega = \{\tilde{\theta} = \{\theta_1, \cdots, \theta_k\} : \theta_i > 0, i = 1, 2, \cdots, k\}$ be the parameter space. Let $A = \{\tilde{a} = \{a_1, \cdots, a_k\} : a_i = 0 \text{ or } 1, i = 1, \cdots, k\}$ be the action space, where $a_i = 1$ means that population π_i is selected as good, $a_i = 0$ means population π_i is excluded as bad.

The loss function we use is

$$L(\tilde{\theta}, \tilde{a}) = \sum_{i=1}^{k} l(\theta_i, a_i) \tag{1.1}$$

with

$$l(\theta_i, a_i) = a_i\theta_i(\theta_0 - \theta_i)I_{[\theta_i < \theta_0]} + (1 - a_i)\theta_i(\theta_i - \theta_0)I_{[\theta_i \geq \theta_0]}.$$

We also assume that θ_i is a realization of a random variable Θ_i, and $\Theta_1, \cdots,$

Θ_k are independently distributed with priors $G_1, \cdots, G_k$ respectively. Let $G = \prod_{i=1}^k G_i(\theta_i)$.

Let $\widetilde{X} = (X_1, \cdots, X_k)$ and $\mathcal{X}$ be the sample space of $\widetilde{X}$. Here X_i may be thought of as a sufficient statistic based on several i.i.d. samples.

The selection procedure $\tilde{\delta} = (\delta_1, \cdots, \delta_k)$, where $\delta_i(\tilde{x})$ is the probability of selecting population π_i as good when $\widetilde{X} = \tilde{x}$ is observed. To ensure that the Bayes rule exists, we assume $\int_0^\infty \theta^2 dG_i(\theta) < \infty$ for $i = 1, \cdots, k$.

Based on previous assumptions, a straightforward computation shows that

$$R(G, \tilde{\delta}) = \sum_{i=1}^k R_i(G, \delta_i) \tag{1.2}$$

and

$$R_i(G, \delta_i) = \int_{\mathcal{X}} \delta_i(\tilde{x})[\prod_{j \neq i} f_j(x_j)] w_i(x_i) h(x_i) d\tilde{x} + T_i \tag{1.3}$$

where

$$\begin{aligned} f_i(x_i) &= \int_0^\infty c(\theta_i) e^{-x_i/\theta_i} h(x_i) dG_i(\theta_i), \\ w_i(x_i) &= \int_0^\infty \theta_i(\theta_0 - \theta_i) c(\theta_i) e^{-x_i/\theta_i} dG_i(\theta_i), \\ T_i &= \int_0^\infty \theta_i(\theta_i - \theta_0) I_{[\theta_i > \theta_0]} dG_i(\theta_i). \end{aligned}$$

Here $f_i(x_i)$ is the marginal density of X_i and T_i is independent of the selection rule $\tilde{\delta}$. Clearly, a Bayes selection produce $\tilde{\delta}_G = (\delta_{Gi}, \cdots, \delta_{Gk})$ is given by

$$\delta_{Gi} = \begin{cases} 1 & \text{if } w_i(x_i) \leq 0, \\ 0 & \text{if } w_i(x_i) > 0. \end{cases} \tag{1.4}$$

Let $\alpha_i(x_i) = \int_0^\infty \theta_i c(\theta_i) e^{-x_i/\theta_i} dG_i(\theta_i)$ and $\psi_i(x_i) = \int_0^\infty \theta_i^2 c(\theta_i) e^{-x_i/\theta_i} dG_i(\theta_i)$. Denote $\phi_i(x_i) = \psi_i(x_i)/\alpha_i(x_i)$, the posterior mean of Θ_i with respect to prior $G_i^*(\theta)$, where $dG_i^*(\theta) = \theta dG_i(\theta) / \int \theta dG_i(\theta)$. Then δ_{Gi} can be expressed as

$$\delta_{Gi} = \begin{cases} 1 & \text{if } \phi_i(x_i) \geq \theta_0, \\ 0 & \text{if } \phi_i(x_i) < \theta_0. \end{cases} \tag{1.5}$$

It should be noted that δ_{Gi} depends on $\tilde{x}$ only through x_i. Also $\phi_i(x)$ is increasing for $i = 1, \cdots, k$. If x_i is large so that $\phi_i(x_i) \geq \theta_0$, we have $\delta_{Gi} = 1$; If x_i is small so that $\phi_i(x_i) < \theta_0$, we have $\delta_{Gi} = 0$. There are two trivial cases: $\delta_{Gi} = 1$ for all $x \in (0, \tau]$ or $\delta_{Gi} = 0$ for all $x \in (0, \tau]$. To exculde these trivial cases, we assume that $\tilde{\delta}_G$ is non-degenerate, i.e.

$$(1.6) \qquad \lim_{x \downarrow 0} \phi_i(x) < \theta_0 < \lim_{x \uparrow \tau} \phi_i(x), \qquad i = 1, \cdots, k.$$

If G_i is unknown, the Bayes rule cannot be applied and the selection cannot be made. The empirical Bayes approach is a way to help one to make the decision when past data are available. Since Robbins (1956, 1964) introduced the empirical Bayes approach, it has become a powerful tool in decision-making.

For each $i = 1, \ldots, k$, let $(X_{ij}, \Theta_{ij}), j = 1, 2, \ldots$ be random vectors associated with population π_i and stage j, where X_{ij} is observable while Θ_{ij} is unobservable. It is assumed that Θ_{ij} has a prior distribution G_i, for all $j = 1, 2, \ldots$, and conditioning on $\Theta_{ij} = \theta_{ij}, X_{ij}$ follows a distribution with density $f(x_{ij}|\theta_{ij})$ and $(X_{ij}, \Theta_{ij}), i = 1, \ldots, k, j = 1, 2, \ldots$ are mutually independent. At the present stage, say, stage $n + 1$, we have observed $\widetilde{X} = \tilde{x}$. The past acumulated observations are denoted by $(\widetilde{X}_1, \cdots, \widetilde{X}_n) = \widetilde{\widetilde{X}}_n$, where $\widetilde{X}_j = (X_{1j}, \cdots, X_{kj})$ is the observation at stage j. Based on $\widetilde{\widetilde{X}}_n$ and $\tilde{x}$, we wish to construct an empirical Bayes rule to select all good populations and to exculde all bad populations. Such an empirical Bayes rule can be expressed as

$$\widetilde{\delta_n}(\tilde{x}, \widetilde{\widetilde{X}}_n) = (\delta_{n1}(\tilde{x}, \widetilde{\widetilde{X}}_n), \cdots, \delta_{nk}(\tilde{x}, \widetilde{\widetilde{X}}_n))$$

where $\delta_{ni}(\tilde{x}, \widetilde{\widetilde{X}}_n)$ is the probability of selecting π_i as good if $\widetilde{\widetilde{X}}_n$ and $\tilde{x}$ are observed. Let $R(G, \tilde{\delta}_n)$ denote the overall Bayes risk of $\tilde{\delta}_n$. Then

$$R(\widetilde{G}, \tilde{\delta}_n) = \sum_{i=1}^{k} R_i(G, \delta_{ni}), \tag{1.7}$$

where

$$R_i(G, \delta_{ni}) = \int_{\mathcal{X}} E[\delta_{ni}(\tilde{x}, \widetilde{\widetilde{X}})] \cdot [\prod_{j \neq i} f_j(x_j)] \cdot w_i(x_i) h(x_i) d\tilde{x} + T_i. \tag{1.8}$$

The regret Bayes risk is defined as $R(G, \tilde{\delta}_n) - R(G, \tilde{\delta}_G)$, which is used to measure the performance of empirical Bayes rule $\tilde{\delta}_n$. If $R(G, \tilde{\delta}_n) - R(G, \tilde{\delta}_G) = o(1)$, we say that $\tilde{\delta}_n$ is asymptotically optimal (a.o.). If $R(G, \tilde{\delta}_n) - R(G, \tilde{\delta}_G) = O(\beta_n)$ for some positive β_n such that $\lim_{n\to\infty} \beta_n = 0$, we say that $\tilde{\delta}_n$ is asymptotically optimal at a rate of $O(\beta_n)$.

The aim of this paper is to construct an empirical Bayes rule for the selection problem described above. Then we show that the rule has a convergence rate of $O(n^{-1})$ under the above general setting or, in some cases, with the additional condition $\int_0^\infty \theta^3 dG(\theta) < \infty$ for most distributions in the family $f(x_i|\theta_i)$.

It should be pointed out that Gupta and Liang (1999) studied the selection problem for $gamma(x|\theta, s)$ populations, a special case of above problem, firstly through an empirical Bayes approach. They constructed an empirical Bayes rule δ_n^* and established its convergence rate $O(n^{-1})$ under some regularity conditions. A rate of $O(n^{-1} \log n)$ was obtained there under the condition that $\Theta_i's$ are bounded.

The paper is organized as follows. We gives the introduction and formulation of the problem in Section 1. In Section 2 an empirical Bayes selection rule $\tilde{\delta}_n$ is constructed. The asymptotic behavior of $\tilde{\delta}_n$ is investigated in Section 3. In Section 4, we provide a few typical examples as applications of our results. The proofs of our results are given in Section 5.

§ 2 Construction of Empirical Bayes Selection Procedure $\tilde{\delta}_n$

The construction of $\tilde{\delta}_n$ can be divided into three steps. First, we construct an estimator of $w_i(x)$. Second, we localize the Bayes rule. And then we complete the construction by mimicking the Bayes rule using the estimator of $w_i(x)$.

The construction of an estimator of $w_i(x)$ follows the idea of Gupta and Liang [1999]. For the loss function (1.1), an unbiased and consistent estimator of $w_i(x)$ can be obtained. For each $i = 1, \cdots, k$, $j = 1, \cdots, n$, and $x > 0$, define

$$V_{ij}(x) = \frac{\theta_0 + x - X_{ij}}{h(X_{ij})} I_{[X_{ij} \in [x, \tau]]}. \tag{2.1}$$

Through a standard calculation, we have $E[V_{ij}(x)] = w_i(x)$. Based on this nice property, an unbiased and consistent estimator of $w_i(x)$ can be constructed as:

$$W_{ni}(x) = \frac{1}{n} \sum_{j=1}^{n} V_{ij}(x), \tag{2.2}$$

for each $i = 1, \cdots, k$, and $x \in (0, \tau]$.

We call the next step as a localization of the Bayes test. Examining the Bayes selection rule $\tilde{\delta}_G$, one will be more likely to take action $a_i = 1$ if the observation of $X_i = x_i$ is relatively large and take action $a_i = 0$ if it is relatively small. By knowing this, we want to find two numbers B_n and L_n such that we select π_i as good if we observe $x_i > L_n$ and exclude it as bad if $x_i < B_n$. Here both B_n and L_n depend on n. This could be understood as follows. As n increases, we have more information from the accumulated data, and we should adapt new B_n and L_n so that our decision can be made more precisely.

Certainly, the exact form of $f(x|\theta)$ and the distribution G affect the choice of B_n and L_n. Since we have no knowledge about G except that $\int_0^\infty \theta_i dG(\theta_i) < \infty$ for $i = 1, \cdots, k$, we rely on $f(x|\theta)$ itself.

If $\lim_{x\downarrow 0} h(x) > 0$, let $B_n = 0$ and $L_n = \theta_0 \log n/3$. If $\lim_{x\downarrow 0} h(x) = 0$, let H_n and L_n be the two sequences of positive numbers such that $H_n e^{L_n/\theta_0} = n^{1/3}$ and $H_n \to \infty$, $L_n \to \infty$ as $n \to \infty$. For example, $H_n = n^{1/4}$ and $L_n = \theta_0 \log n/12$. Then define $B_n = \inf\{x < 1 : h(x) \leq 1/H_n\}$. It follows that $B_n \to 0$ since $H_n \to 0$ as $n \to \infty$.

According to what we mentioned at the beginning of this section, we propose the following empirical Bayes procedure: For each $i = 1, \cdots, k$, and x_i,

$$\delta_{ni}(x_i) = \begin{cases} 1 \text{ if } & (x_i > L_n^*) \text{ or } (B_n \leq x_i \leq L_n^* \text{ and } W_{ni}(x_i) \leq 0), \\ 0 \text{ if } & (x_i < B_n) \text{ or } (B_n \leq x_i \leq L_n^* \text{ and } W_{ni}(x_i) > 0), \end{cases} \tag{2.3}$$

where $L_n^* = L_n$ if $\tau = \infty$ and $L_n^* = L_n \wedge \tau_0$ if $\tau = \tau_0 < \infty$. This empirical Bayes procedure says that, at stage $n+1$, if the present observation x_i from π_i is relatively big or small, a decision will be made based on x_i only. If it is not too small or too big, we have to resort to past data information and use $W_{ni}(x)$, the estimator of $w_i(x)$, to make the decision.

§3 Asymptotic Optimality of $\tilde{\delta}_n(\tilde{x})$

In this section, the asymptotic behavior of $\tilde{\delta}_n$ is investigated. We derive the regret Bayes risk first. From (1.2) and (1.3), the Bayes risk of $\tilde{\delta}_G$ is

$R(G,\tilde{\delta}_G) = \Sigma_{i=1}^k R_i(G,\delta_{Gi})$ with

$$R_i(G,\delta_{Gi}) = \int_0^T \delta_{Gi}(\tilde{x})w_i(x_i)h(x_i)dx_i + T_i.$$

From (1.7) and (1.8), the Bayes risk of $\tilde{\delta}_n(\tilde{x})$ is $R(G,\tilde{\delta}_n) = \Sigma_{i=1}^k R_i(G,\delta_{ni})$ with

$$R_i(G,\delta_{ni}) = \int_0^T E[\tilde{\delta}_{ni}(\tilde{x})]w_i(x_i)h(x_i)dx_i + T_i.$$

Thus, the regret Bayes risk of $\tilde{\delta}_n$ is

$$R(G,\tilde{\delta}_n) - R(G,\tilde{\delta}_G) = \sum_{i=1}^k [R_i(G,\delta_{ni}) - R_i(G,\delta_{Gi})], \tag{3.1}$$

and $R_i(G,\delta_{ni}) - R_i(G,\delta_{Gi})$ can be written as

$$\begin{aligned} &R_i(G,\delta_{ni}) - R(G,\delta_{Gi}) \\ = &\int_{B_n}^{L_n^*} P(W_{ni}(x) \le 0)w_i(x)I_{[w_i(x)>0]}h(x)dx + \int_{B_n}^{L_n^*} P(W_{ni}(x) > 0)w_i(x)I_{[w_i(x)<0]}h(x)dx \end{aligned} \tag{3.2}$$

Under the assumption $\int_0^\infty \theta^2 dG_i(\theta) < \infty$, we have $\int_0^\infty |w_i(x)|h(x)dx < \infty$ from the inequality

$$\int_0^T |w_i(x)|h(x)dx \le \theta_0 \int_0^T \alpha_i(x)h(x)dx + \int_0^T \psi_i(x)h(x)dx \le \theta_0 \int_0^\infty \theta dG_i(\theta) + \int_0^\infty \theta^2 dG_i(\theta).$$

Since $W_n(x)$ is a consistent estimator of $w_i(x)$, $P(W_{ni}(x) \le 0) \to 0$ if $w_i(x) > 0$, and $P(W_{ni}(x) > 0) \to 0$ if $w_i(x) < 0$. Applying the dominated convergence theorem, we have $R(G_i,\delta_{ni}) - R(G_i,\delta_{Gi}) = o(1)$. Thus we have the following theorem.

Theorem 3.1 *Assume that* $\int_0^\infty \theta^2 dG_i(\theta) < \infty$ *for each* $i = 1, 2, \cdots, k$. *Then* $\tilde{\delta}_n$*, as defined by (2.3), is asymptotically optimal.*

Besides the asymptotic optimality, the convergence rate of an empirical Bayes procedure is also an important factor to be considered when the procedure is applied. The following discussion shows that the procedure $\tilde{\delta}_n$ achieves the rate $O(n^{-1})$.

From now on, we consider only those members of the family $f(x|\theta)$ in which $\lim_{x\uparrow\tau} h(x) > 0$ and $h(x)$ is bounded from below for any inner closed subset of $(0, \tau]$.

These members belong to one of the following cases:

Case 1. $\lim_{x\uparrow\tau} \frac{h(x)}{x} > 0$ and $\lim_{x\downarrow 0} h(x) > 0$.

Case 2. $\lim_{x\uparrow\tau} \frac{h(x)}{x} > 0$ and $\lim_{x\downarrow 0} h(x) = 0$.

Case 3. $\lim_{x\uparrow\tau} \frac{h(x)}{x} = 0$ and $\lim_{x\downarrow 0} h(x) > 0$.

Case 4. $\lim_{x\uparrow\tau} \frac{h(x)}{x} = 0$ and $\lim_{x\downarrow 0} h(x) = 0$.

The main result about the convergence rate of $\tilde{\delta}_n$ for the various cases is given in the following Theorem 3.2 and Corollary 3.3.

Theorem 3.2 *Assume that* $\int_0^\infty \theta^2 dG_i(\theta) < \infty$ *for* $i = 1, \cdots, k$, *and the Bayes rule* $\tilde{\delta}_G$ *is non-degenerate. In Case 3 and Case 4, we also assume that* $\int_1^\infty \theta^4 c(\theta) dG_i(\theta) < \infty$ *for* $i = 1, \cdots, k$. *Then*

$$R(G, \tilde{\delta}_n) - R(G, \tilde{\delta}_G) = O(n^{-1}). \tag{3.3}$$

Proof. The proof is given in Section 5.

In Case 3 and Case 4, the assumptions $\int_0^\infty \theta^2 dG_i(\theta) < \infty$ and $\int_1^\infty \theta^4 c(\theta) dG_i(\theta) < \infty$ can be simplified into $\int_0^\infty \theta^3 dG_i(\theta) < \infty$. So we have the following corollary.

Corollary 3.3 *In Case 3 and Case 4, if* $\int_0^\infty \theta^3 dG_i(\theta) < \infty$ *for* $i = 1, \cdots, k$, *and the Bayes rule* $\tilde{\delta}_G$ *is non-degenerate, then*

$$R(G, \tilde{\delta}_n) - R(G, \tilde{\delta}_G) = O(n^{-1}). \tag{3.4}$$

Proof. If $\tau = \tau_0 < \infty$, then $\lim_{n\to\tau} \frac{h(x)}{x} > 0$. It says that $\tau = \infty$ in Case 3 and Case 4. Note that $\theta c(\theta) = \theta[\int_0^\infty exp(-x/\theta)h(x)dx]^{-1}$ and for $\theta > 1$,

$$\theta^{-1}\int_0^\infty e^{-x/\theta}h(x)dx = \int_0^\infty e^{-y/\theta}h(y\theta)dy \geq e^{-2}\int_1^2 h(y\theta)dy > e^{-2}[\min_{t\geq 1} h(t)]. \tag{3.5}$$

It follows that $\theta c(\theta)$ is bounded for $\theta > 1$. Thus $\int_0^\infty \theta^3 dG_i(\theta) < \infty$ implies both $\int_0^\infty \theta^2 dG_i(\theta) < \infty$ and $\int_1^\infty \theta^4 c(\theta)dG_i(\theta) < \infty$. Then (3.4) follows (3.3).

From Theorem 3.2, one sees a rate of order $O(n^{-1})$ is obtained under a (quite) weak condition. If $\tilde{\delta}_G$ is non-degenerate, we only require $\int_0^\infty \theta^2 dG_i(\theta) < \infty$ in Case 1 and Case 2. The assumption $\int_0^\infty \theta^2 dG_i(\theta) < \infty$ guarantees the existence of the Bayes rule. This assumption is natural and not very stingent. In Case 3 and Case 4, we require one moment condition, $\int_0^\infty \theta^3 dG_i(\theta) < \infty$.

The applications of our results to a few typical distributions are presented in the following section. It includes the construction of $\tilde{\delta}_n$ and the statement of convergence rate for each distribution there.

§ 4 Examples and Results

We select a few distributions as examples. Our results certainly work for most lifetime distributions since most of them satisfy $\lim_{x\to\infty} h(x) > 0$.

Example 4.1 (*exp* (θ)*-family*). Consider the exponential populations having density

$$f(x_i|\theta_i) = \frac{1}{\theta_i}e^{-x_i/\theta_i}, \qquad x_i > 0,\ \ \theta_i > 0,\ \ i = 1, \cdots, k. \tag{4.1}$$

Here $h(x) \equiv 1$. This family belongs to Case 3. Take $B_n = 0$, $L_n = \theta_0 \log n/3$

and construct $\widetilde{\delta_n}$ as

$$\text{(4.2)} \qquad \delta_{ni}(x_i) = \begin{cases} 1 \ \text{ if } \ (x_i > L_n) \text{ or } (0 < x_i \le L_n \text{ and } W_{ni}(x_i) \le 0), \\ 0 \ \text{ if } \ (0 < x_i \le L_n \text{ and } W_{ni}(x_i) > 0). \end{cases}$$

Then applying Corollary 3.3, we have the following.

Result 4.1 *If X_i has density $f(x_i|\theta_i)$ given in (4.1), $\int_0^\infty \theta^3 dG_i(\theta) < \infty$ for all $i = 1, \cdots, k$, and the Bayes rule $\tilde{\delta}_G$ is non-degenerate, then $\tilde{\delta}_n$, as constructed in (4.2), has a rate of convergence of order $O(n^{-1})$.*

Example 4.2 (*Gamma (θ, s)-family with known $s > 1$*). Consider the gamma populations having density

$$\text{(4.3)} \qquad f(x_i|\theta_i) = \frac{x_i^{s-1}}{\Gamma(s)\theta_i^s} e^{-x_i/\theta_i}, \qquad x_i > 0, \ \ \theta_i > 0, \ \ i = 1, \cdots, k.$$

Here $h(x) = x^{s-1}$. This family belongs to Case 2. Let $H_n = n^{1/4}$ and $L_n = \theta_0 \log n/12$. Then $B_n = n^{-1/[4(s-1)]}$. Construct $\widetilde{\delta_n}$ as:

$$\text{(4.4)} \qquad \delta_{ni}(x_i) = \begin{cases} 1 \ \text{ if } \ (x_i > L_n) \text{ or } (B_n \le x_i \le L_n \text{ and } W_{ni}(x_i) \le 0), \\ 0 \ \text{ if } \ (x_i < B_n) \text{ or } (B_n \le x_i \le L_n \text{ and } W_{ni}(x_i) > 0). \end{cases}$$

Then applying Theorem 3.2, we have the following.

Result 4.2 *If X_i has density $f(x_i|\theta_i)$ given in (4.3), $\int_0^\infty \theta^2 dG_i(\theta) < \infty$ for all $i = 1, \cdots, k$, and the Bayes rule $\tilde{\delta}_G$ is non-degenerate, then $\tilde{\delta}_n$, as constructed in (4.4), has a rate of convergence of order $O(n^{-1})$.*

Example 4.3 (*Truncated Gamma (θ, s)-family with known $s > 1$*). Consider the gamma populations having density

$$\text{(4.5)} \qquad f(x_i|\theta_i) = c(\theta_i) x_i^{s-1} e^{-x_i/\theta_i}, \qquad x_i \in (0, \tau_0], \ \ \theta_i > 0, \ \ i = 1, \cdots, k.$$

Here $h(x) = x^{s-1}$. This family belongs to Case 2. Let $B_n = n^{1/4}$ and

$L_n = \theta_0 \log n/12$. Then $B_n = n^{-1/[4(s-1)]}$. Construct $\widetilde{\delta_n}$ as:

(4.6)

$$\delta_{ni}(x_i) = \begin{cases} 1 \text{ if } & (x_i > L_n) \text{ or } (B_n \le x_i \le L_n \wedge \tau_0 \text{ and } W_{ni}(x_i) \le 0), \\ 0 \text{ if } & (x_i < B_n) \text{ or } (B_n \le x_i \le L_n \wedge \tau_0 \text{ and } W_{ni}(x_i) > 0). \end{cases}$$

Then applying Theorem 3.2, we have the following.

Result 4.3 *If X_i has density $f(x_i|\theta_i)$ given in (4.5), $\int_0^\infty \theta^2 dG_i(\theta) < \infty$ for all $i = 1, \cdots, k$, and the Bayes rule $\tilde{\delta}_G$ is non-degenerate, then $\tilde{\delta}_n$, as constructed in (4.6), has a rate of convergence of order $O(n^{-1})$.*

Example 4.4 (*A population having the density with infinite many discontinuities*). Consider the exponential populations having density

(4.7)

$$f(x_i|\theta_i) = c(\theta_i) e^{-x_i/\theta_i} \sum_{l=0}^{\infty} (l+1) I_{[l<x_i\le l+1]}, \qquad x_i > 0, \ \ \theta_i > 0, \ \ i = 1, \cdots, k.$$

Here $h(x) = \Sigma_{l=0}^{\infty}(l+1)I_{[l<x\le l+1]}$. This family belongs to Case 1. Take $B_n = 0$, $L_n = \theta_0 \log n/3$ and construct $\tilde{\delta}_n$ as

$$(4.8) \qquad \delta_{ni}(x_i) = \begin{cases} 1 \text{ if } & (x_i > L_n) \text{ or } (0 < x_i \le L_n \text{ and } W_{ni}(x_i) \le 0), \\ 0 \text{ if } & (0 < x_i \le L_n \text{ and } W_{ni}(x_i) > 0). \end{cases}$$

Then applying Theorem 3.2, we have the following

Result 4.4 *If X_i has density $f(x_i|\theta_i)$ given in (4.7), $\int_0^\infty \theta^2 dG_i(\theta) < \infty$ for all $i = 1, \cdots, k$, and the Bayes rule $\tilde{\delta}_G$ is non-degenerate, then $\tilde{\delta}_n$ as constructed in (4.8), has a rate of convergence of order $O(n^{-1})$.*

Remark. Gupta and Liang (1999) considered the same selection problem for the gamma population (4.3). In that paper, an empirical Bayes rule was constructed as

$$\delta^*_{ni}(x_i) = \begin{cases} 1 & \text{if } W_{ni}(x_i) \le 0, \\ 0 & \text{if } W_{ni}(x_i) > 0. \end{cases}$$

The convergence rate of $\tilde{\delta}^*_n$ is affected by the tail probability of the underlying distributions. In our paper, we cut the interval $(0, \infty)$ into three parts $(0, B_n)$, $[B_n, L_n]$ and (L_n, ∞) by localizing the Bayes test. Then we construct the empirical Bayes rule as (4.4). So the influence of the tail probability of the underlying distributions is controlled and a rate of $O(n^{-1})$ is obtained under quite weak conditions as shown in Result 4.2.

§ 5 Proof of Theorem 3.2

The ideas of the proof are similar to those in Gupta and Li (1999). The main idea is to use a classic result about the non-uniform estimation of the difference between the normal distribution and the distribution of the sum of i.i.d. random variables.

Recall that $\tilde{\delta}_G$ is non-degenerate. That is,

$$\lim_{x \downarrow 0} \phi_i(x) < \theta_0 < \lim_{x \uparrow \tau} \phi_i(x), \qquad i = 1, \cdots, k, \tag{5.1}$$

Then G must be non-degenerate and $\phi_i(x)$ must be strictly increasing. Therefore there exists a point b_i such that $\phi_i(b_i) = \theta_0$, $\phi_i(x) > 0$ for $x > b_i$ and $\phi_i(x) < 0$ for $x < b_0$. Since we consider the asymptotic behavior of $\tilde{\delta}_n$, we assume $b_i \in (B_n, L^*_n)$ for $i = 1, \cdots, k$ without loss of generality.

We prove (3.3) only for $\tau = \infty$. The proof is similar if $\tau = \tau_0 < \infty$.

Lemma 5.1 *For each $i = 1, \cdots, k$, $w'_i(b_i) < 0$ and further there is a* *neighborhood of* b_i*, denoted by* $N(b_i, \epsilon_i)$*, such that* $N(b_i, \epsilon_i) \subset (B_n, L_n)$ *and*

$$A_i = \min_{x \in N(b_i, \epsilon_i)} |w_i'(x)| > 0. \tag{5.2}$$

Denote $b_{i1} = b_i - \epsilon_i$*,* $b_{i2} = b_i + \epsilon_i$*. Then for all* $x \in [B_n, b_{i1}] \cup [b_{i2}, L_n]$*,*

$$|w_i(x)| \geq M_i e^{-L_n/\theta_0}, \tag{5.3}$$

where $M_i = \epsilon_i A_i \int_{\theta_0}^{\infty} \theta c(\theta) dG(\theta) / \int_0^{\infty} \theta c(\theta) e^{-b_{i1}/\theta} dG_i(\theta) > 0$*.*

Proof. For $x > 0$, the derivative of $w_i(x)$ exists and can be expressed as

$$w_i'(x) = -\theta_0 \int_0^{\infty} e^{-x/\theta} c(\theta) dG_i(\theta) + \int_0^{\infty} \theta e^{-x/\theta} c(\theta) dG_i(\theta).$$

Under (5.1), G is non-degenerate. From Jensen's inequality, we see that for $x > 0$

$$\frac{\int_0^{\infty} \theta e^{-x/\theta} c(\theta) dG_i(\theta)}{\int_0^{\infty} e^{-x/\theta} c(\theta) dG_i(\theta)} < \frac{\int_0^{\infty} \theta^2 e^{-x/\theta} c(\theta) dG_i(\theta)}{\int_0^{\infty} \theta e^{-x/\theta} c(\theta) dG_i(\theta)}.$$

Plugging b_i for x in the above inequality, we have

$$\frac{\int_0^{\infty} \theta e^{-b_i/\theta} c(\theta) dG_i(\theta)}{\int_0^{\infty} e^{-b_i/\theta} c(\theta) dG_i(\theta)} < \theta_0.$$

This implies that $w_i'(b_i) < 0$.

Note that $w_i'(x)$ is continuous in $(0, \infty)$. We can find an ϵ_i-neighborhood of b_i, denoted by $N(b_i, \epsilon_i)$ such that $N(b_i, \epsilon_i) \subset (B_n, L_n)$ and

$$A_i = \min_{x \in N(b_i, \epsilon_i)} |w_i'(x)| > 0.$$

Then (5.2) is proved. On the other hand, rewrite $w_i(x)$ as

$$w_i(x) = \alpha_i(x)[\theta_0 - \phi_i(x)].$$

For $x \in [B_n, b_{i1}]$, noting $\phi_i(x)$ is strictly increasing in x, $\theta_0 - \phi_i(x) \geq \theta_0 - \phi_i(b_{i1})$. For $x \leq L_n$,

$$\alpha_i(x) \geq \int_{\theta_0}^{\infty} \theta c(\theta) e^{-x/\theta} dG_i(\theta) \geq e^{-L_n/\theta} \int_{\theta_0}^{\infty} \theta c(\theta) dG_i(\theta).$$

Thus

$$|w_i(x)| \geq [\theta_0 - \phi_i(x)]e^{-L_n/\theta} \int_{\theta_0}^{\infty} \theta c(\theta) dG_i(\theta).$$

Similarly, for $x \in [b_{i2}, L_n]$,

$$|w_i(x)| \geq [\phi_i(b_{i2}) - \theta_0]e^{-L_n/\theta} \int_{\theta_0}^{\infty} \theta c(\theta) dG_i(\theta),$$

Using (5.2) and the mean value theorem, we have $w_i(b_{i1}) \geq A_i\epsilon_i$. Then $\theta_0 - \phi_i(b_{i1}) \geq A_i\epsilon_i/\alpha_i(b_{i1})$. Similarly, $\phi_i(b_{i2}) - \theta_0 \geq A_i\epsilon_i/\alpha_i(b_{i2}) \geq A_i\epsilon_i/\alpha_i(b_{i1})$. Since $\alpha_i(b_{i1}) \geq \alpha_i(b_{i2})$. Then for $m_i = \epsilon_i A_i \int_{\theta_0}^{\infty} \theta dG_i(\theta)/\alpha_i(b_{i1})$, (5.3) holds.

This completes the proof of Lemma 5.1.

Next lemma deals with the bounds of the moments of $W_{ni}(x)$.

In Case 1 and Case 3, $\min_{0<x<\infty} h(x) > 0$. Let $S_n \equiv 1/\min_{0<x<\infty} h(x)$. In Case 2 and Case 4, Let $S_n = H_n \vee [1/\min_{1<x<\infty} h(x)]$. Then $h(x) \geq S_n$ for $x > B_n$ in all four cases. Recall $L_n = \theta_0 \log n/3$ in Case 1 ans Case 3 and $H_n e^{L_n/\theta_0} = n^{1/3}$ in Case 2 and Case 4. Then we have $S_n e^{L_n/\theta_0} \sim n^{1/3}$ as $n \to \infty$ in all four cases.

In Case 3 and Case 4, we know $\int_1^{\infty} \theta^4 c(\theta) dG_i(\theta) < \infty$ and let $C_i = \int_1^{\infty} \theta^4 c(\theta) dG_i(\theta)$.

Without loss of generality, we assume $h(x) \geq x$ for $x > 1$ in Case 1 and Case 2.

Lemma 5.2 *Let* $\sigma_i^2(x) = E[(V_{ij}(x) - w_i(x))^2]$ *and* $\gamma_i(x) = E[|V_{ij}(x) - w_i(x)|^3]$. *Then for* $x \in [B_n, L_n]$,

$$(5.4) \qquad \sigma_i^2(x) \leq \begin{cases} [2S_n(\theta_0+1)+1]^2 & \text{for Case 1 and Case 2}, \\ S_n[(\theta_0+1)^2\alpha_i(x) + 2(\theta_0+1)C_i] & \text{for Case 3 and Case 4}, \end{cases}$$

and

$$(5.5) \qquad \gamma_i(x) \le \begin{cases} 3[2S_n(\theta_0+1)+1]^3 + 3|w_i(x)|^3 & \textit{for Case 1 and Case 2}, \\ 9S_n^2[(\theta_0^2+6)\alpha_i(x)+6C_i] + 3|w_i(x)|^3 & \textit{for Case 3 and Case 4}. \end{cases}$$

For $x \in [b_{i1}, b_{i2}]$, there exist two constants $C_{i\sigma} > 0$ and $C_{i\gamma} > 0$ such that

$$(5.6) \qquad \sigma_i^2(x) \le C_{i\sigma}^2, \qquad \gamma_i \le C_{i\gamma}$$

For $x \in [B_n, b_{i1}] \cup [b_{i2}, L_n]$ and large n,

$$(5.7) \qquad n^{3/8}|w_i(x)|/|\sigma_i(x)| \ge 1.$$

Proof. Consider $x \in [B_n, L_n]$. Note that $h(x) \ge S_n^{-1}$. In Case 1 and Case 2, if $x \ge 1$, $h(x) \ge x$. Then

$$|V_{ij}(x)| \le I_{[X_j \ge x]}\theta_0/h(X_j) + I_{[X_j \ge x]}(\theta_0 + x - X_j)/h(X_j) \le \theta_0 S_n + 1.$$

If $x > 1$, it can be shown that $|V_{ij}(x)| \le 2S_n(\theta_0+1)+1$. Thus

$$\sigma_i^2(x) \le E[|V_{ij}(x)|^2] \le [2S_n(\theta_0+1)+1]^2.$$

For $\gamma_i(x)$, using $|a+b|^3 \le 3|a|^3 + 3|b|^3$, we have

$$\gamma_i(x) \le 3E[|V_{ij}(x)|^3] + 3|w_i(x)|^3 \le 3[2S_n(\theta_0+1)+1]^3 + 3|w_i(x)|^3.$$

In Case 3 and Case 4, a simple calculation shows that

$$\sigma_i^2(x) \le S_n[\theta_0^2\alpha_i(x) + 2\theta_0\psi_i(x) + 2\int_0^\infty \theta^3 c(\theta)e^{-x/\theta}dG_i(\theta)].$$

By breaking the interval $(0,\infty)$ into $(0,1)$ and $[1,\infty)$, we have $\int_0^\infty \theta^3 c(\theta)e^{-x/\theta}dG_i(\theta) \le C_i + \alpha_i(x)$ and $\psi_i(x) \le C_i + \alpha_i(x)$. Thus

$$\sigma_i^2(x) \le S_n[(\theta_0+1)^2\alpha_i(x) + 2(\theta_0+1)C_i].$$

Similarly,

$$\gamma_i(x) \le 9S_n^2[(\theta_0^3+6)\alpha_i(x)+6C_i]+3|w_i(x)|^3.$$

Now consider $x \in [b_{i1}, b_{i2}]$. It is easy to see that

$$\sigma_i^2(x) \le \begin{cases} \frac{[1+2(\theta_0+1)]^2}{[\min_{x>b_{i1}} h(x)]^2} \equiv C_{i\sigma}^2 & \text{in Case 1 and Case 2} \\ \frac{(\theta_0+1)^2\alpha_i(b_{i1})+2(\theta_0+1)C_i}{\min_{x\ge b_{i1}} h(x)} \equiv C_{i\sigma}^2 & \text{in Case 3 and Case 4} \end{cases}$$

and

$$\gamma_i(x) \le \begin{cases} 3[\frac{1+2(\theta_0+1)}{[\min_{x>b_{i1}} h(x)]}]^3 + 3\max_{b_{i1}\le x\le b_{i2}} |W_i(x)|^3 \equiv C_{i\gamma} & \text{in Case 1 and Case 2} \\ 9\frac{(\theta_0^3+6)\alpha_i(b_{i1})+6C_i}{[\min_{x\ge b_{i1}} h(x)]^2} + 3\max_{b_{i1}\le x\le b_{i2}} |W_i(x)|^3 \equiv C_{i\gamma} & \text{in Case 3 and Case 4} \end{cases}$$

Then (5.6) holds. Next we prove (5.7). From (5.3), $|w_i(x)| \ge M_i e^{-L_n/\theta_0}$ for $x \in [B_n, b_{i1}] \cup [b_{i2}, L_n]$. In Case 1 and Case 2,

$$|\frac{w_i(x)}{\sigma_{in}(x)}| \ge \frac{M_i e^{-L_n/\theta_0}}{2S_n(\theta_0+1)+1} \sim S_n^{-1}e^{-L_n/\theta_0} = O(n^{-3/9}).$$

In Case 3 and Case 4

$$|\frac{w_i(x)}{\sigma_{in}(x)}| \ge \frac{|\theta_0 - \phi_i(x)|}{S_n^{1/2}[(\theta_0+1)^2/\alpha_i(x)+2(\theta_0+1)C_i/[\alpha_i(x)]^2]^{1/2}}.$$

It is easy to see that $|\theta_0 - \phi_i(x)| \ge \min\{|\theta_0 - \phi_i(b_{i1})|, |\theta_0 - \phi_i(b_{i2})|\}$. We know from the proof of Lemma 5.1 that $\alpha_i(x) \ge e^{-L_n/\theta_0}\int_{\theta_0}^{\infty} \theta c(\theta)dG_i(\theta)$. Then

$$S_n^{1/2}[(\theta_0+1)^2/\alpha_i(x)+2(\theta_0+1)C_i/[\alpha_i(x)]^2]^{1/2} \sim S_n^{1/2}e^{L_n/\theta}.$$

Thus $|w_i(x)/\sigma_{in}(x)| = O(S_n^{-1/2}e^{-L_n/\theta_0}) = O(S_n^{1/2}n^{-1/3})$.

This completes the proof of Lemma 5.2.

Note that $V_{ij}(x)$ are i.i.d random variables for fixed x. For large n, the central limit theorem tells us that $\sum_{j=1}^n [V_{ij}(x) - w_i(x)]/[\sigma_i(x)\sqrt{n}]$ is close to $N(0,1)$ in distribution. Furthermore, we have the following non-uniform estimation of the difference between the normal distribution and the distribution of the sum of i.i.d random variables. This result can be found in Petrov (1975, pp125) or Michel (1981). Michel proved $A < 30.54$ in his paper.

Fact *Let $X_1, X_2, \cdots, X_n$ be i.i.d random variables, $EX_1 = 0$, $EX_1^2 = \sigma^2 > 0$, $E|X_1|^3 < \infty$. Then for all x*

$$|F_n(x) - \Phi(x)| \leq A \frac{\rho}{\sqrt{n}(1+|x|)^3}. \tag{5.8}$$

Here $\Phi(x)$ is the c.d.f. of $N(0,1)$, $F_n(x)$ and ρ are given by

$$F_n(x) = P\left(\frac{1}{\sigma\sqrt{n}} \sum_{j=1}^n X_j \leq x\right), \qquad \rho = \frac{E|X_1|^3}{\sigma^3}.$$

Now, we are ready to prove our main result.

Proof of Theorem 3.2 It suffices to prove $R_i(G, \delta_{ni}) - R_i(G, \delta_{Gi}) = O(n^{-1})$. Rewrite $P(W_{ni}(x) < 0)$ as

$$P\left(\frac{1}{\sqrt{n\sigma_i^2(x)}} \sum_{j=1}^n [V_{ij}(x) - w_i(x)] \leq -\frac{\sqrt{n}w_i(x)}{\sigma_i(x)}\right).$$

Then applying (5.8), we have

$$P(W_{ni}(x) < 0) \leq \Phi\left(-\frac{\sqrt{n}|w_i(x)|}{\sigma_i(x)}\right) + \frac{A\gamma_i(x)}{\sqrt{n}(\sigma_i(x) + \sqrt{n}|w_i(x)|)^3}.$$

Similarly,

$$P(W_{ni}(x) > 0) \leq 1 - \Phi\left(\frac{\sqrt{n}|w_i(x)|}{\sigma_i(x)}\right) + \frac{A\gamma_i(x)}{\sqrt{n}(\sigma_i(x) + \sqrt{n}|w_i(x)|)^3}.$$

Plugging above two inequalities in (3.2), we obtain

$$\begin{aligned}
& R_i(G, \delta_{ni}) - R_i(G, \delta_{Gi}) \\
= & \int_{B_n}^{b_i} \left[\Phi\left(-\frac{\sqrt{n}w_i(x)}{\sigma_i(x)}\right) + \frac{A\gamma_i(x)}{\sqrt{n}(\sigma_i(x) + \sqrt{n}|w_i(x)|)^3}\right] w_i(x) h(x) dx \\
& + \int_{b_0}^{L_n} \left[1 - \Phi\left(\frac{\sqrt{n}w_i(x)}{\sigma_i(x)}\right) + \frac{A\gamma_i(x)}{\sqrt{n}(\sigma_i(x) + \sqrt{n}|w_i(x)|)^3}\right] |w_i(x)| h(x) dx \\
\equiv & I + II.
\end{aligned}$$

From (5.5), (5.6), (5.7) and (5.8), we see that $w_i(x)$, $\sigma_i^2(x)$ and $\gamma_i(x)$ have different behavior for different x. So we decompose I into four parts.

$$\begin{aligned} I &\le \int_{B_n}^{b_{i1}} \Phi(-\frac{\sqrt{n}w_i(x)}{\sigma_i(x)})w_i(x)h(x)dx + \int_{b_{i1}}^{b_i} \Phi(-\frac{\sqrt{n}w_i(x)}{\sigma_i(x)})w_i(x)h(x)dx \\ &+ \int_{B_n}^{b_{i1}} \frac{A\gamma_i(x)w_i(x)h(x)}{\sqrt{n}(\sigma_i(x)+\sqrt{n}|w_i(x)|)^3}dx + \int_{b_{i1}}^{b_i} \frac{A\gamma_i(x)w_i(x)h(x)}{\sqrt{n}(\sigma_i(x)+\sqrt{n}|w_i(x)|)^3}dx \\ &\equiv I_1 + I_2 + I_3 + I_4. \end{aligned}$$

Consider I_1 first. According to (5.7), as n is large, $w_i(x)/\sigma_{in}(x) \ge n^{-3/8}$ for $x \in [B_n, b_{i1}]$, It follows that $\sqrt{n}w_i(x)/\sigma_i(x) \ge n^{1/8}$. Then applying it to I_1, we have

$$I_1 \le \Phi(n^{-1/8}) \int_{B_n}^{b_{i1}} w_i(x)h(x)dx = O(n^{-1}).$$

For I_2, since $x \in [b_{i1}, b_i]$, $h(x)$ is bounded and $\sigma_i(x) \le C_{i\sigma}$. Thus

$$I_2 \le [\max_{b_{i1}\le x\le b_i} h(x)] \int_{b_{i1}}^{b_i} \Phi(-\frac{\sqrt{n}w_i(x)}{C_{i\sigma}})w_i(x)dx.$$

To prove $I_2 = O(n^{-1})$, it is sufficient to prove $\int_{b_{i1}}^{b_i} \Phi(-\sqrt{n}w_i(x)/C_{i\sigma})w_i(x)dx = O(n^{-1})$. Using (5.2)

$$\begin{aligned} &\int_{b_{i1}}^{b_i} \Phi(-\sqrt{n}w_i(x)/C_{i\sigma})w_i(x)dx \\ \le\ &\frac{1}{A_i}\int_{b_{i1}}^{b_i} \Phi(-\sqrt{n}w_i(x)/C_{i\sigma})w_i(x)w_i'(x)dx \\ \le\ &\frac{C_{i\sigma}^2}{A_i n}\int_0^{\frac{\sqrt{n}w_i(b_{i1})}{C_{i\sigma}}} \Phi(-y)ydy \\ =\ &O(n^{-1}). \end{aligned}$$

Next we consider I_3. From (5.3), $|w_i(x)| \ge M_i e^{-L_n/\theta_0}$ for $x \in [B_n, b_{i1}]$. In Case 1 and Case 2, applying (5.5), we have

$$\begin{aligned} I_3 &\le \int_{B_n}^{b_{i1}} \frac{3[2S_n(\theta_0+1)+1]^3 + 3|w_i(x)|^3}{\sqrt{n}(\sigma_i(x)+\sqrt{n}|w_i(x)|^3)}w_i(x)h(x)dx \\ &\le \frac{3[2S_n(\theta_0+1)+1]^3}{n^2M_i^3e^{-3L_n/\theta_0}}\int_{B_n}^{b_{i1}} w_i(x)h(x)dx + \frac{3}{n^2}\int_{B_n}^{b_{i1}} w_i(x)h(x)dx \\ &= O(n^{-1}) \end{aligned}$$

In Case 3 and Case 4, using (5.5) again,

$$\begin{aligned} I_3 &\le \int_{B_n}^{b_{i1}} \frac{9S_n^2[(\theta_0^2+6)\alpha_i(x)+6C_i]+3|w_i(x)|^3}{\sqrt{n}(\sigma_i(x)+\sqrt{n}|w_i(x)|^3)} w_i(x)h(x)dx \\ &\le \frac{9S_n^2(\theta_0^2+6)}{n^2M_i^2e^{-2L_n/\theta_0}} \int_{B_n}^{b_{i1}} \alpha_i(x)h(x)dx + \frac{54S_n^2C_i}{n^2M_i^3e^{-3L_n/\theta_0}} \int_{B_n}^{b_{i1}} w_i(x)h(x)dx + \\ &\quad \frac{3}{n^2}\int_{B_n}^{b_{i1}} w_i(x)h(x)dx \\ &= O(n^{-1}) \end{aligned}$$

For $x \in [b_i, b_{i1}]$, $\gamma_i(x) \le C_{ir}$ and $w_i(x) \ge M_i e^{-L_n/\theta_0}$ from (5.6) and (5.3). Then

$$I_4 \le AC_{i\gamma}\int_{b_{i1}}^{b_i} \frac{w_i(x)h(x)dx}{\sqrt{n}(\sigma_i(x)+\sqrt{n}|w_i(x)|)^3} \le \frac{AC_{i\gamma}}{n^2M_i^3e^{-3L_n/\theta_0}}\int_{b_{i1}}^{b_i} w_i(x)h(x)dx = O(n^{-1}).$$

Thus we have $I = O(n^{-1})$. Similarly we can prove $II = O(n^{-1})$. Then the proof is complete.

References

[1] Gupta, S. S. and Hsiao, P. (1983). Empirical Bayes rules for selecting good populations. *J. Statist. Plann. Inference.* **8**, 87-101.

[2] Gupta, S. S. and Li, J. (1999). Empirical Bayes tests in some continuous exponential family. *Technical Report #99-21, Department of Statistics, Purdue University.*

[3] Gupta, S. S. and Liang, T. (1999). Selecting good exponential populations compared with a control : A nonparametric Bayes approach. *Sankhya, Ser. B.* **61**, 289-304.

[4] Gupta, S. S. and Sobel, M. (1958). On selecting a subset which contains all populations better than a standard. *Ann. Math. Statist.* **29**, 235-244.

[5] Lehmann, E. L. (1961). Some Model I problems of selection. *Ann. Math. Statist.* **32**, 2 990-1012.

[6] Michel, R. (1981). On the constant in the nonuniform version of the Berry-Esseen Theorem. *Z. W.* **55**, 109-117.

[7] Petrov, V. V. (1975). *Sums of Independent Random Variables.* Springer-verlag, Berlin Heidelberg New York.

[8] Robbins, H. (1956). An empirical Bayes approach to statistics. *Proc. Third. Berkeley Symp. Math. Statist, Probab.* **1**, 157-163, Univ. California Press, Berkeley.

[9] Robbins, H. (1964). The empirical Bayes approach to statistical decision problems. *Ann. Math. Statist.* **35**, 1-20.

[10] Van Houwelingen, J.C. (1976). Monotone empirical Bayes tests for the continuous one-parameter exponential family. *Ann. Statist.* **4**, 981-989.

SF 298 MASTER COPY KEEP THIS COPY FOR REPRODUCTION PURPOSES

REPORT DOCUMENTATION PAGE			*Form Approved* *OMB NO. 0704-0188*

1. AGENCY USE ONLY *(Leave blank)*

2. REPORT DATE
November, 1999

3. REPORT TYPE AND DATES COVERED
Technical Report, November 1999

4. TITLE AND SUBTITLE
An Empirical Bayes Procedure for Selecting Good Populations in Some Positive Exponential Family

5. FUNDING NUMBERS
DAAAH04-95-1-0165

6. AUTHOR(S)

Shanti S. Gupta and Jianjun Li

7. PERFORMING ORGANIZATION NAMES(S) AND ADDRESS(ES)

Purdue University
Department of Statistics
West Lafayette, IN 47907-1399

8. PERFORMING ORGANIZATION REPORT NUMBER
Technical Report #99-25

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)

U.S. Army Research Office
P.O. Box 12211
Research Triangle Park, NC 27709-2211

10. SPONSORING / MONITORING AGENCY REPORT NUMBER

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

12a. DISTRIBUTION / AVAILABILITY STATEMENT

Approved for public release; distribution unlimited.

12 b. DISTRIBUTION CODE

13. ABSTRACT *(Maximum 200 words)*

This paper deals with the problem of selecting good ones compared with a control from $k(\geq 2)$ populations. The random variable associated with population π_i is assumed to be positive-valued and has density $f(x_i|\theta_i) = c(\theta_i)\exp(-x_i/\theta_i)h(x_i)$ with unknown parameter θ_i, for each $i=1,\ldots,k$. The distributions of parameters θ_i's are also unknown. A nonparametric empirical Bayes approach is used to construct the selection procedure. It is shown that this procedure is asymptotically optimal with a rate of order $O(n^{-1})$. The results are applicable to data arising from (most) life-test experiments.

14. SUBJECT TERMS
Empirical Bayes; exponential family; rate of convergence; regret Bayes risk

15. NUMBER IF PAGES
21

16. PRICE CODE

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