

REPORT DOCUMENTATION PAGE

Form Approved  
OMB NO. 0704-0188

Public reporting burden for this collection of information is estimated to average 1 hour per response including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comment regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Budget, Paperwork Reduction Project (0704-0188), Washington, DC 20503.

1. AGENCY USE ONLY (Leave blank)		2. REPORT DATE March 2001	3. REPORT TYPE AND DATES COVERED Technical Report March 2001	
4. TITLE AND SUBTITLE Optimal Rate of Convergence of Monotone Empirical Bayes Tests for a Normal Mean			5. FUNDING NUMBERS DAAD 19-00-1-0502	
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 #01-04C	
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 40940.4-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.				
13. ABSTRACT (Maximum 200 words) This paper studies monotone empirical Bayes tests for a normal mean under a linear loss. The optimal rate of convergence of the monotone empirical Bayes tests is obtained. Applying a few techniques and using the non-uniform estimate of the remainder in the central limit theorem, we are able to construct a monotone empirical Bayes test and show that it achieves the best possible rate over a broad class of prior distributions, while the best possible rate is obtained through an idea of Donoho and Liu by constructing the "hardest two-point subproblem". This answers the question raised recently by Karunamuni and Liang. The result indicates that $n^{-1}$ may not be an attainable lower bound for the monotone empirical Bayes tests in the continuous one-parameter exponential family. A method to construct the monotone empirical Bayes test achieving the optimal rate is also discussed in this paper.				
14. SUBJECT TERMS Empirical Bayes; regret; optimal rate of convergence; minimax			15. NUMBER OF PAGES 14	
			16. PRICE CODE	
17. SECURITY CLASSIFICATION OF REPORT UNCLASSIFIED	18. SECURITY CLASSIFICATION OF THIS PAGE UNCLASSIFIED	19. SECURITY CLASSIFICATION OF ABSTRACT UNCLASSIFIED	20. LIMITATION OF ABSTRACT UL	

20010904 062

OPTIMAL RATE OF CONVERGENCE OF MONOTONE  
EMPIRICAL BAYES TESTS FOR A NORMAL MEAN

by

Shanti S. Gupta and Jianjun Li  
Purdue University

Technical Report #01-04

Department of Statistics  
Purdue University  
West Lafayette, IN USA

March 2001

# Optimal Rate of Convergence of Monotone Empirical Bayes Tests for a Normal Mean

Shanti S. Gupta      *and*      Jianjun Li

Department of Statistics      Department of Statistics

Purdue University      Purdue University

W. Lafayette, IN 47907      W. Lafayette, IN 47907

*Abstract:* This paper studies monotone empirical Bayes tests for a normal mean under a linear loss. The optimal rate of convergence of the monotone empirical Bayes tests is obtained. Applying a few techniques and using the non-uniform estimate of the remainder in the central limit theorem, we are able to construct a monotone empirical Bayes test and show that it achieves the best possible rate over a broad class of prior distributions, while the best possible rate is obtained through an idea of Donoho and Liu by constructing the “hardest two-point subproblem”. This answers the question raised recently by Karunamuni and Liang. The result indicates that  $n^{-1}$  may not be an attainable lower bound for the monotone empirical Bayes tests in the continuous one-parameter exponential family. A method to construct the monotone empirical Bayes test achieving the optimal rate is also discussed in this paper.

AMS 1991 *subject classification:* Primary 62C12; secondary 62F03, 62C20.

*Keywords and phrases:* Empirical Bayes, regret, optimal rate of convergence, minimax.

**1. Introduction.** Let  $X$  denote a  $N(\theta, 1)$  random variable, where  $\theta$  is the parameter, which is distributed according to an unknown prior distribution  $G$  on  $(-\infty, \infty)$ . We consider the problem of testing the hypotheses  $H_0 : \theta \leq 0$  versus  $H_1 : \theta > 0$ . The loss function is  $l(\theta, 0) = \max\{\theta, 0\}$  for accepting  $H_0$  and  $l(\theta, 1) = \max\{-\theta, 0\}$  for accepting  $H_1$ . A test  $\delta(x)$  is defined to be a measurable mapping from  $(-\infty, \infty)$  into  $[0, 1]$  so that  $\delta(x) = P\{\text{accepting } H_1 | X = x\}$ , 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  $\delta$  when  $G$  is a prior distribution. Given that  $E[|\theta|] < \infty$ , a Bayes test  $\delta_G$  is found as

$$\delta_G(x) = 1 \text{ if } E[\theta | X = x] \geq 0, \text{ and } \delta_G(x) = 0 \text{ if } E[\theta | X = x] < 0.$$

Because  $E[\theta | X = x]$  involves  $G$ , the above solution works only if the prior  $G$  is known. If  $G$  is unknown, this testing problem is formed as a compound decision problem and the empirical Bayes approach is used. Let  $X_1, X_2, \dots, X_n$  be the observations from  $n$  independent past experiences. Based on  $\widetilde{X}_n = (X_1, X_2, \dots, X_n)$  and  $X$ , an empirical Bayes rule  $\delta_n(X, \widetilde{X}_n)$  can be constructed. The performance of  $\delta_n$  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.

The empirical Bayes approach was introduced by Robbins (1956, 1964). Since then, it has been widely used in statistics. For its applications in testing problems, much research has been done. For example, Johns and Van Ryzin (1972) studied the empirical Bayes tests for the general continuous one-parameter exponential family. Van Houwelingen (1976) constructed the monotone empirical Bayes tests for the same family and showed that the tests have good performance for large samples and small samples as well. Stijnen (1985) studied the asymptotic behavior of both the monotone empirical Bayes rules and non-monotone rules. Karunamuni and Yang (1995) also studied monotone rules and their asymptotic behavior.

For the problem described above, Karunamuni (1996) “claimed” that he obtained the optimal rate of convergence of monotone empirical Bayes tests (in minimax sense). Later, Liang (2000a) and Liang (2000b) obtained a faster rate than Karunamuni’s “optimal rate”. So an interesting question arises: what is the optimal rate of empirical Bayes tests for the normal mean? We shall answer the question in this paper.

After introducing some preliminary results in Section 2, we start our answer with considering monotone empirical Bayes tests for a single prior in Section 3. A method to construct monotone empirical Bayes tests is suggested. A typical rule is constructed from this method and an upper bound of its regret is obtained using the non-uniform estimate of the remainder in the central limit theorem (Theorem 3.1). In Section 4, we use the results in Section 3 to get a upper bound of monotone empirical Bayes tests over a broad class of prior distributions (Theorem 4.1). And a lower bound is obtained by careful construction of the “hardest 2-point subproblem” (Lemma 4.3 and Theorem 4.2). Then we find the optimal rate of monotone empirical Bayes tests. And clearly, all the empirical tests based on the method in Section 3 achieve the optimal rate. All proofs are given in Section 5.

**2. Preliminary.** To ensure that the Bayes analysis can be carried out, we assume  $\mu_G \equiv \int |\theta| dG(\theta) < \infty$ . Also, assume  $P(\theta > 0) \cdot P(\theta < 0) > 0$  in the following. If  $P(\theta > 0) = 0$  or  $P(\theta < 0) = 0$ , it is known which action one should take regardless of the value of  $x$ . So both these cases are excluded from the decision problem.

Denote the density of  $X$  by  $f(x|\theta) = c(\theta) \exp(\theta x) h(x)$ , where  $c(\theta) = \exp(-\theta^2/2)/\sqrt{2\pi}$  and  $h(x) = \exp(-x^2/2)$ . Let  $f_G(x) = \int f(x|\theta) dG(\theta)$  be the marginal density of  $X$ . Denote  $\phi_G(x) = E[\theta|X = x]$  and  $w(x) = -\int \theta f(x|\theta) dG(\theta) = -f_G(x) \phi_G(x)$ . Since  $\mu_G < \infty$ ,  $f_G(x)$ ,  $\phi_G(x)$  and  $w(x)$  are infinitely differentiable.

Noting that  $f_G(x) > 0$  and  $\phi_G(x)$  is increasing, the Bayes rule stated in Section 1 can be

represented as

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

where  $c_G = \sup\{x : w(x) > 0\}$ .  $c_G$  is called the critical point corresponding to  $G$ .

study of (1985)) for discussions

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

$$\delta_n = 1 \text{ if } x \geq c_n, \text{ and } \delta_n = 0 \text{ if } x < c_n. \quad (2.2)$$

Note that  $R(G, \delta) = \int_0^\infty \theta dG(\theta) + \int \delta(x)w(x)dx$ . Then the regret of  $\delta_n$  is expressed as

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

**3. A class of MEBT's.** Before considering the optimal rate of MEBT's over a class of prior distributions, we consider MEBT's for a single prior in this section.

Let  $k(x)$  be a kernel function of form  $k(x) = (2\pi)^{-1} \int \exp(itx)\lambda(t)dt$ , where  $\lambda(t)$  satisfies  $\lambda(t) = 1$  in a neighborhood of the origin. This type of kernels could be found in Devroye and Györfi (1985). Two typical examples are

$$k(x) = (\pi x)^{-1} \sin x \quad \text{or} \quad k(x) = (4/\pi x^2) \{[\sin(x/2)]^2 - [\sin(x/4)]^2\}.$$

See Hall and Marron (1988). MEBT's can be constructed based on these kernels and the asymptotic behaviour for the MEBT's is the same. For simplicity, we use  $k(x) = (\pi x)^{-1} \sin x$  in the following. For this  $k(x)$ ,  $\lambda(t) = I_{|t| \leq 1}$ . Let  $u = u_n = (\ln n)^{-1/2}$  ( $u_n = 1$  if  $n = 1$ ).

Denote

$$W_n(x) = n^{-1} \sum_{j=1}^n \{[k'((X_j - x)/u)/u^2] - [(X_j/u)k((X_j - x)/u)]\}. \quad (3.1)$$

It is shown later that  $W_n(x)$  is a consistent estimator of  $w(x)$ .

Liang (2000a, 2000b) have constructed empirical Bayes rules based on (3.1) by mimicking the Bayes rule (2.1). The approach we are using here is different from his.

Let  $\xi = \xi_n = (\ln \ln n)^{1/2}$ . Observe that  $c_G = \int_{-\xi}^{\xi} I_{[w(x)>0]} dx - \xi$  as  $n$  is large. Then define

$$c_n = \int_{-\xi}^{\xi} I_{[W_n(x)>0]} dx - \xi, \quad (3.2)$$

and propose  $\delta_n(x)$  as

$$\delta_n = 1 \quad \text{if} \quad x \geq c_n \quad \text{and} \quad \delta_n = 0 \quad \text{if} \quad x < c_n. \quad (3.3)$$

To consider the convergence rate of  $\delta_n$ , we first express the regret of  $\delta_n$  through  $c_n - c_G$ .

Throughout this section, assume that  $E[|\theta|] < \infty$  and  $P(\theta > 0) \cdot P(\theta < 0) > 0$ .

**Lemma 3.1.**  $-\infty < c_G < \infty$  and  $-w'(c_G) = \int \theta^2 f(c_G|\theta) dG(\theta) > 0$ .

**Lemma 3.2.** For  $\epsilon > 0$ , let  $A_\epsilon = \inf_{x \in [c_G - \epsilon, c_G + \epsilon]} [-w'(x)]$  and  $\bar{w}_\epsilon = \sup_{x \in [c_G - \epsilon, c_G + \epsilon]} |w'(x)|$ .

Then  $\exists \epsilon_G > 0$  such that for  $\epsilon < \epsilon_G$ ,  $A_\epsilon \geq A_{\epsilon_G} > 0$  and

$$R(G, \delta_n) - R(G, \delta_G) \leq 1/2 \bar{w}_\epsilon E[(c_n - c_G)^2] + \mu_G \epsilon^{-4} E[(c_n - c_G)^4]. \quad (3.4)$$

As  $n$  is large,  $c_G \in [-\xi, \xi]$  and  $c_n - c_G = -\int_{-\xi}^{c_G} I_{[W_n(x) \leq 0]} dx + \int_{c_G}^{\xi} I_{[W_n(x) > 0]} dx$ . To study the rate of  $c_n$  going to  $c_G$ , we rewrite  $W_n(x)$  as  $W_n(x) = n^{-1} \sum_{j=1}^n V_n(X_j, x)$ , where  $V_n(X_j, x) = [k'((X_j - x)/u)/u^2] - [(X_j/u)k((X_j - x)/u)]$ . Note that  $V_n(X_j, x)$  are i.i.d. for fixed  $x$  and  $n$ . So the non-uniform estimate of the remainder in the central limit theorem can be used to find  $P(W_n(x) > 0)$  and  $P(W_n(x) \leq 0)$  for each  $x \in [-\xi, \xi]$ . Combining the properties of  $w(x)$  on  $[-\xi, \xi]$ , the following result is derived in Subsection 5.3.

**Theorem 3.1.**  $\delta_n$  has a rate of convergence of  $(\ln n)^{1.5}/n$ . Moreover,

$$\lim_{n \rightarrow \infty} \left\{ n^{-1} (\ln n)^{1.5} [R(G, \delta_n) - R(G, \delta_G)] \right\} \leq [\pi\sqrt{3} \int \theta^2 f(c_G|\theta) dG(\theta)]^{-1}. \quad (3.5)$$

**Remark 3.1.** Liang (2000a) studied the problem under a critical condition that  $c_G \in [-A, A]$ . He constructed an empirical Bayes rule  $\delta_n^*$  with a rate  $(\ln n)^{1.5}/n$ . Later Liang (2000b) constructed another rule with rate  $(\ln n)^{1.5+\epsilon}/n$  without the assumption  $c_G \in [-A, A]$ . Since  $\delta_n^*$  requires  $c_G \in [-A, A]$  and  $A$  must be given in the construction of  $\delta_n^*$ ,  $\delta_n^*$  does not achieve the best possible rate as  $\delta_n$  does in Theorem 4.1 (below). To illustrate this, let  $\mathcal{G}_0 = \{G_i : P_{G_i}(\theta < 0) \cdot P_{G_i}(\theta > 0) > 0, i = 1, \dots, m\}$  be a finite set of prior distributions. Then  $\mathcal{G}_0 \subset \mathcal{G}$  for some (unknown)  $\mu_0, b$  and  $L$  ( $\mathcal{G}$  is defined in (4.1) below). From Theorem 4.1,  $\delta_n$  has the rate  $(\ln n)^{1.5}/n$  over  $\mathcal{G}_0$  clearly and  $\delta_n^*$  does not necessarily. Even though  $\delta_n^*$  has the rate  $(\ln n)^{1.5}/n$  for a single prior  $G$ ,  $\delta_n^*$  is not robust and the assumption  $c_G \in [-A, A]$  is difficult to check in applications.

**Remark 3.2.** In (3.2) we use the integration of  $I_{[W_n(x) > 0]}$ . This technique is similar to an idea used by Brown, Cohen, and Strawderman (1976), Van Houwelingen (1976) and Stijnen (1985). Another technique used in (3.2) is localization. We have the integration only from  $-\xi$  to  $\xi$  in (3.2) through localization. As  $n \rightarrow \infty$ ,  $[-\xi, \xi]$  expands to the whole interval. But it is a compact interval for each  $n$ . Instead of considering  $W_n(x)$  and  $w(x)$  for  $x \in (-\infty, \infty)$ , we consider them only for  $x \in [-\xi, \xi]$  and therefore many crucial properties of  $W_n(x)$  and  $w(x)$  can be obtained. For more mathematical details, see Lemma 5.1 and the proof of Theorem 3.1 in Subsection 5.3. Statistically, the rationale behind (3.2) is that, according to the monotonicity of  $\phi_G(x)$ , one would like to accept  $H_1$  if  $x$  is quite large and accept  $H_0$  if  $x$  is quite small. Here we use  $-\xi$  and  $\xi$  as cut-off points since  $c_n \in [-\xi, \xi]$ .

**Remark 3.3.** Theorem 3.1 gives a useful formula for estimating the constant in the

upper bound. For example, if  $G$  is symmetric with support  $[-1, 1]$  and  $P(|\theta| > 0.5) > 0.5$ , then  $\{\pi\sqrt{3} \int \theta^2 f(c_G|\theta) dG(\theta)\}^{-1} < 6.2$ .

**4. Optimal Rate.** We obtain the optimal rate over a broad class of prior distributions in this section. Define,

$$\mathcal{G} = \{G : \mu_G < \mu_0, |c_G| < b, \int \theta^2 f(c_G|\theta) dG(\theta) > L\}. \quad (4.1)$$

where  $\mu_0 > 0$ ,  $b > 0$  and  $L > 0$  may be unknown. We assume that  $\mathcal{G}$  is a broad class so that  $G(\theta) = N(\theta, 1) \in \mathcal{G}$ . Let  $\psi(x) = -\int \theta c(\theta) \exp(\theta x) dG(\theta)$ . Clearly,  $-\psi'(x) > 0$ . Actually,  $-\psi'(x)$  has a (positive) uniform lower bound on  $[-b, b]$  over  $\mathcal{G}$ .

**Lemma 4.1.** *For some  $\psi_0 > 0$ ,  $\inf_{G \in \mathcal{G}} \inf_{x \in [-b-1, b+1]} |\psi'(x)| > \psi_0$ .*

Following the proof of Theorem 3.1 and applying Lemma 4.1, we have the following theorem.

**Theorem 4.1.** *For some  $l > 0$ ,  $\sup_{G \in \mathcal{G}} [R(G, \delta_n) - R(G, \delta_G)] \leq l \cdot (\ln n)^{1.5}/n$ .*

Next we shall find a lower bound of MEBT's over  $\mathcal{G}$ . In the following, let  $l_1, l_2, \dots$  denote positive constants, which may have different values on different occasions.

Let  $\mathcal{C}$  be the set of estimators  $c_n^*$  of  $c_G$  and  $\mathcal{D}$  be the set of empirical Bayes rules of type (2.2) with  $c_n = c_n^* \in \mathcal{C}$ . Let  $\bar{\mathcal{C}} = \{c_n^* \vee (-b) \wedge b : c_n^* \in \mathcal{C}\}$ . For  $c_n^* \in \mathcal{C}$ , denote  $\bar{c}_n = c_n^* \vee (-b) \wedge b$ .

Then by Taylor expansion and Lemma 4.1

$$\int_{c_n^*}^{c_G} w(x) dx \geq h(b) \int_{\bar{c}_n}^{c_G} \psi(x) dx = -1/2h(b)\psi'(\bar{c}_n)(\bar{c}_n - c_G)^2 \geq l_1(\bar{c}_n - c_G)^2, \quad (4.2)$$

Note that  $R(G, \delta_n^*) - R(G, \delta_G) = E[\int_{c_n^*}^{c_G} w(x) dx]$  and  $\bar{\mathcal{C}} \subset \mathcal{C}$ .

$$\inf_{\delta_n^* \in \mathcal{D}} \sup_{G \in \mathcal{G}} [R(G, \delta_n^*) - R(G, \delta_G)] \geq l_1 \inf_{\bar{c}_n \in \bar{\mathcal{C}}} \sup_{G \in \mathcal{G}} E(\bar{c}_n - c_G)^2 \geq l_1 \inf_{c_n^* \in \mathcal{C}} \sup_{G \in \mathcal{G}} E(c_n^* - c_G)^2. \quad (4.3)$$

Let  $\mathcal{F} = \{f_G(x) : G \in \mathcal{G}\}$  and  $c_f$  be the critical point corresponding to  $f$ . For  $f_1, f_2 \in \mathcal{F}$ ,

let  $\chi^2(f_1, f_2) = \int \{f_1(x) - f_2(x)\}^2 f_1^{-1}(x) dx$  be the  $\chi^2$  distance of  $f_1$  and  $f_2$ . Then

$$\inf_{c_n^* \in \mathcal{C}} \sup_{G \in \mathcal{G}} E(c_n^* - c_G)^2 \geq l_1 \sup\{(c_{f_1} - c_{f_2})^2 : \chi^2(f_1, f_2) \leq l_2/n, \forall f_1, f_2 \in \mathcal{F}\}, \quad (4.4)$$

(4.4) was proved in Donoho and Liu (1991) and others. We shall find a lower bound of RHS of (4.4) through a careful construction of “hardest 2-point subproblem” (Donoho and Liu (1991)), i.e., we need to construct  $f_1$  and  $f_2$  such that the supremum of RHS of (4.4) is obtained. This type of construction is often used to find a lower bound for various problems; see Fan (1991, 1993) for example. But the construction for this empirical Bayes testing problem appears so different: we cannot find  $f_1$  first and then find  $f_2$  in the  $\chi^2$ -distance ball around  $f_1$ . Here the center of the ball is moving too. Let  $f_i(x) = \int f(x|\theta)g_i(\theta)d\theta$ , where

$$g_1(\theta) = m_1 c(\theta)[1 + u\theta I(\theta > 0)] \quad \text{and} \quad g_2(\theta) = m_2 [g_1(\theta) + u^v c(\theta)H(\sqrt{2}\theta/u)]$$

with (i)  $v$  such that  $u^{2v+1} = n^{-1}$ , (ii)  $m_i$  satisfies  $\int g_i(\theta)d\theta = 1$  for  $i = 1, 2$ , (iii)  $H(x) = (2\pi)^{-1} \int \lambda_H(t) \exp(itx) dt$ , and  $\lambda_H(t) = \exp(t^2/(2u^2))I_{[|t| \leq 1]}$ .

**Lemma 4.2.** *As  $n$  is large,  $f_i \in \mathcal{F}$ ,  $\chi^2(f_1, f_2) \leq l_2/n$ , and  $(c_{f_1} - c_{f_2})^2 \geq l_3 \cdot (\ln n)^{1.5}/n$ .*

Based on (4.3), (4.4) and Lemma 4.2, the next theorem follows naturally.

**Theorem 4.2.** *For some  $l > 0$ ,  $\inf_{\delta_n^* \in \mathcal{D}} \sup_{G \in \mathcal{G}} [R(G, \delta_n^*) - R(G, \delta_G)] \geq l \cdot (\ln n)^{1.5}/n$ .*

**Remark 4.1.** Theorem 4.2 tells us that the best possible rate of MEBT's is  $(\ln n)^{1.5}/n$ .

Based on Theorem 4.1 and Theorem 4.2, one sees that the optimal rate of convergence of MEBT's is  $(\ln n)^{1.5}/n$  and  $\delta_n$  achieves this optimal rate.

**Remark 4.2.** For a long time, it was thought that  $n^{-1}$  is a lower bound of empirical Bayes rule for the continuous exponential family (including the normal distribution); see Singh (1979) for his conjecture about the estimation problem. Surprisingly, we obtain that

the best possible rate for the normal distribution is  $(\ln n)^{1.5}/n$ . So, even through  $n^{-1}$  is a lower bound for general continuous exponential family (see Gupta and Li (2000)), we believe that  $n^{-1}$  is not obtainable.

## 5. Proofs.

**5.1. Proof of Lemma 3.1.** Note that  $P(\theta > \theta_\epsilon) > 0$  for some  $\theta_\epsilon > 0$ . And also

$$\phi_G(x) \geq \frac{\int_{-\infty}^{\theta_\epsilon} \theta c(\theta) \exp(\theta x - \theta_\epsilon x) dG(\theta) + \theta_\epsilon \int_{\theta_\epsilon}^{\infty} c(\theta) \exp(\theta x - \theta_\epsilon x) dG(\theta)}{\int_{-\infty}^{\theta_\epsilon} c(\theta) \exp(\theta x - \theta_\epsilon x) dG(\theta) + \int_{\theta_\epsilon}^{\infty} c(\theta) \exp(\theta x - \theta_\epsilon x) dG(\theta)}.$$

Then  $\lim_{x \rightarrow \infty} \phi_G(x) \geq \theta_\epsilon > 0$ . Therefore  $c_G < \infty$ . Similarly  $c_G > -\infty$ . It is clear that  $-w'(c_G) = \int \theta^2 f(c_G|\theta) dG(\theta) < \infty$ . This completes the proof of Lemma 3.1.

**5.2. Proof of Lemma 3.2.** Since  $w'(x)$  is continuous,  $A_{\epsilon_G} > 0$  for some  $\epsilon_G$ . As  $\epsilon < \epsilon_G$ ,

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

where  $\int_{c_n}^{c_G} w(x) dx \leq \int |w(x)| dx \leq \mu_G$  and by Taylor expansion

$$I_{|c_n - c_G| \leq \epsilon} \int_{c_n}^{c_G} w(x) dx = -1/2 \times w'(\hat{c}_n) (c_n - c_G)^2 I_{|c_n - c_G| \leq \epsilon} \leq 1/2 \bar{w}_\epsilon (c_n - c_G)^2.$$

**5.3. Proof of Theorem 3.1.** We prove Theorem 3.1 in two steps.

**Step 1:** We present two lemmas. Their proofs are in Subsection 5.7 and 5.8.

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

Let  $p = 1/\sqrt{4\pi\sqrt{3}}$ ,  $d_n = 1/\sqrt{nu^3}$  and  $q_n = 1 - (p\pi)^{-1} \mu_G u^{5/2}$ .

**Lemma 5.1.** *The following statements hold (as  $n \geq 5$ ).*

(i) For  $\epsilon > 0$ ,  $\exists M_\epsilon > 0$  such that  $|w(x)| > M_\epsilon (\ln n)^{-1}$  for  $x \in [-\xi, \xi] \setminus [c_G - \epsilon, c_G + \epsilon]$ .

(ii) For  $x \in (-\infty, \infty)$ ,  $|w_n(x) - w(x)| \leq \pi^{-1} \mu_G \cdot u \exp(-1/(2u^2)) \equiv d_{1n}$ .

(iii) For  $x \in [-\xi, \xi]$ ,  $\sigma_n \leq d_{2n} u^{-3/2}$ ,  $d_{2n} = (3\pi)^{-1/2} + u^{1/4}$ .

(iv) For  $x \in [-\xi, \xi]$ ,  $\gamma_n \leq \gamma u^{-5}$ ,  $\gamma = 1 + 2\mu_G$ .

**Lemma 5.2.** If  $x \in (-\xi, \xi)$  and  $w(x) > pd_n$ ,

$$P(W_n(x) \leq 0) \leq \Phi(-\sqrt{nu^3}q_n w(x)/d_{2n}) + A\gamma/\{q_n^3 u^5 n^2 [w(x)]^3\}. \quad (5.1)$$

If  $x \in (-\xi, \xi)$  and  $w(x) < -pd_n$ ,

$$P(W_n(x) > 0) \leq [1 - \Phi(-\sqrt{nu^3}q_n w(x)/d_{2n})] + A\gamma/\{q_n^3 u^5 n^2 |w(x)|^3\}. \quad (5.2)$$

where  $A$  is some constant and  $\Phi(\cdot)$  is the c.d.f. of  $N(0, 1)$ .

**Step 2:** We present the main proof. Since (3.4) holds for any small  $\epsilon$  and  $A_\epsilon \rightarrow -w'(c_G)$  as  $\epsilon \rightarrow 0$ , we only need to show that

$$\lim_{n \rightarrow \infty} \{(nu^3)E[(c_n - c_G)^2]\} \leq 2/(\pi\sqrt{3}[w'(c_G)]^2), \quad \lim_{n \rightarrow \infty} \{(nu^3)E[(c_n - c_G)^4]\} = 0. \quad (5.3)$$

Let  $I = \int_{-\xi}^{c_G} I_{[W_n(x) \leq 0]} dx$  and  $II = \int_{c_G}^{\xi} I_{[W_n(x) > 0]} dx$ . Then  $c_n - c_G = -I + II$ . For  $\epsilon < \epsilon_G$ , let  $\eta_1 = c_G - \epsilon$ . As  $n$  is large,  $\eta_1 > -\xi$ . Then  $I^2 \leq 2\xi I_1 + 2I_2^2 + 2I_3^2$ , where

$$I_1 = \int_{-\xi}^{\eta_1} I_{[W_n(x) \leq 0]} dx, \quad I_2 = \int_{\eta_1}^{c_G} I_{[w(x) \leq pd_n]} dx, \quad I_3 = \int_{\eta_1}^{c_G} I_{[W_n(x) \leq 0, w(x) > pd_n]} dx.$$

For  $x \in [-\xi, \eta_1]$ ,  $w(x) \geq M_\epsilon(\ln n)^{-1} > pd_n$  from Lemma 5.1. Then by Lemma 5.2

$$\xi E[I_1] \leq \xi \int_{-\xi}^{\eta_1} P(W_n(x) \leq 0) dx \leq l_1 \Phi(-n^{1/3}) + l_2 n^{-3/2} = o(n^{-1}). \quad (5.4)$$

For  $x \in [\eta_1, c_G]$ ,  $-w'(x) \geq A_\epsilon (\geq A_{\epsilon_G} > 0)$ . Thus by letting  $y = w(x)/(pd_n)$ ,

$$I_2 \leq A_\epsilon^{-1} \int_{\eta_1}^{c_G} I_{[w(x) \leq pd_n]} [-w'(x)] dx \leq pd_n A_\epsilon^{-1} \int_0^\infty I_{[y \leq 1]} dy = pd_n A_\epsilon^{-1}. \quad (5.5)$$

By Holder inequality and Lemma 5.2,

$$\begin{aligned} E[I_3^2] &\leq \left[ \int_{\eta_1}^{c_G} w^{-3}(x) I_{[w(x) > pd_n]} dx \right] \cdot \left[ \int_{\eta_1}^{c_G} P(W_n(x) \leq 0) w^3(x) I_{[w(x) > pd_n]} dx \right] \\ &\leq [(2A_\epsilon)^{-1} p^{-2} d_n^{-2}] \cdot [A_\epsilon^{-1} d_{2n}^4 q_n^{-4} (nu^3)^{-2} \int_0^\infty \Phi(-y) y^3 dy + A\gamma \epsilon q_n^{-3} u^{-5} n^{-2}] \end{aligned} \quad (5.6)$$

From (5.4)-(5.6),

$$\lim_{n \rightarrow \infty} \{(nu^3)E[I^2]\} = \lim_{\epsilon \rightarrow 0} \lim_{n \rightarrow \infty} \{(nu^3)E[I^2]\} \leq 1/(\pi\sqrt{3}[w'(c_G)]^2).$$

Similarly,  $\lim_{n \rightarrow \infty} \{(nu^3)E[II^2]\} \leq 1/(\pi\sqrt{3}[w'(c_G)]^2)$ . Then the first part of (5.3) is proved.

The second can be proved similarly. The details are omitted.

**5.4. Proof of Lemma 4.1.** Note that  $c_G \in [-b, b]$  and

$$L \leq \int \theta^2 c(\theta) \exp(\theta c_G) dG(\theta) \leq \int \theta^2 c(\theta) \exp(b|\theta|) dG(\theta) \leq l \cdot \int \theta^2 \exp(-\theta^2/4) dG(\theta).$$

Then we can find  $\theta_{01} > 0$ ,  $\theta_{02} > 0$  and  $\epsilon_\theta > 0$  such that  $P(\theta_{01} \leq |\theta| \leq \theta_{02}) \geq \epsilon_\theta$  for all  $G \in \mathcal{G}$ .

Therefore for  $x \in [-b-1, b+1]$ ,

$$|\psi'(x)| = \int \theta^2 c(\theta) \exp(\theta x) dG(\theta) \geq \int \theta^2 c(\theta) \exp(-|\theta|(b+1)) I_{[\theta_{01} \leq |\theta| \leq \theta_{02}]} dG(\theta) > 0.$$

**5.5. Proof of Theorem 4.1.** Based on the proof of Theorem 3.1, in order to prove Theorem 4.1, it is sufficient to show that there is  $0 < \epsilon_0 < 1$  such that as  $\xi > b+1$

- (a)  $\inf_{G \in \mathcal{G}} \inf_{x \in [-\xi, c_G - \epsilon_0] \cup [c_G + \epsilon_0, \xi]} |w(x)| > l_1 / \ln n;$
- (b)  $\inf_{G \in \mathcal{G}} \inf_{x \in [c_G - \epsilon_0, c_G + \epsilon_0]} [-w'(x)] > l_2 (> 0);$
- (c)  $\sup_{G \in \mathcal{G}} \sup_{x \in [c_G - \epsilon_0, c_G + \epsilon_0]} [-w'(x)] < l_3 (< \infty).$

Recall  $\psi(x) = -\int \theta c(\theta) \exp(\theta x) dG(\theta)$ . Then  $|\psi'(x)| \leq \int \theta^2 c(\theta) \exp((b+1)|\theta|) dG(\theta)$  for  $x \in [-b-1, b+1]$ . Note that  $\theta^2 c(\theta) \exp((b+1)|\theta|)$  is bounded. Therefore

$$\sup_{G \in \mathcal{G}} \sup_{x \in [-b-1, b+1]} |\psi'(x)| < l_4. \quad (5.7)$$

Let  $\epsilon_0 = [1/2\psi_0 h(b+1)/((b+1)l_4)] \wedge (1/2)$ , where  $\psi_0$  is defined in Lemma 4.1. It is easy to check that (a), (b) and (c) hold for this  $\epsilon_0$ . The details are omitted. theorem.

**5.6. Proof of Lemma 4.3.** We prove it in three steps.

**Step 1:** To prove  $f_i \in \mathcal{F}$  as  $n$  is large. Clearly  $g_1(\theta) > 0$  and  $m_1 \rightarrow 1$ . Simple algebra computations show that  $u^v |H(\sqrt{2}\theta/u)| \leq 2\sqrt{u}$  and  $|\int c(\theta) H(\sqrt{2}\theta/u) d\theta| \leq u$ . Then  $g_2(\theta) > 0$  as  $n$  is large and  $(1 - m_2)^2 = O(u^{2v+2})$ . Let  $w_i(x) = -\int \theta f(x|\theta) g_i(\theta) d\theta$  for  $i = 1$  and 2. One

can see that  $w_1(-u/3) > 0$  and  $w_1(-u) < 0$  as  $n$  is large. Therefore  $-u < c_{f_1} < -u/3$  and  $g_1 \in \mathcal{G}$  for large  $n$ . Similarly,  $g_2 \in \mathcal{G}$ . Therefore  $f_i(x) \in \mathcal{F}$ .

**Step 2:** To prove  $\chi^2(f_1, f_2) \leq l_2/n$ . Note that  $f_1(x) \geq m_1 \int c(\theta) f(x|\theta) d\theta \geq l_1 \exp(-x^2/4)$  and

$$[f_2(x) - f_1(x)]^2 \leq 2(1 - m_2)^2 f_1^2(x) + 2u^{2v} m_2^2 \left[ \int f(x|\theta) c(\theta) H(\sqrt{2}\theta/u) d\theta \right]^2.$$

Then

$$\chi^2(f_1, f_2) \leq O(u^{2v+2}) + l_1 u^{2v} \int \left[ \int \exp(-(\theta - x/2)^2) H(\sqrt{2}\theta/u) d\theta \right]^2 dx.$$

It turns out that  $\chi^2(f_1, f_2) = O(u^{2v+1}) \leq l_2/n$  since using Parseval identity

$$\begin{aligned} \int \left[ \int \exp(-(\theta - x/2)^2) H(\sqrt{2}\theta/u) d\theta \right]^2 dx &\leq l_3 \int \left[ \int \exp(-(\eta - y)^2/2) H(\eta/u) d\eta \right]^2 dy \\ &= l_4 u \int |\lambda_H(t)|^2 \exp(-t^2/u^2) dt \\ &\leq 2l_4 u. \end{aligned}$$

**Step 3:** To prove  $(c_{f_1} - c_{f_2})^2 \geq l_3(\ln n)^{1.5}/n$ . Note that  $|w_2'(x)|$  is bounded for all  $x \in [-b, b]$  and all  $n$ . Then  $[w_2(c_{f_1})]^2 = [w_2(c_{f_2}) - w_2(c_{f_1})]^2 \leq l_1(c_{f_2} - c_{f_1})^2$  and  $(c_{f_2} - c_{f_1})^2 \geq l_2[w_2(c_{f_1})]^2$ . Let  $x_0 = c_{f_1}/\sqrt{2}$ . Then  $-u/\sqrt{2} < x_0 < -u/(3\sqrt{2})$ . Using integration by parts,

$$\begin{aligned} |w_2(c_{f_1})| &= l_3 u^v \exp(-x_0^2/2) \cdot \left| \int \eta \exp(-(\eta - x_0)^2/2) H(\eta/u) d\eta \right| \\ &\geq l_4 u^{v-1} \left| \int \exp(-(\eta - x_0)^2/2) H'(\eta/u) d\eta \right| \\ &\quad - l_4 u^v |x_0| \left| \int \exp(-(\eta - x_0)^2/2) H(\eta/u) d\eta \right| \\ &\geq l_5 u^{v-1} \int_0^1 t \sin(t/6) dt - l_6 u^{v+1}. \end{aligned}$$

Then  $(c_{f_2} - c_{f_1})^2 \geq l_7 u^{2v-2} = l_7(\ln n)^{1.5}/n$ .

**5.7. Proof of Lemma 5.1.** For  $x \in (-\xi, \xi)$ ,  $h(x) \geq (\ln n)^{-1}$  and  $|w(x)| \geq (\ln n)^{-1} |\psi(x)|$ . Since  $\psi(x)$  is decreasing and  $\psi(c_G) = 0$ , then (i) holds with  $M_\epsilon = [|\psi(c_G - \epsilon)| \wedge |\psi(c_G + \epsilon)|]$ . (ii)-(iv) are simple algebra calculations. The details are omitted.

**5.8. Proof of Lemma 5.2.** if  $w(x) > pd_n$ ,

$$\frac{w_n(x)}{w(x)} \geq \frac{w(x) - pd_n + pd_n - d_{1n}}{w(x) - pd_n + pd_n} \geq \frac{pd_n - d_{1n}}{pd_n} = 1 - (p\pi)^{-1} \mu_G u^{5/2}.$$

Then

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

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

$$P(W_n(x) \leq 0) \leq \Phi(-\sqrt{n}q_n w(x)/\sigma_n) + A\gamma_n / \{\sqrt{n}[\sigma_n + \sqrt{n}q_n w(x)]^3\}.$$

Then (5.1) follows Lemma 5.1. (5.2) can be proved similarly. The details are omitted.

**Acknowledgment.** The author wish to thank Prof. Shanti S. Gupta for his encouragements and helpful comments.

This research was supported in part by a US Army Research Office Grant at Purdue University.

#### 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] Devroye, L. and Györfi, L. *Nonparametric Density Estimation: the  $L^1$  View*. Wiley, New York.
- [4] Fan, J. (1991). On the optimal rates of convergence for nonparametric deconvolution problems. *Ann. Statist.* **19**, 1257-1272.
- [5] Fan, J. (1993). Local linear regression smoothers and their minimax efficiencies. *Ann. Statist.* **21**, 196-216.
- [6] Gupta, S. S. and Li, J. (2000). Monotone empirical Bayes tests based on kernel sequence

estimation. *Purdue University, Department of Statistics, Technical Report.*

- [7] Hall, P. and Marron, J. (1988). Choice of kernel order in density estimation. *Ann. Statist.* **16**, 161-173.
- [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 empirical Bayes tests for the continuous one-parameter exponential family. *Ann. Statist.* **24**, 212-231.
- [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. for a truncation parameter.
- [11] Liang, T. (2000a). On an empirical Bayes test for a normal mean. *Ann. Statist.* **28**, 648-655.
- [12] Liang, T. (2000b). Empirical Bayes testing for a normal mean. *Personal Communication. Submitted.*
- [13] Petrov, V. V. (1995). *Limit Theorems of Probability Theory.* Clarendon Press · Oxford.
- [14] Robbins, H. (1956). An empirical Bayes approach to statistics. *Proc. Third Berkeley Symp. Math. Statist. Probab.* **1**, 157-163, University of California Press, Berkeley.
- [15] Robbins, H. (1964). The empirical Bayes approach to statistical decision problems. *Ann. Math. Statist.* **35** 1-20.
- [16] Singh, R. S. (1979). Empirical Bayes estimation in Lebesgue-exponential family with rates near the best possible rate. *Ann. Statist.* **7**, 890-902.
- [17] Stijnen, T. (1985). On the asymptotic behavior of empirical Bayes tests for the continuous one-parameter exponential family. *Ann. Statist.* **13**, 403-412.
- [18] Van Houwelingen, J. C. (1976). Monotone empirical Bayes tests for the continuous one-parameter exponential family. *Ann. Statist.* **4**, 981-989.