MASTER COPY KEEP FOR REPRODUCTION PURPOSES

REPORT DOCUMENTATION PAGE

Form Approved
OMB NO. 0704-0188

1. AGENCY USE ONLY (Leave blank)

2. REPORT DATE
February 2001

3. REPORT TYPE AND DATES COVERED
Technical Report: February 2001

4. TITLE AND SUBTITLE
Monotone empirical Bayes tests based on kernel sequence estimation

5. FUNDING NUMBERS
DAAD19-00-1-0502

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
#01-03

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
ARO40940.2-MA

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)

Empirical Bayes inference problems involve the estimation of unknown functions (a density and its derivative). It is well known that this can be done through the kernel method, i.e. using a fixed index kernel and varied window bandwidth. In this paper, we introduce the kernel sequence method which considers using a sequence of kernel functions and allows the kernel index and window bandwidth to vary simultaneously in the estimates. This method usually produces better estimates since varied kernels give us more flexibility to do so.

We apply the above method to the construction of the monotone empirical Bayes test for the general continuous one-parameter exponential family. The rule we construct is shown to have a rate of convergence of $(\ln n)^{3+\epsilon}/n$ for any $\epsilon > 0$. This rate is a substantial improvement over the previous results. Note that this rate is much closer to $1/n$, which is proved here to be a lower bound for the monotone empirical Bayes tests. So the rule has good large sample behavior. Since the rule is monotone, it also has good performance for small samples.

14. SUBJECT TERMS
Empirical Bayes, regret Bayes risk, optimal rate of convergence, minimax

15. NUMBER IF PAGES
21

16. PRICE CODE

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

MONOTONE EMPIRICAL BAYES TESTS BASED ON KERNEL SEQUENCE ESTIMATION

by

Shanti S. Gupta and Jianjun Li
Purdue University Purdue University

Technical Report # 01-03

Department of Statistics
Purdue University
West Lafayette, IN USA

February 2001

Monotone Empirical Bayes Tests Based on Kernel Sequence Estimation [1]

Shanti S. Gupta and Jianjun Li

Abstract: Empirical Bayes inference problems involve the estimation of unknown functions (a density and its derivative). It is well known that this can be done through the kernel method, i.e. using a fixed index kernel and varied window bandwidth. In this paper, we introduce the kernel sequence method which considers using a sequence of kernel functions and allows the kernel index and window bandwidth to vary simultaneously in the estimates. This method usually produces better estimates since varied kernels give us more flexibility to do so.

We apply the above method to the construction of the monotone empirical Bayes test for the general continuous one-parameter exponential family. The rule we construct is shown to have a rate of convergence of $(\ln n)^{3+\epsilon}/n$ for any $\epsilon > 0$. This rate is a substantial improvement over the previous results. Note that this rate is much closer to $1/n$, which is proved here to be a lower bound for the monotone empirical Bayes tests. So the rule has good large sample behavior. Since the rule is monotone, it also has good performance for small samples.

[1]This research was supported in part by US Army Research Office, Grant DAAD19-00-1-0502 at Purdue University.

AMS Classification: 62C12.

Keywords: Empirical Bayes, regret Bayes risk, optimal rate of convergence, minimax.

1. Introduction. Assume that X is an observation from the distribution with density

$$f(x|\theta) = c(\theta)\exp\{\theta x\}h(x), \qquad -\infty \le a < x < b \le +\infty, \tag{1.1}$$

where $h(x)$ is continuous. positive for $x \in (a,b)$, θ is the parameter, which is distributed according to an unknown prior G on the parameter space Ω, a subset of the natural parameter space $\{\theta : c(\theta) > 0\}$.

We consider the problem of testing the hypotheses $H_0 : \theta \le \theta_0$ versus $H_1 : \theta > \theta_0$, where θ_0 is known. The loss function is $l(\theta, 0) = \max\{\theta - \theta_0, 0\}$ for accepting H_0 and $l(\theta, 1) = \max\{\theta_0 - \theta, 0\}$ for accepting H_1. A test $\delta(x)$ is defined to be a measurable mapping from (a, b) into $[0, 1]$ so that $\delta(x) = P\{$ accepting $H_1|X = x\}$, i.e., $\delta(x)$ is the probability of accepting H_1 when $X = x$ is observed. Let $R(G, \delta)$ denote the Bayes risk of a test δ when G is a prior distribution. Let $\phi_G(x) = E[\theta|X = x]$. Given that $E[|\theta|] < \infty$, a Bayes test δ_G is found as

$$\delta_G(x) = \begin{cases} 1 & \text{if } \phi_G(x) \ge \theta_0, \\ 0 & \text{if } \phi_G(x) < \theta_0. \end{cases} \tag{1.2}$$

Because $\phi_G(x)$ involves G, the above solution works only if the prior G is given. If G is unknown, this testing problem is formed as a compound decision problem and the empirical Bayes approach is used. Let $X_1, X_2, \cdots, X_n$ be the observations from n independent past experiences and let X be the present observation. Based on $\widetilde{X}_n = (X_1, X_2, \cdots, X_n)$ and X, an empirical Bayes rule $\delta_n(X, \widetilde{X}_n)$ can be constructed. The performance of δ_n is measured by $R(G, \delta_n) - R(G, \delta_G)$, where $R(G, \delta_n) = E[R(G, \delta_n|\widetilde{X}_n)]$. The quantity $R(G, \delta_n) - R(G, \delta_G)$ is referred as the regret Bayes risk (or regret) in the literature.

Denote $\alpha_G(x) = \int c(\theta)\exp(\theta x)dG(\theta)$, $\psi_G(x) = \int \theta c(\theta)\exp(\theta x)dG(\theta)$. It is clear that $\phi_G(x) = \psi_G(x)/\alpha_G(x)$ and $\phi_G(x) \ge \theta_0 \iff w(x) \equiv \theta_0\alpha_G(x) - \psi_G(x) \le 0$. So the construction of δ_n involves the estimation of $\alpha_G(x)$ and $\phi_G(x)$. This is usually done using the kernel method. In this paper, we introduce the kernel sequence method and apply it to obtain the estimates of $\alpha_G(x)$ and $\phi_G(x)$. The kernel sequence method considers using a sequence of kernel functions, and the kernel index and window bandwidth are allowed to vary simultaneously in the estimate(s). This method usually produces better estimates since varied kernels give us more flexibility to do so.

Based on the estimates of $\alpha_G(x)$ and $\phi_G(x)$, we construct an empirical Bayes rule δ_n for the testing problem mentioned above. Then we show that δ_n has a rate of convergence of $(\ln n)^{3+\epsilon}/n$ for any $\epsilon > 0$ with the assumption $E[|\theta|] < \infty$, which is a substantial improvement over the previous results. Note that this rate is much closer to $1/n$, which is proved here to be a lower bound for the monotone empirical Bayes tests. So the rule has good large sample behaviour. Since the rule is monotone, it also has good performance for small samples.

The readers interested in empirical Bayes approach may refer to two introductory papers of Robbins (1956, 1964). For the above empirical Bayes testing problem, Johns and Van Ryzin (1972) made an early contribution. Van Houwelingen (1976) used the monotonicity of the problem and constructed the monotone empirical Bayes tests, which achieve the rate of $O(n^{-2r/(2r+1)}(\ln n)^2)$ if $E[|\theta|^{r+1}] < \infty$. Van Houwelingen also showed that his rules have a good performance for small samples since they are monotone. Karunamuni and Yang (1995) studied monotone rules and their asymptotic behavior. With one more assumption $c_G \in [-A, A]$, they obtained the rate of $O(n^{-2r/(2r+1)})$. Karunamuni (1996) tried to find the optimal rate of convergence of the monotone empirical Bayes rule. But he failed; see Liang (2000a) and Liang (2000b), Gupta and Li(2000). Another related work is from Stijnen (1985). He studied the asymptotic behaviour of both the monotone empirical Bayes rules and non-monotone rules.

This paper is organized as follows: In Section 2 we introduce a few preliminary results. In Section 3 we introduce the idea of kernel sequence method. In Section 4, we construct the monotone empirical Bayes test δ_n and obtain its rate of convergence. Section 5 gives a lower bound of monotone empirical Bayes tests, which is n^{-1}. Section 6 contains the proofs of the main results in Section 4 and Section 5. In the appendix, we provide the proofs of a few lemmas used in Section 6.

2. Preliminary. We assume $\int |\theta| dG(\theta) < \infty$ throughout this paper. Note that $\alpha_G(x)$ and $\phi_G(x)$ exist for all $x \in (a,b)$ under the assumption $\int |\theta| dG(\theta) < \infty$. Therefore they are infinitely differentiable for $x \in (a,b)$. Furthermore, $\phi_G'(x) \geq 0$ and $\phi_G(x)$ is an increasing function. If $\lim_{x \downarrow a} \phi_G(x) \geq \theta_0$, then $\phi_G(x) \geq \theta_0$ and $\delta_G(x) \equiv 1$ for all $x \in (a,b)$; If $\lim_{x \uparrow b} \phi_G(x) \leq \theta_0$, then $\phi_G(x) \leq \theta_0$ and $\delta_G(x) \equiv 0$ for all $x \in (a,b)$. In both cases, we call that $\delta_G(x)$ is degenerate. We assume that $\delta_G(x)$ is non-degenerate in the following, i.e., we assume that $\lim_{x \downarrow a} \phi_G(x) < \theta_0 < \lim_{x \uparrow b} \phi_G(x)$. Then G is non-degernate and $\phi_G'(x) > 0$. Therefore there exists the unique point $c_G \in (a,b)$ such that $\phi_G(x) > \theta_0$ for $x > c_G$, $\phi_G(x) = \theta_0$ for $x = c_G$ and $\phi_G(x) < \theta_0$ for $x < c_G$ (see Van Houwelingen (1976) and others). Note that $w(x) = \theta_0 \alpha_G(x) - \psi_G(x)$. Then c_G is the unique root of $w(x)$.

Based on the previous discussion, the Bayes rule stated in Section 1 can be represented as

$$\delta_G(x) = \begin{cases} 1 & \text{if } \phi_G(x) \geq \theta_0 \iff w(x) \leq 0 \iff x \geq c_G, \\ 0 & \text{if } \phi_G(x) < \theta_0 \iff w(x) > 0 \iff x < c_G. \end{cases} \tag{2.1}$$

Noting that the Bayes rule δ_G is characterized by a single number c_G, a monotone empirical Bayes test (MEBT) can be constructed through estimating c_G by $c_n(X_1, X_2, \cdots, X_n)$, say, and defining

$$\delta_n = \begin{cases} 1 & \text{if} \quad x \geq c_n, \\ 0 & \text{if} \quad x < c_n. \end{cases} \tag{2.2}$$

Then the regret of δ_n is

$$R(G, \delta_n) - R(G, \delta_G) = E \int_{c_n}^{c_G} w(x) h(x) dx. \tag{2.3}$$

Remark 2.1. The assumption that $\delta_G(x)$ is non-degenerate is not crucial in this empirical Bayes testing problem. It can be reduced for the particular case of (1.1); see Gupta and Li (2000).

3. Kernel Sequence method. The kernel method has been used by many authors over the years. Here we introduce the kernel sequence method which uses a sequence of kernel functions instead of the single one. As the number of observations n increases, the kernel function and the kernel window bandwidth are set to vary simultaneously.

For each $i = 0, 1$ and $m = 1, 2, \cdots$, let $K_{im}(y)$ be a Borel-measurable function such that $K_{im}(y)$ vanishes outside the interval $[A_{im}, B_{im}]$, and for $K_{0m}(y)$

$$\int y^j K_{0m}(y)dy \begin{cases} = 1 & \text{if} \quad j = 0, \\ = 0 & \text{if} \quad j = 1, 2, \cdots, m-1, \cdots, k_{0m} - 1, \\ \neq 0 & \text{if} \quad j = k_{0m}, \end{cases} \tag{3.1}$$

and for $K_{1m}(y)$

$$\int y^j K_{1m}(y)dy \begin{cases} = 0 & \text{if} \quad j = 0, 2, 3, \cdots, m, \cdots, k_{1m} - 1, \\ = 1 & \text{if} \quad j = 1. \\ \neq 0 & \text{if} \quad j = k_{1m}. \end{cases} \tag{3.2}$$

Let $u = u_n$ be a sequence of positive numbers and $v = v_n$ be a sequence of positive integer numbers. For any $x \in (a, b)$, define

$$\alpha_n(x) = \frac{1}{nu}\sum_{j=1}^{n} K_{0v}\left(\frac{X_j - x}{u}\right)/h(X_j), \qquad \psi_n(x) = \frac{1}{nu^2}\sum_{j=1}^{n} K_{1v}\left(\frac{X_j - x}{u}\right)/h(X_j).$$

For u and v being properly chosen, $\alpha_n(x)$ and $\psi_n(x)$ are the estimates of $\alpha_G(x)$ and $\psi_G(x)$ respectively. In these kernel estimates, u is called the kernel (window) bandwidth and v is called the kernel index.

Note that the kernel indices of functions K_{0v} and K_{1v} will change as n increases. The method here is a little different from the traditional fixed index kernel method. Here both the kernel indices and window bandwidths vary in the construction.

4. MEBT For General Exponential Family. We use the idea of the kernel sequence method to find the estimators of $\alpha_G(x)$ and $\psi_G(x)$. Then we construct c_n based on these estimators.

We present the two sequences of kernel functions used in this paper. Define K_{0v} as follows: For odd v, $K_{0v}(y) = K_{0(v+1)}(y)$; for even v,

$$K_{0v}(y) = \begin{cases} p_v y^v + p_{v-1} y^{v-1} + \cdots + p_0, & \text{if } -1 \le y \le 1, \\ 0, & \text{otherwise}, \end{cases} \tag{4.1}$$

where

$$p_i = \begin{cases} 0, & \text{if } i \text{ is odd}, \\ \frac{(-1)^{i/2} v!(v+i)!v(v-i)}{i!(i+1)2^{2v+1}[(\frac{v}{2})!]^2(\frac{v+i}{2})!(\frac{v-i}{2})!}, & \text{if } i \text{ is even}. \end{cases}$$

Define $K_{1v}(y)$ as follows: For even v, $K_{1v}(y) = K_{1(v+1)}(y)$; for odd v,

$$K_{1v}(y) = \begin{cases} q_v y^v + q_{v-1} y^{v-1} + \cdots + q_0, & \text{if } -1 \le y \le 1, \\ 0, & \text{otherwise}. \end{cases} \tag{4.2}$$

where

$$q_i = \begin{cases} 0, & \text{if } i \text{ is even}. \\ \frac{(-1)^{(i+1)/2}(v+1)!(v+i)!(v-1)(v-i)}{i!(i+2)2^{2v+1}(\frac{v-1}{2})!(\frac{v+1}{2})!(\frac{v+i}{2})!(\frac{v-i}{2})!}, & \text{if } i \text{ is odd}. \end{cases}$$

Then $K_{0v}(y)$ defined by (4.1) satisfies (3.1) with $A_{0v} = -1$, $B_{0v} = 1$, $k_{0v} = v$ if v is even and $k_{0v} = v + 1$ if v is odd; $K_{1v}(y)$ defined by (4.2) satisfies (3.2) with $A_{1v} = -1$, $B_{1v} = 1$, $k_{1v} = v$ if v is odd and $k_{1v} = v + 1$ if v is even; see Gasser, Muller and Mammitzsch (1985).

Let ϵ_n be a sequence of positive numbers with $\epsilon_n \to 0$. Denote $u = u_n = \epsilon_n^{1/3}$. Let $v = v_n$ be a sequence of integer numbers such that $u^v \sim n^{-1}$. For any $x \in (a, b)$, define

$$\alpha_n(x) = \frac{1}{nu} \sum_{j=1}^{n} K_{0v}(\frac{X_j - x}{u})/h(X_j), \qquad \psi_n(x) = \frac{1}{nu^2} \sum_{j=1}^{n} K_{1v}(\frac{X_j - x}{u})/h(X_j). \tag{4.3}$$

It is shown later that $\alpha_n(x)$ and $\phi_n(x)$ are consistent estimators of $\alpha_G(x)$ and $\phi_G(x)$ respectively. Therefore $W_n(x) = \theta_0 \alpha_n(x) - \psi_n(x)$ is a consistent estimator of $w(x)$.

Since c_G is the unique root of $w(x)$, we are going to use $W_n(x)$ to construct c_n. Before doing this, let us examine δ_G. Note that δ_G is a monotone rule. If x is larger than c_G, we accept H_1; If x is smaller than c_G, we accept H_0. Since G is unknown, we do not know at which point we should accept H_0 or reject it. But, one will be more likely to accept H_1 if the present observation x is quite large and accept H_0 if it is quite small. By knowing this, we want to find two numbers c_{1n} and c_{2n} such that we accept H_1 if we observe $x > c_{2n}$ and accept H_0 if we observe $x < c_{1n}$. Here both cutoff points c_{1n} and c_{2n} 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 c_{1n} and c_{2n} so that our decision can be made more precisely. Once proper c_{1n} and c_{2n} are found, we can concentrate our effort on $x \in [c_{1n}, c_{2n}]$ in our construction.

The idea of splitting (a, b) into (a, c_{1n}), $[c_{1n}, c_{2n}]$ and (c_{2n}, b) is called the localization technique. To implement the localization technique, the following lemma is necessary.

Lemma 4.1. Four sequences of numbers $\{a_n, \bar{a}_n, b_n, \bar{b}_n\}$ can be found such that $a_n \downarrow a$, $b_n \uparrow b$, and as n is large

(i) $-[(\ln\ln n) \wedge u^{-1}] \le a_n < b_n \le [(\ln\ln n) \wedge u^{-1}]$;

(ii) $\min_{a_n < x < b_n} h(x) \ge u$;

(iii) $\int_{\bar{a}_n}^{a_n} h(t)dt \ge 2u, \qquad \int_{b_n}^{\bar{b}_n} h(t)dt \ge 2u.$

Let $c_{1n} = a_n + u + u^{1/3}$ and $c_{2n} = b_n - u - u^{1/3}$. From Lemma 4.1, we know that $c_{1n} \downarrow a$ and $c_{2n} \uparrow b$. So c_G will fall in $[c_{1n}, c_{2n}]$ for large values of n. Then we define c_n as in the following:

$$c_n = \int_{c_{1n}}^{c_{2n}} I_{[W_n(x)>0]}dx + c_{1n}. \tag{4.4}$$

A monotone empirical Bayes test $\delta_n(x)$ is now proposed as follows:

$$\delta_n = \begin{cases} 1 & \text{if} \quad x \geq c_n, \\ 0 & \text{if} \quad x < c_n. \end{cases} \tag{4.5}$$

It is obvious that $c_n \in [c_{1n}, c_{2n}]$. So if $x > c_{2n}$, we will accept H_1, and if $x < c_{1n}$, we will accept H_0. If $x \in [c_{1n}, c_{2n}]$, we will calculate c_n and compare x with c_n to make the decision.

The use of the localization technique helps us avoid the boundary effect of kernel estimates. It gives us nice bounds on the moments of $W_n(x)$ for $x \in [c_{1n}, c_{2n}]$(see Lemma 6.3 below). Also it results in a nice lower bound of $|w(x)|$ for $x \in [c_{1n}, c_G - \epsilon_G] \cup [c_G + \epsilon_G, c_{2n}]$ and $\epsilon_G > 0$ (see Lemma 6.2 below), which is crucial to get the desired rate of convergence in Section 6. For more uses of this technique, please see Gupta and Li (1999a), Gupta and Li (1999b), Gupta and Li (2000) and Li and Gupta (2000).

Note that since $W_n(x)$ is an estimate of $w(x)$, a natural construction of the empirical Bayes rule should be $\delta_n = 1$ if $W_n(x) \leq 0$ and $\delta_n = 0$ if $W_n(x) > 0$. Unfortunately this construction will lead to a non-monotone rule. So we use the integration of $I_{[W_n(x)>0]}$ in (4.4) instead. This technique is borrowed from Brown, Cohen, and Strawderman (1976), Van Houwelingen (1976) and Stijnen (1985).

Now we study the large sample behaviour of δ_n. The next two lemmas enable us to express the regret of δ_n through $c_n - c_G$.

Lemma 4.2. $w'(c_G) < 0$.

Since $w'(x)$ is continuous in (a, b), we can find $N_{\epsilon_G}(c_G)$, a neighborhood of c_G, such that $N_{\epsilon_G}(c_G) \subset (c_{1n}, c_{2n}) \subset (a, b)$ (as n is large), and $A_\epsilon = \min_{x \in N_{\epsilon_G}(c_G)}[-w'(x)] > 0$. Denote $\eta_1 = c_G - \epsilon_G$ and $\eta_2 = c_G + \epsilon_G$ in the following.

Lemma 4.3. *Let* $\bar{h} = \sup\{h(x) : x \in [\eta_1, \eta_2]\}$ *and* $\bar{w} = \sup\{-w'(x) : x \in [\eta_1, \eta_2]\}$. *Then*

$$R(G, \delta_n) - R(G. \delta_G) \leq 1/2\bar{h}\bar{w}E(c_n - c_G)^2 + (\theta_0 + E[|\theta|])\epsilon_G^{-4}E(c_n - c_G)^4.$$

Following (4.4) and $c_G \in [c_{1n}, c_{2n}]$, we have $c_n - c_G = -\int_{c_{1n}}^{c_G} I_{[W_n(x)\leq 0]}dx + \int_{c_G}^{c_{2n}} I_{[W_n(x)>0]}dx$. So a upper bound of $c_n - c_G$ is easy to obtain through the properties of $W_n(x)$ and $w(x)$. Note that $W_n(x)$ can be written as

$$W_n(x) = \frac{1}{n}\sum_{j=1}^{n} V_n(X_j, x), \quad \text{where } V_n(X_j, x) = \frac{\theta_0}{u} \cdot \frac{K_{0v}(\frac{X_j - x}{u})}{h(X_j)} - \frac{1}{u^2} \cdot \frac{K_{1v}(\frac{X_j - x}{u})}{h(X_j)}.$$

For fixed n and x, $V_n(X_j, x)$ are i.i.d. random variables. So $W_n(x)$ is the sum of the i.i.d. random variables. After applying the results in Petrov (1995), we have the following result.

Lemma 4.4. $\lim_{n\to\infty}[n\epsilon_n(\ln n)^3 E(c_n - c_G)^2] = 0, \quad \lim_{n\to\infty}[n\epsilon_n(\ln n)^3 E(c_n - c_G)^4] = 0.$

The proofs of Lemma 4.1-4.4 are given in Section 6. Lemma 4.3 and Lemma 4.4 lead us to the following theorem.

Theorem 4.1. *Assume that* $\int |\theta| dG(\theta) < \infty$ *and the Bayes rule* δ_G *is nondegenerate. Then for any* $\epsilon > 0$, $R(G, \delta_n) - R(G, \delta_G) = o((\ln n)^{3+\epsilon}/n)$.

Remark 4.1. In this paper, we get a faster rate of convergence for the general exponential family. This is mainly due to the use of the kernel sequence in the construction of estimate of $w(x)$. The previous papers in the literature constructed the empirical Bayes rules based on the kernel estimation with fixed kernel functions and varied window bandwidths. So the resulting rates are not fast. Now we let kernel functions and window bandwidths vary simultaneously. Then a better rate of convergence is obtained.

Remark 4.2. To apply the kernel sequence method, a key question is how to construct this sequence of kernel functions. In this paper we use the result obtained by Gasser, Muller and Mammitzsch (1985). We expect that the rate here will be improved if a "better" kernel sequence is found.

Remark 4.3. Note that the rule δ_n is monotone. It has the weak admissibility (see Van Houwelingen (1976)). So it also has good performance for small samples.

Remark 4.4. The result (4.6) is a rate of convergence for the general distribution (1.1). For some special member of the exponential family, the special property of that family member may be incorporated in the construction. Therefore, a better rate can possibly be obtained. See Liang (2000a) and Liang (2000b), Gupta and Li (2000).

5. Lower bound. We shall prove that $1/n$ is a lower bound for any MEBT even if θ is bounded.

As presented in Section 2, the problem of constructing a monotone empirical Bayes rule is essentially equivalent to finding an estimator c_n^* of c_G, a functional of the marginal distribution $f_G(x)$ of X, based on the i.i.d. sample $X_1, \cdots, X_n$. So a lower bound of MEBT's can be found through obtaining a lower bound of c_n^* going to c_G. This will be done using the ideas from Donoho and Liu (1991) or Fan (1991) and then constructing carefully the hardest two-point subproblem. In the following, l_1, l_2, $\cdots$ stand for the positive constants, which may have different values on different occasions.

Let $\mathcal{G}$ be the set of prior distributions with bounded supports inside $[\theta_0 - \theta_d, \theta_0 + \theta_d] \subset \Omega$ for some $\theta_d > 0$. Let $\mathcal{C}$ be the set of estimators c_n^* of c_G ($a < c_n^* < b$) and $\mathcal{D}$ be the set of empirical Bayes rules of type (2.2) with $c_n = c_n^* \in \mathcal{C}$. In order to find a minimax lower bound of MEBT's over $\mathcal{G}$, we first define $\mathcal{G}_0$, a subset of $\mathcal{G}$.

Denote $\theta_{01} = \theta_0 - \theta_d/2$ and $\theta_{02} = \theta_0 + \theta_d/2$. Choose any $c_0 \in (a, b)$. Let

$$g_0(\theta) = m_0 \exp(-\theta c_0)/c(\theta) I_{[\theta_{01} \le \theta \le \theta_{02}]}, \quad g_1(\theta) = m_1 \exp(-\theta x_d) g_0(\theta),$$

where (i) m_i is normalizing constant satisfying $\int g_i(\theta)d\theta = 1$ for $i = 1, 2$, (ii) x_d satisfies that $w_0'(x) < 1/2 w_0'(c_0) < 0$ for all $x \in [c_0 - x_d, c_0 + x_d] \subset (a, b)$, $w_0(x) = w(x)$ associated with $G \sim g_0$ $(dG(\theta) = g_0(\theta)d\theta)$. Let $\mathcal{F} = \{f_G(x) = \int f(x|\theta)dG(\theta) : G \in \mathcal{G}_0\}$, where

$$\mathcal{G}_0 = \{G : G \sim g = (1 + \sqrt{m})^{-1}[\sqrt{m} g_1(\theta) + g_0(\theta)], m = 0, 1, \cdots, \infty\}.$$

The next lemma tells us that finding a lower bound of MEBT's is equivalent to finding a lower bound of the hardest two-point subproblem.

Lemma 5.1. *Let c_i be the critical point corresponding to f_i, $i = 1, 2$. Then*

$$\begin{aligned} & \inf_{\delta_n^* \in \mathcal{D}} \sup_{G \in \mathcal{G}} [R(G, \delta_n^*) - R(G, \delta_G)] \\ \ge\ & \inf_{\delta_n^* \in \mathcal{D}} \sup_{G \in \mathcal{G}_0} [R(G, \delta_n^*) - R(G, \delta_G)] \\ \ge\ & l_1 \sup\{(c_1 - c_2)^2 : \int [\sqrt{f_1(x)} - \sqrt{f_2(x)}]^2 dx \le l_2/n, \quad f_1, f_2 \in \mathcal{F}\}. \end{aligned}$$

The lemma 5.1 is proved based on a result of Donoho and Liu (1991). From this lemma, we need to identify f_1 and f_2 in $\mathcal{F}$ to find the minimax lower bound.

Lemma 5.2. Let $g_2(\theta) = (1 + \sqrt{n})^{-1}[\sqrt{n} g_1(\theta) + g_0(\theta)]$. Let $f_i(x) = \int f(x|\theta) g_i(\theta) d\theta$ for $i = 1, 2$. Then $f_i \in \mathcal{F}$ and

$$\int [\sqrt{f_1(x)} - \sqrt{f_2(x)}]^2 dx \le \frac{l_2}{n}, \qquad (c_2 - c_1)^2 \ge \frac{l_3}{n}.$$

As a natural conclusion of Lemma 5.1 and Lemma 5.2, we have the following theorem.

Theorem 5.1. *For some $l > 0$, $\inf_{\delta_n^* \in \mathcal{D}} \sup_{G \in \mathcal{G}} [R(G, \delta_n^*) - R(G, \delta_G)] \ge l/n$.*

Remark 5.1. A natural question for empirical Bayes inference problems is: what is a lower (or the best lower) bound of monotone empirical Bayes rules for general exponential family. For empirical estimation problem, Singh (1979) conjectured that n^{-1} is a lower bound and also it is not obtainable even if θ is bounded. For the testing problem, we know now that n^{-1} is a lower bound for the monotone empirical Bayes rules.

Remark 5.2. Since the optimal rate of monotone rules for $N(\theta, 1)$ is $(\ln n)^{1.5}/n$ (see Gupta and Li (2000)), n^{-1} may not be the best lower bound or obtainable lower bound for general exponential family (1.1). Also we believe that it is not possible to find the obtainable lower bound for family (1.1) once. It must be done for each distribution individually and the information stored in that particular distribution must be incoporated.

6. Proofs. We shall prove the results in the previous sections. First we state some lemmas which will be used in this section. Their proofs are provided in the appendix.

6.1. Some Lemmas. As n is large, we have the following lemmas.

Lemma 6.1. *Let* $\bar{\alpha}_n = \max\{\alpha_G(x) : x \in [a_n, b_n]\}$. *Then* $\bar{\alpha}_n \leq (2u)^{-1}$.

Lemma 6.2. *For* $x \in [c_{1n}, c_{2n}]$, $|w(x)| \leq 2/u^2$;

For $x \in [c_{1n}, \eta_1] \cup [\eta_2, c_{2n}]$, $|w(x)| \geq M \cdot u(\ln n)^{-B}$, *where* $M > 0$, $B > 0$.

Lemma 6.3. *Let* $w_n(x) = E[V_n(X_j, x)]$, $Z_{jn} = V_n(X_j, x) - w_n(x)$, $\sigma_n^2(x) = E[|Z_{jn}|^2]$ *and* $\gamma_n(x) = E[|Z_{jn}|^3]$. *Then*

(i) *For* $x \in [c_{1n}, c_{2n}]$, $|w_n(x) - w(x)| \leq 1/\sqrt{n}$.

(ii) *For* $x \in [c_{1n}, c_{2n}]$, $\sigma_n(x) \leq l_1 v^{3/2} u^{-5/2}$; *for* $x \in [\eta_1, \eta_2]$, $l_2 \leq \sigma_n(x) \leq l_3 (v/u)^{3/2}$.

(iii) *For* $x \in [c_{1n}, c_{2n}]$, $\gamma_n(x) \leq l_4 v^{13} 36^v u^{-6}$.

Lemma 6.4. *Let $d_n = \sqrt{v^3/nu^3}$. For $x \in [c_{1n}, c_{2n}]$,*

$$w(x) > d_n \Longrightarrow w_n(x) \geq w(x)/2, \qquad w(x) < -d_n \Longrightarrow w_n(x) \leq w(x)/2.$$

6.2. Proof of Lemma 4.1. Lemma 4.1 is obvious intuitively. We also give a rigorous proof here. Let $h(a+) = \lim_{x\downarrow a} h(x)$ and $h(b-) = \lim_{x\uparrow b} h(x)$. Choose any $\xi \in (a,b)$. Let

$$h_a = \begin{cases} \max\{a < x < \xi : h(x) \leq u\} & \text{if } h(a+) = 0, \\ a & \text{if } 0 < h(a+) \leq \infty, \end{cases}$$

$$h_b = \begin{cases} \min\{\xi < x < b : h(x) \leq u\} & \text{if } h(b-) = 0, \\ b & \text{if } 0 < h(b-) \leq \infty, \end{cases}$$

And

$$S_a = \begin{cases} \max\{a < x < \xi : \int_a^\xi h(t)dt \leq 2u\} & \text{if } \int_a^\xi h(t)dt < \infty, \\ a & \text{if } \int_a^\xi h(t)dt = \infty, \end{cases}$$

$$S_b = \begin{cases} \min\{\xi < x < b : \int_\xi^b h(t)dt \leq 2u\} & \text{if } \int_\xi^b h(t)dt < \infty, \\ b & \text{if } \int_\xi^b h(t)dt = \infty. \end{cases}$$

Then we define a_n and b_n as follows:

$$\begin{cases} a_n = h_a \vee S_a \vee (a + 1/n) \vee (-\ln\ln n) \vee (-1/u), \\ b_n = h_b \wedge S_b \wedge (b - 1/n) \wedge (\ln\ln n) \wedge (1/u). \end{cases}$$

And let

$$\bar{a}_n = \begin{cases} a & \text{if } \int_a^\xi h(t)dt < \infty, \\ x_a \in \{a < x < \xi : \int_x^{a_n} h(t)dt \geq 2u\} & \text{if } \int_a^\xi h(t)dt = \infty, \end{cases}$$

$$\bar{b}_n = \begin{cases} b & \text{if } \int_\xi^b h(t)dt < \infty, \\ x_b \in \{\xi < x < b : \int_{b_n}^x h(t)dt \geq 2u\} & \text{if } \int_\xi^b h(t)dt = \infty. \end{cases}$$

Then it is easy to see that $a_n \downarrow a$, $b_n \uparrow b$, (i), (ii) and (iii) in Lemma 4.1 hold.

6.3. Proof of Lemma 4.2. Note that $\alpha_G(x)$ is infinitely differentiable, $\alpha'_G(x) = \psi_G(x)$ and $w'(x) = \theta_0\psi_G(x) - \psi'_G(x)$. If $\psi_G(c_G) = 0$, then $w'(c_G) = -\int \theta^2 c(\theta)e^{\theta c_G}dG(\theta) < 0$. If $\psi_G(c_G) > 0$, by Jesen Inequality $\psi'_G(c_G)/\psi_G(c_G) > \psi_G(c_G)/\alpha_G(c_G) = \theta_0$. Thus $w'(c_G) < 0$. Similarly, if $\psi_G(c_G) < 0$, $w'(c_G) < 0$. The proof of Lemma 4.2 is complete.

6.4. Proof of Lemma 4.3. From (2.3),

$$\begin{aligned} R(G,\delta_n) - R(G,\delta_G) &\le E[I_{[|c_n - c_G| > \epsilon_G]}\int_{c_n}^{c_G} w(x)h(x)dx] + \bar{h}E[I_{[|c_n - c_G| \le \epsilon_G]}\int_{c_n}^{c_G} w(x)dx] \\ &\le (\theta_0 + \mu_G)\epsilon_G^{-4}E(c_n - c_G)^4 + 1/2\bar{h}\bar{w}E(c_n - c_G)^2, \end{aligned}$$

where $\int_{c_n}^{c_G} w(x)h(x)dx \le (\theta_0 + \mu_G)$ and by Taylor expansion

$$I_{[|c_n - c_G| \le \epsilon_G]}\int_{c_n}^{c_G} w(x)dx = -1/2 \times w'(\hat{c}_n)(c_n - c_G)^2 I_{[|c_n - c_G| \le \epsilon_G]} \le 1/2\bar{w}(c_n - c_G)^2.$$

6.5. Proof of Lemma 4.4. From (4.4),

$$E(c_n - c_G)^2 \le E[\int_{c_{1n}}^{c_G} I_{[W_n(x)\le 0]}dx]^2 + E[\int_{c_G}^{c_{2n}} I_{[W_n(x)>0]}dx]^2 \equiv r_{1n} + r_{2n}. \tag{6.1}$$

It turns out by Holder inequality and a little algebra that

$$r_{1n} \le 2(c_{2n} - c_{1n})I_1 + 2I_2 + 2I_3, \tag{6.2}$$

where $I_1 = \int_{c_{1n}}^{\eta_1} P(W_n(x) \le 0)dx$, $I_2 = (\int_{\eta_1}^{c_G} I_{[w(x)\le d_n]}dx)^2$, $I_3 = E[\int_{\eta_1}^{c_G} I_{[W_n(x)\le 0, w(x) > d_n]}dx]^2$. For $w(x) > d_n$, $w_n(x) > 1/2w(x)$ from Lemma 6.4. Then we have

$$P(W_n(x) \le 0) = P(\frac{1}{\sqrt{n\sigma_n^2}}\sum_{j=1}^n Z_{jn} \le \frac{-\sqrt{n}w_n(x)}{\sigma_n}) \le P(\frac{1}{\sqrt{n\sigma_n^2}}\sum_{j=1}^n Z_{jn} \le \frac{-\sqrt{n}w(x)}{2\sigma_n}).$$

Applying Theorem 5.16 on page 168 in Petrov (1995) to the LHS of the above inequality,

$$P(W_n(x) \le 0) \le \Phi(-\frac{\sqrt{n}w(x)}{2\sigma_n}) + \frac{8A\gamma_n(x)}{\sqrt{n}[2\sigma_n + \sqrt{n}w(x)]^3} \equiv S_n(x) + T_n(x), \tag{6.3}$$

where A is a constant and $\Phi(\cdot)$ is the cdf of $N(0,1)$. For $x \in [c_{1n}, \eta_1]$, $w(x) \ge Mu(\ln n)^{-B}$ and certainly $w(x) > d_n$ as n is large. Also note that $\sigma_n \le l_1 u^{-5/2}v^{3/2}$ and $\gamma_n(x) \le l_4 v^{13} 36^v u^{-6}$.

It follows that $S_n(x) \le \Phi(-n^{1/4})$ and $T_n(x) \le n^{-3/2}$ for large n. Thus

$$(c_{2n} - c_{1n})I_1 = (c_{2n} - c_{1n}) \int_{c_{1n}}^{\eta_1 \vee c_{1n}} P(W_n(x) \le 0)dx = o(n^{-1}). \tag{6.4}$$

For $x \in [\eta_1, c_G]$, $|w'(x)| \ge A_\epsilon$. Thus $I_2 \le A_\epsilon^{-2}[\int_{\eta_1}^{c_G} I_{[w(x) \le d_n]} w'(x)dx]^2$. Letting $y = w(x)/d_n$, $I_2 \le A_\epsilon^{-2} d_n^2 \int_0^\infty I_{[y \le 1]}dy = A_\epsilon^{-2} d_n^2$. Therefore

$$I_2 = O(d_n^2) = o((\ln n)^3/(n\epsilon_n)), \tag{6.5}$$

By Holder inequality again,

$$I_3 \le \int_{\eta_1}^{c_G} P(W_n(x) \le 0)[w(x)]^{3/2} I_{[w(x)>d_n]}dx \times \int_{\eta_1}^{c_G} [w(x)]^{-3/2} I_{[w(x)>d_n]}dx.$$

Letting $y = w(x)/d_n$, $\int_{\eta_1}^{c_G} [w(x)]^{-3/2} I_{[w(x)>d_n]}dx \le 2/[A_\epsilon \sqrt{d_n}]$. Using the previous two inequalities and (6.3), we have

$$I_3 \le 2/(A_\epsilon {d_n}^{1/2})\{\int_{\eta_1}^{c_G} S_n(x)[w(x)]^{3/2}dx + \int_{\eta_1}^{c_G} T_n(x)[w(x)]^{3/2}dx\}. \tag{6.6}$$

For $x \in [\eta_1, c_G]$, $l_2 \le \sigma_n \le l_3\sqrt{v^3/u^3}$ and $\gamma_n(x) \le l_4 v^{13} 36^v u^{-6}$. Therefore

$$\int_{\eta_1}^{c_G} S_n(x) w^{\frac{3}{2}}(x)dx \le \frac{1}{A_\epsilon} \int_{\eta_1}^{c_G} \Phi(-\frac{\sqrt{nu^3}w(x)}{2l_3\sqrt{v^3}})[w(x)]^{\frac{3}{2}}dw(x) \le \frac{(2l_3 d_n)^{5/2}}{A_\epsilon} \int_0^\infty \Phi(-y)y^{\frac{3}{2}}dy, \tag{6.7}$$

and

$$\int_{\eta_1}^{c_G} T_n(x)[w(x)]^{3/2}dx \le \frac{8Al_4 v^{13} 36^v}{A_\epsilon n^3 u^6} \int_0^\infty \frac{y^{3/2}}{[2l_2 + y]^3}dy. \tag{6.8}$$

Combining (6.6)-(6.8), we have $I_3 = o((\ln n)^3/(n\epsilon_n))$. This together with (6.4) and (6.5) yields $r_{1n} = o((\ln n)^3/(n\epsilon_n))$. Similarly $r_{2n} = o((\ln n)^3/(n\epsilon_n))$. Then $E(c_n - c_G)^2 = o((\ln n)^3/(n\epsilon_n))$. Similarly, $E(c_n - c_G)^4 = o((\ln n)^3/(n\epsilon_n))$. This completes the proof of Lemma 4.4.

6.6. Proof of Lemma 5.1. Let $w_1(x) = w(x)$ with $G \sim g_1$. Then $w_1(x) = m_1 w_0(x - x_d)$ and $c_1 = c_0 + x_d$. Since $w_1(c_0) > 0$ and $w_0(c_1) < 0$, $c_G \in [c_0, c_1]$ for $G \in \mathcal{G}_0$. Since

$w_0'(x) < 1/2w_0'(c_0)$ for $x \in [c_0 - x_d, c_0 + x_d]$, $-w'(x) > -(m_1 \wedge 1)w_0'(c_0)/2 \equiv \underline{w} > 0$ for $x \in [c_0, c_1]$ and $G \in \mathcal{G}_0$.

Let $\bar{\mathcal{C}} = \{c_n^* \vee c_0 \wedge c_1 : c_n^* \in \mathcal{C}\}$. For $c_n^* \in \mathcal{C}$, denote $\bar{c}_n = c_n^* \vee c_0 \wedge c_1$. Note that $h(x)$ is bounded on $[c_0, c_1]$. Then for any $G \in \mathcal{G}_0$, $\int_{c_n^*}^{c_G} w(x)h(x)dx \geq l_1 \int_{\bar{c}_n}^{c_G} w(x)dx$. From (2.3)

$$\inf_{\delta_n^* \in \mathcal{D}} \sup_{G \in \mathcal{G}_0} [R(G, \delta_n^*) - R(G, \delta_G)] \geq l_1 \inf_{c_n^* \in \mathcal{C}} \sup_{G \in \mathcal{G}_0} E[\int_{\bar{c}_n}^{c_G} w(x)dx].$$

By Taylor expansion, $\int_{\bar{c}_n}^{c_G} w(x)dx = -1/2 \times w'(\tilde{c}_n^*)(\bar{c}_n - c_G)^2 \geq 1/2\underline{w}(\bar{c}_n - c_G)^2$. Therefore

$$\inf_{c_n^* \in \mathcal{C}} \sup_{G \in \mathcal{G}_0} E[\int_{\bar{c}_n}^{c_G} w(x)dx] \geq l_2 \inf_{c_n^* \in \mathcal{C}} \sup_{G \in \mathcal{G}_0} E[(\bar{c}_n - c_G)^2].$$

Since $\bar{\mathcal{C}} \subset \mathcal{C}$,

$$\inf_{c_n^* \in \mathcal{C}} \sup_{G \in \mathcal{G}_0} E[(\bar{c}_n - c_G)^2] = \inf_{\bar{c}_n \in \bar{\mathcal{C}}} \sup_{G \in \mathcal{G}_0} E[(\bar{c}_n - c_G)^2] \geq \inf_{c_n^* \in \mathcal{C}} \sup_{G \in \mathcal{G}_0} E[(c_n^* - c_G)^2].$$

From the results in Donoho and Liu (1991) (Theorem 3.1 and the remark after Lemma 3.3),

$$\inf_{c_n^* \in \mathcal{C}} \sup_{G \in \mathcal{G}_0} E[(c_n^* - c_G)^2] \geq l_1 \sup\{(c_1 - c_2)^2 : \int [\sqrt{f_1(x)} - \sqrt{f_2(x)}]^2 dx \leq l_2/n, \quad f_1, f_2 \in \mathcal{F}\}.$$

Then Lemma 5.1 follows.

6.7. Proof of Lemma 5.2. Note that $f_2(x) - f_1(x) = (1 + \sqrt{n})^{-1}[-f_1(x) + f_0(x)]$, where $f_0(x) = \int c(\theta)\exp(\theta x)h(x)g_0(\theta)d\theta$. For all $x \in (a, b)$

$$f_0(x)[f_1(x)]^{-1} = [\int_{\theta_{01}}^{\theta_{02}} \exp(\theta(x - c_0))d\theta] \cdot [m_1 \int_{\theta_{01}}^{\theta_{02}} \exp(\theta(x - x_d - c_0))d\theta]^{-1} \leq l_1.$$

Then $\int[\sqrt{f_1(x)} - \sqrt{f_2(x)}]^2 dx \leq \int [f_1(x) - f_2(x)]^2/f_1(x)dx \leq (1 + l_2)/n$.

Denote $w_2(x) = w(x)$ with $G \sim g_2$. Then $w_2(x) = (1 + \sqrt{n})^{-1}[\sqrt{n}m_1 w_0(x - x_d) + w_0(x)]$. Note that $|w_2'(x)| \leq l_3$ for $x \in (c_0, c_1)$ and $|w_2(c_1)|^2 = [w_2(c_2) - w_2(c_1)]^2 \leq l_3^2(c_2 - c_1)^2$, Then $(c_2 - c_1)^2 \geq l_4|w_2(c_1)|^2 = l_4(1 + \sqrt{n})^{-2}[w_0(c_1)]^2$. The proof of Lemma 5.2 is complete now.

Appendix.

Lemma A.1. *The following statements hold.*

(i) $|K_{iv}(y)| \leq kv^{10}36^v$, $i = 0, 1$, k *is some constant.*

(ii) $v^{-1} \int |K_{0v}(y)|^2 dy \to \pi^{-1}$.

(iii) $v^{-3} \int |K_{1v}(y)|^2 dy \to (3\pi)^{-1}$.

Proof. (i) is obtained by simple calculations. It is omitted here. From our definition of K_{0v} and K_{1v}, and Theorem 1 of Gasser, Muller and Mammitzsch (1985), for an even v

$$\int_{-1}^{1} K_{0v}^2(y)dy = \frac{v^2[(v-1)!!]^2}{2[v!!]^2}, \qquad \int_{-1}^{1} K_{1v}^2(y)dy = \frac{v^2[(v+1)!!]^2}{6[v!!]^2}.$$

Since $s[(2s-1)!!]^2/[(2s)!!]^2 \to \pi^{-1}$ as $s \to \infty$, (ii) and (iii) are obvious. The case of odd v can be proved similarly.

Proof of Lemma 6.1. Note that $\alpha_G''(x) = \int \theta^2 c(\theta) e^{\theta x} dG(\theta) > 0$ for $x \in (a, b)$. Then $\alpha_G(x)$ is a convex function and $\bar{\alpha}_n = \alpha_G(a_n) \vee \alpha_G(b_n)$. We prove $\alpha_G(a_n) \leq (2u)^{-1}$ in the following. The proof of $\alpha_G(b_n) \leq (2u)^{-1}$ is similar. Since $c(\theta) = 1/\{\int_a^b h(x) e^{\theta x} dx\}$ and $\alpha_G(a_n) = \int c(\theta) e^{\theta a_n} dG(\theta)$, it follows

$$\alpha_G(a_n) \leq \int_{[\theta \geq 0]} \frac{1}{\int_{a_n}^{\bar{b}_n} h(x) \exp(\theta(x - a_n)) dx} dG(\theta) + \int_{[\theta < 0]} \frac{1}{\int_{\bar{a}_n}^{a_n} h(x) \exp(\theta(x - a_n)) dx} dG(\theta).$$

Note that $\int_{a_n}^{\bar{b}_n} \exp(\theta(x - a_n)) h(x) dx \geq 2u$ as $\theta \geq 0$ and $\int_{\bar{a}_n}^{a_n} \exp(\theta(x - a_n)) h(x) dx \geq 2u$ as $\theta < 0$ from Lemma 4.1. Then Lemma 6.1 holds.

Proof of Lemma 6.2. Since $\psi_G(x) = \int \theta c(\theta) \exp(\theta x) dG(\theta)$ and $u|\theta| \leq \exp(u|\theta|)$,

$$|\psi_G(x)| \leq u^{-1}[\int_{[\theta \geq 0]} c(\theta) \exp(\theta(x+u)) dG(\theta) + \int_{[\theta < 0]} c(\theta) \exp(\theta(x-u)) dG(\theta)].$$

From Lemma 6.1, for $x \in [c_{1n}, c_{2n}]$, $\alpha_G(x) \leq 1/(2u)$. Then $|\psi_G(x)| \leq 1/u^2$ and $|w(x)| \leq 2/u^2$ as n is large. Assume that $B > 0$ such that $\int_{[|\theta| \leq B]} dG(\theta) > 0$. Denote $\Omega_B = \Omega[|\theta| \leq B]$. Since $1/c(\theta)$ is a convex function of θ on Ω and therefore $c(\theta)$ is bounded on Ω_B. Thus $\int_{\Omega_B} c(\theta) dG(\theta)$ is finite.

Recall that $w(x) = \alpha_G(x)[\theta_0 - \phi_G(x)]$. Since $\phi_G(x)$ is increasing and $\phi_G(c_G) = 0$, then for $x \in [c_{1n}, \eta_1]$, $\theta_0 - \phi_G(x) \geq \theta_0 - o_G(\eta_1) > 0$; for $x \in [\eta_2, c_{2n}]$, $\phi_G(x) - \theta_0 \geq \phi_G(\eta_2) - \theta_0 > 0$. For $x \in [c_{1n}, c_{2n}]$, $|x| \leq \ln\ln n$ and

$$\alpha_G(x) \geq \int_{\Omega_B} c(\theta)\exp(-\theta|\ln\ln n|)dG(\theta) \geq (\ln n)^{-B}\int_{\Omega_B} c(\theta)dG(\theta).$$

Let $M = \{[\theta_0 - \phi_G(\eta_1)] \wedge [\phi_G(\eta_2) - \theta_0]\} \cdot \int_{\Omega_B} c(\theta)dG(\theta)$. Then Lemma 6.2 is proved.

Proof of Lemma 6.3. We prove (i) for even v only. It is similar for odd v. Using Taylor expansion of $e^{\theta ux}$, simple calculations show that

$$E[\frac{K_{0v}(\frac{X_j-x}{u})}{uh(X_j)}] = \int c(\theta)e^{\theta x}dG(\theta) + u^v\int \theta^v c(\theta)e^{\theta x}[\int_{-1}^{1}\frac{K_{0v}(t)t^v e^{\theta ut^*}}{v!}dt]dG(\theta),$$

and

$$E[\frac{K_{1v}(\frac{X_j-x}{u})}{u^2h(X_j)}] = \int \theta c(\theta)e^{\theta x}dG(\theta) + u^v\int \theta^{v+1}c(\theta)e^{\theta x}[\int_{-1}^{1}\frac{K_1(t)t^{v+1}e^{\theta ut^{**}}}{(v+1)!}dt]dG(\theta),$$

where $|t^*|$, $|t^{**}| < 1$. Then $E[V_n(X_j, x)] = w(x) + u^{v/2}d_n(x)$ and

$$\begin{aligned} d_n(x) &= \theta_0 u^{v/2}\int \frac{\theta^v}{v!}c(\theta)e^{\theta x}[\int_{-1}^{1} K_{0v}(t)t^v e^{\theta ut^*}dt]dG(\theta) \\ &\quad -u^{v/2}\int \frac{\theta^{v+1}}{(v+1)!}c(\theta)e^{\theta x}[\int_{-1}^{1} K_{1v}(t)t^{v+1}e^{\theta ut^{**}}dt]dG(\theta). \end{aligned}$$

Since $(u^{1/3}\theta)^v/v! \leq \exp(|\theta|u^{1/3})$ and $(u^{1/3}\theta)^{v+1}/(v+1)! \leq \exp(|\theta|u^{1/3})$, for $x \in [c_{1n}, c_{2n}]$

$$\begin{aligned} |d_n(x)| &\leq u^{v/6-1}\int c(\theta)e^{\theta x+|\theta|u+|\theta|u^{1/3}}dG(\theta)\cdot[|\theta_0|\int_{-1}^{1}|K_{0v}(t)|dt + \int_{-1}^{1}|K_{1v}(t)|dt]. \\ &\leq u^{v/6-1}\bar{\alpha}_n\{|\theta_0|[2\int_{-1}^{1}|K_{0v}(y)|^2dy]^{1/2} + [2\int_{-1}^{1}|K_{1v}(y)|^2dy]^{1/2}\}. \end{aligned}$$

From Lemma A.1 and Lemma 6.1, $|d_n(x)| \to 0$ uniformly for $x \in [c_{1n}, c_{2n}]$. Then (i) is proved. Next we prove (ii). For $x \in [c_{1n}, c_{2n}]$, $h(x+u) \geq u$ from Lemma 4.1 and

$$\begin{aligned} \sigma_n^2(x) &\leq E[\theta_0\frac{K_{0v}(\frac{X_j-x}{u})}{uh(X_j)} - \frac{K_{1v}(\frac{X_j-x}{u})}{u^2h(X_j)}]^2 \\ &= u^{-3}\int\int_{-1}^{1}[\theta_0 uK_{0v}(t) - K_{1v}(t)]^2c(\theta)e^{\theta x}e^{\theta ut}[h(x+ut)]^{-1}dtdG(\theta) \\ &\leq l_1^2u^{-4}v^3\int c(\theta)e^{\theta x}e^{|\theta|u}dG(\theta) \\ &\leq l_1^2u^{-5}v^3. \end{aligned}$$

Especially, for $x \in [\eta_1, \eta_2]$, letting $\underline{h} = \min\{h(x+ut) : x \in [\eta_1, \eta_2], |t| \le 1\}$,

$$\sigma_n^2(x) \le l_5 u^{-3} \bar{h}^{-1} v^3 \int c(\theta) e^{\theta x} e^{|\theta| u} dG(\theta) \le l_3^2 u^{-3} v^3.$$

It is easy to see that $\sigma_n^2(x) > l_2^2$. We prove (iii) next. From Lemma A.1, for $i = 0$ or 1, $|K_{iv}(t)| \le kv^{10}36^v$. Also note that $|K_{iv}(t)| = 0$ if $|t| > 1$. Then

$$|K_{iv}((y-x)/u)/h(y)| I_{[c_{1n} \le x \le c_{2n}]} \le kv^{10}36^v/h(y) I_{[c_{1n} \le y \le c_{2n}+u]} \le kv^{10}36^v u^{-1}.$$

For $x \in [c_{1n}, c_{2n}]$, $E[|Z_{jn}(x)|^3] \le 2kv^{10}36^v u^{-1} E[Z_{jn}^2(x)] \le l_4 v^{13} 36^v u^{-6}$. The proof of Lemma 6.3 is completed.

Proof of Lemma 6.4. From lemma 6.3, we have that $|w_n(x) - w(x)| \le 1/\sqrt{n}$ for all $x \in [c_{1n}, c_{2n}]$. If $w(x) > d_n$ and n is large,

$$\frac{w_n(x)}{w(x)} \ge \frac{w(x) - d_n + d_n - |w_n(x) - w(x)|}{w(x) - d_n + d_n} \ge \frac{d_n - |w_n(x) - w(x)|}{d_n} \ge \frac{1}{2}.$$

Similarly, we can prove that $w(x) < -d_n \Longrightarrow w_n(x) \le w(x)/2$.

References

[1] Brown, L. D., Cohen, A. and Strawderman W. E. (1976). A complete class theorem for strict monotone likelihood ratio with applications. *Ann. Statist.* **4**, 712-722.

[2] Donoho, D. L. and Liu, R. C. (1991). Geometrizing rates of convergence. II. *Ann. Statist.* **19**, 633-667.

[3] Fan, J. (1991). On the optimal rates of convergence for nonparametric deconvolution problems. *Ann. Statist.* **19**, 1257-1272.

[4] Gasser, T., Muller, H-G and Mammitzsch, V. (1985). Kernels for nonparametric curve estimation. *J. R. Statist. Soc. B.* **47**, 238-252.

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

[6] Gupta, S. S. and Li, J. (1999b). An empirical Bayes selection rule for positive exponential family. *Technical Report #99-25, Department of Statistics, Purdue University.*

[7] Gupta, S. S. and Li, J. (2000). Optimal rate of convergence of monotone empirical Bayes tests for the normal family.

[8] Johns, M. V., Jr. and Van Ryzin, J. R. (1972). Convergence rates for empirical Bayes two-action problems, II. Continuous Case. *Ann. Math. Statist.* **43**, 934-947.

[9] Karunamuni, R. J. (1996). Optimal rates of convergence of monotone empirical Bayes tests for the continuous one-parameter exponential family. *Ann. Statist.* **43**, 934-947.

[10] Karunamuni, R. J. and Yang, H. (1995). On convergence rates of monotone empirical Bayes tests for the continuous one-parameter exponential family. *Statistics and Decisions* **13**, 181-192.

[11] Li, J. and Gupta, S. (2000). Empirical Bayes tests with optimal rate for a truncation paprameter. *To appear in Statistics and Decisions.*

[12] Liang, T. (2000a). On an empirical Bayes test for a normal mean. *Ann. Statist.* **28**, 648-655.

[13] Liang, T. (2000b). Empirical Bayes testing for a normal mean. Personal Communication. Submitted.

[14] Petrov, V. V. (1995). *Limit Theorems of Probability Theory.* Clarendon Press · Oxford.

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

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

[17]Singh, R. S. (1979). Empirical Bayes estimation in Lebesgue-exponential family with rates near the best possible rate. *Ann. Statist.* **7**, 890-902.

[18] Stijnen, T. (1985). On the asymptotic behaviour of empirical Bayes tests for the con-

tinuous one-parameter exponential family. *Ann. Statist.* **13**, 403-412.

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