T E C H N I C A L
R E P O R T

0440

BAYES AND EMPIRICAL BAYES TESTS FOR THE LIFE PARAMETER

L. WANG

I A P S T A T I S T I C S
N E T W O R K

INTERUNIVERSITY ATTRACTION POLE

http://www.stat.ucl.ac.be/IAP

BAYES AND EMPIRICAL BAYES TESTS FOR THE LIFE PARAMETER *

Lichun Wang

Abstract. We study the one-sided testing problem for the exponential distribution via the empirical Bayes (EB) approach. Under a weighted linear loss function, a Bayes test is established. Using the past samples, we construct an EB test and exhibit its optimal rate of convergence. When the past samples are not directly observable, we work out an EB test by using deconvolution kernel method and obtain its asymptotic optimality.

2000 Mathematics Subject Classification: Primary 62C12; Secondary 62F12.
Key words and phrases: empirical Bayes, asymptotic optimality, rate of convergence, exponential distribution.

1. Introduction

Let us consider the problem of testing the hypothesis

$$H_0 : \theta \leq \theta_0 \longleftrightarrow H_1 : \theta > \theta_0 \tag{1.1}$$

in the exponential distribution

$$f_\theta(x) = \frac{1}{\theta}\exp(-\frac{x}{\theta}), \quad x > 0, \tag{1.2}$$

where $f_\theta(x)$ denotes the conditional probability density function (pdf) of random variable (r.v.) X given θ.

*Financial support from the IAP research network (# P5/24) of the Belgian Government (Belgian Science Policy) is gratefully acknowledged

In practice, the distribution (1.2) appears very often and is important, and it can be used to describe many models of survival analysis, reliability theory, engineering and environmental sciences. Usually, the data observed from this distribution denotes the lifetime of an individual in survival analysis and reliability problem. Since the expectation of r.v. X is equal to θ, we call θ the life parameter.

We adopt a weighted linear error loss function defined as follows

$$L(\theta, d_m) = (1-m)\frac{\theta-\theta_0}{\theta} I_{[\theta>\theta_0]} + m\frac{\theta_0-\theta}{\theta} I_{[\theta_0\geq\theta]}, \tag{1.3}$$

where d_m denotes an action in favor of H_m ($m = 0, 1$), and $I_{[A]}$ is the indicator of the set A. Obviously, the loss function (1.3) is more reasonable for the life parameter than the ordinary linear loss since it can remove the influence of the measurement unit. Suppose the parameter θ is distributed according to a prior $G(\theta)$ with support on $\Theta = (0, \infty)$.

Let

$$\delta(x) = P(\text{accepting } H_0 | X = x). \tag{1.4}$$

Then the Bayes risk of the test $\delta(x)$ is given by

$$\begin{aligned} R(\delta(x), G(\theta)) &= \int_0^\infty \int_\Theta [L(\theta, d_0)\delta(x) + L(\theta, d_1)(1-\delta(x))] f_\theta(x) dG(\theta) dx \\ &\hat{=} \int_0^\infty \alpha(x)\delta(x)dx + \int_\Theta \theta^{-1}(\theta_0-\theta) I_{[\theta_0\geq\theta]} dG(\theta) \end{aligned} \tag{1.5}$$

with

$$\alpha(x) = \int_\Theta \theta^{-1}(\theta-\theta_0) f_\theta(x) dG(\theta) = f(x) + \theta_0 f^{(1)}(x), \tag{1.6}$$

where $f(x) = \int_\Theta f_\theta(x) dG(\theta)$ is the marginal pdf of r.v. X, and $f^{(1)}(x)$ denotes the derivative of $f(x)$.

Hence, the best Bayes test minimizing $R(\delta(x), G(\theta))$ would have the form

$$\delta_G(x) = \begin{cases} 1 & \alpha(x) \leq 0 \\ 0 & \alpha(x) > 0 \end{cases}. \tag{1.7}$$

The minimum Bayes risk is

$$R(\delta_G(x), G(\theta)) = \int_0^\infty \alpha(x)\delta_G(x)dx + \int_\Theta \theta^{-1}(\theta_0-\theta) I_{[\theta_0\geq\theta]} dG(\theta). \tag{1.8}$$

Define $\beta(x) = \alpha(x)/f(x)$. Then by Cauchy-Schwarz inequality, it is easy to see that the derivative $\beta^{(1)}(x) \geq 0$. Assume that the prior $G(\theta)$ satisfies the following condition

$$\lim_{x\to\infty} \beta(x) > 0 > \lim_{x\to 0} \beta(x). \tag{1.9}$$

Obviously, the condition (1.9) implies that the prior $G(\theta)$ is non-degenerate and $\beta(x)$ is a strictly increasing function. Therefore, by the continuity of $\beta(x)$, we know there exists a unique point a_G such that $\beta(a_G) = 0$. Then

$$\delta_G(x) = \begin{cases} 1 & \alpha(x) \leq 0 \\ 0 & \alpha(x) > 0 \end{cases} = \begin{cases} 1 & \beta(x) \leq 0 \\ 0 & \beta(x) > 0 \end{cases} = \begin{cases} 1 & x \leq a_G \\ 0 & x > a_G \end{cases}. \tag{1.10}$$

Remark 1. In applications, suppose that the life parameter θ has a prior pdf

$$g(\theta) = \frac{dG(\theta)}{d\theta} = \frac{1}{\Gamma(b-2)}(\frac{1}{\theta})^{b-1}\exp(-\frac{1}{\theta}), \quad b > 2, \quad \theta > 0.$$

For example, let $b = 3$, we have $f(x) = (x+1)^{-2}, x > 0$. It is readily seen that $\beta(x) = 1 - 2\theta_0(x+1)^{-1}$ and $a_G = 2\theta_0 - 1$, then we get

$$\delta_G(x) = \begin{cases} 1 & x \leq 2\theta_0 - 1 \\ 0 & x > 2\theta_0 - 1 \end{cases}.$$

But in many situations, since the prior $G(\theta)$ may be unknown to us, the Bayes test $\delta_G(x)$ (1.10) is unavailable to use. As an alternative we can use the EB approach to estimate $\alpha(x)$ in (1.6) so as to obtain an EB test $\delta_n(x)$.

EB approach was first introduced to statistical problems by Robbins [6, 7] and has been applied in a wide range of paradigms and to numerous real-life problems. Some earlier papers such as [2], which discussed the EB testing problem for the discrete case, whereas, [8] and [9] concentrated on the EB testing problems in the continuous one-parameter exponential family. Recently the author of [5], who follows the results [3], [9] and [4], has considered the EB testing problem in a positive exponential family, and obtains a better rate of convergence under the assumption that the critical point a_G is within some known compact interval.

In this paper, we discuss the EB test problem for the life parameter in the exponential distribution, firstly, under the condition that the past samples are not contaminated and secondly, that they are contaminated.

The rest of the paper is organized as follows. In Section 2 we propose an EB test and exhibit the optimal convergence rate. In Section 3 we discuss the case when the past samples are contaminated by a normal error variable.

2. Empirical Bayes test and rate of convergence

In the empirical Bayes framework, we usually make the following assumptions: let $(X_i, \theta_i), i = 1, 2, \cdots$, be the independently identically distributed (i.i.d.) copies of the (X, θ), where $X_i,\ i = 1, 2, \cdots$, are observable, but $\theta_i,\ i = 1, 2, \cdots$, are not observable. At time $n+1$, we observe $X \hat{=} X_{n+1}$ and plan to test the hypothesis: $H_0 : \theta \leq \theta_0 \longleftrightarrow H_1 : \theta > \theta_0$, where $\theta \hat{=} \theta_{n+1}$. Usually, the $(X_1, \cdots, X_n)$ denote the past samples and X is the present sample.

In order to construct an EB test, we use two kernel functions $K_l(x)$ $(l = 0, 1)$ which are Borel measurable bounded real functions vanishing off (0,1) such that

$$\int_0^1 x^p K_l(x)dx = \begin{cases} 1 & p = l \\ 0 & p \neq l, p = 1, 2, \cdots, s-1 \end{cases}, \quad \int_0^1 x^s |K_l(x)|dx < \infty.$$

where $s \geq 2$ is an arbitrary but fixed integer. It is easy to find that there exist some polynomials which satisfy the above conditions.

Define the kernel estimation of $f(x)$ and $f^{(1)}(x)$, respectively , as

$$\begin{aligned} f_n(x) &= \frac{1}{nh_n} \sum_{i=1}^n K_0\left(\frac{X_i - x}{h_n}\right); \\ f_n^{(1)}(x) &= \frac{1}{nh_n^2} \sum_{i=1}^n K_1\left(\frac{X_i - x}{h_n}\right), \end{aligned} \tag{2.1}$$

where $0 < h_n \to 0$ $(n \to \infty)$ denotes the bandwidth.

Then, we have

$$\alpha_n(x) = f_n(x) + \theta_0 f_n^{(1)}(x). \tag{2.2}$$

We consider those prior distribution $G(\theta)$ for which $G(\theta) \in \mathcal{F} = \{G(\theta) : 0 < A_1 \leq a_G \leq A_2 < \infty,\ A_1, A_2 \text{ are known constants}\}$. By the fact that $G(\theta) \in \mathcal{F}$ and the Bayes test (1.10), we propose an EB test defined as follows:

$$\delta_n(x)=\begin{cases}1 & x<A_1 \quad \text{or} \quad (A_1\le x\le A_2 \quad \text{and} \quad \alpha_n(x)\le 0)\\ 0 & x>A_2 \quad \text{or} \quad (A_1\le x\le A_2 \quad \text{and} \quad \alpha_n(x)>0)\end{cases}. \tag{2.3}$$

Hence, the Bayes risk of $\delta_n(x)$ is

$$R(\delta_n(x),G(\theta))=\int_0^\infty \alpha(x)E_n[\delta_n(x)]dx+\int_\Theta \theta^{-1}(\theta_0-\theta)I_{[\theta_0\ge\theta]}dG(\theta), \tag{2.4}$$

where E_n denotes the expectation with respect to the joint distribution of $(X_1,\cdots,X_n)$.

By definition, the EB test $\delta_n(x)$ is said to be asymptotically optimal relative to the prior $G(\theta)$ if $R(\delta_n(x),G(\theta))-R(\delta_G(x),G(\theta))=o(1)$. If for some $q>0$, $R(\delta_n(x),G(\theta))-R(\delta_G(x),G(\theta))=O(n^{-q})$, then the convergence rate of the EB test $\delta_n(x)$ is said to be the order $O(n^{-q})$.

Remark 2. Usually, there are two different forms of EB test in the literature. One form, which used by [2, 3] and some other papers, suggested the following test for the above problem

$$\delta_n(x)=\begin{cases}1 & \alpha_n(x)\le 0\\ 0 & \alpha_n(x)>0\end{cases}.$$

They do not make the assumption that the critical point a_G is in the compact interval $[A_1,A_2]$, accordingly, they not consider the monotonicity of the $\beta(x)$. The other form, i.e. (2.3), appeared in [9], [4], and [5], which is named the monotone EB test. In author's opinion, the EB test $\delta_n(x)$ (2.3) is relatively reasonable since it divides the interval $(0,\infty)$ into three parts, but one will has to make some additional assumption about the critical point .

LEMMA 1. Let $f_n^{(l)}(x)$ $(l=0,1)$ be defined in (2.1). If $E(\theta^{-(s+1)})<\infty$, then

$$|E_nf_n^{(l)}(x)-f^{(l)}(x)|=O(h_n^{s-l}),\quad l=0,1,$$

where $s\ge 2$ is an arbitrary but fixed integer.

Proof. Expanding $f(x+uh_n)$ at point x and using the properties of the $K_l(x)$,

$$E_nf_n^{(l)}(x)=\frac{1}{h_n^{l+1}}\int_0^\infty K_l\left(\frac{y-x}{h_n}\right)f(y)dy=\frac{1}{h_n^l}\int_0^1 K_l(u)f(x+uh_n)du$$
$$=\frac{1}{h_n^l}\int_0^1\left[f(x)+\cdots+\frac{f^{(s-1)}(x)}{(s-1)!}(uh_n)^{s-1}+\frac{f^{(s)}(x+\xi uh_n)}{s!}(uh_n)^s\right]$$

$$
\begin{aligned}
&\times K_l(u)du \\
= \quad & f^{(l)}(x) + \frac{h_n^{s-l}}{s!}\int_0^1 K_l(u) f^{(s)}(x+\xi u h_n) u^s du, \quad 0<\xi<1. \qquad (2.5)
\end{aligned}
$$

By $E(\theta^{-(s+1)}) < \infty$, we know $\sup_x |f^{(s)}(x)| < \infty$. So LEMMA 1 is true.

We now represent $\alpha(x)$ by

$$
\begin{aligned}
\alpha(x) &= \frac{1}{nh_n}\sum_{i=1}^n K_0\left(\frac{X_i - x}{h_n}\right) + \theta_0 \frac{1}{nh_n^2}\sum_{i=1}^n K_1\left(\frac{X_i - x}{h_n}\right) \\
&\hat{=} \frac{1}{n}\sum_{i=1}^n R(x, X_i, h_n) \qquad (2.6)
\end{aligned}
$$

with

$$
R(x, X_i, h_n) = \frac{1}{h_n} K_0\left(\frac{X_i - x}{h_n}\right) + \frac{\theta_0}{h_n^2} K_1\left(\frac{X_i - x}{h_n}\right). \qquad (2.7)
$$

Note that $R(x, X_i, h_n)$ $(i = 1, \cdots, n)$ are i.i.d. r.v. such that

$$
|R(x, X_i, h_n) - E_n R(x, X_i, h_n)| \le 2\left(\frac{M_0}{h_n} + \frac{\theta_0}{h_n^2} M_1\right) \qquad (2.8)
$$

and

$$
\begin{aligned}
&Var(R(x, X_i, h_n)) \le \frac{2}{h_n^2} Var\left(K_0\left(\frac{X_i - x}{h_n}\right)\right) + \frac{2\theta_0^2}{h_n^4} Var\left(K_1\left(\frac{X_i - x}{h_n}\right)\right) \\
\le \quad & \frac{2}{h_n^2} E\left(K_0\left(\frac{X_i - x}{h_n}\right)\right)^2 + \frac{2\theta_0^2}{h_n^4} E\left(K_1\left(\frac{X_i - x}{h_n}\right)\right)^2 \\
= \quad & \frac{2}{h_n}\int_0^1 K_0^2(u) f(x + uh_n) du + \frac{2\theta_0^2}{h_n^3}\int_0^1 K_1^2(u) f(x + uh_n) du \\
\le \quad & 2c(h_n^{-1} + h_n^{-3}), \qquad (2.9)
\end{aligned}
$$

where $M_l > 0$ $(l = 0, 1)$ denote the bound of kernel function $K_l(x)$ $(l = 0, 1)$, and c is a positive constant that does not depend on n.

Denote $A_G = \min_{A_1 \le x \le A_2} f(x)$, and let $a_{1n} < a_G < a_{2n}$ be the point such that $-\beta(a_{1n}) = 2ch_n^{s-1}/A_G = \beta(a_{2n})$. Since $\beta(x)$ is continuous, we know that $\lim_{n\to\infty} a_{1n} = \lim_{n\to\infty} a_{2n} = a_G$.

It follows from (1.8) and (2.4)

$$
0 \le R(\delta_n(x), G(\theta)) - R(\delta_G(x), G(\theta)) = \int_0^\infty [E_n\delta_n(x) - \delta_G(x)]\alpha(x) dx
$$

$$
\begin{aligned}
&= \int_{A_1}^{a_G} [P(\alpha_n(x) \le 0) - 1]\alpha(x)dx + \int_{a_G}^{A_2} P(\alpha_n(x) \le 0)\alpha(x)dx \\
&= -\int_{A_1}^{a_{1n}} P(\alpha_n(x) > 0)\alpha(x)dx - \int_{a_{1n}}^{a_G} P(\alpha_n(x) > 0)\alpha(x)dx \\
&\quad + \int_{a_G}^{a_{2n}} P(\alpha_n(x) \le 0)\alpha(x)dx + \int_{a_{2n}}^{A_2} P(\alpha_n(x) \le 0)\alpha(x)dx \\
&\hat{=} \sum_{i}^{4} I_i. \qquad (2.10)
\end{aligned}
$$

It is easy to see that

$$
I_2 \le -\int_{a_{1n}}^{a_G} \alpha(x)dx \le -\beta(a_{1n}) \int_{a_{1n}}^{a_G} f(x)dx = O(h_n^{s-1}). \qquad (2.11)
$$

Similarly, we get

$$
I_3 = O(h_n^{s-1}). \qquad (2.12)
$$

Note that

$$
\alpha(x) \le \beta(a_{1n})f(x) \le \beta(a_{1n})A_G = -2ch_n^{s-1}, \quad A_1 \le x \le a_{1n}. \qquad (2.13)
$$

Furthermore, by LEMMA 1, we know $E_n\alpha_n(x) \le \alpha(x) + ch_n^{s-1} \le \alpha(x)/2$. Hence, for $A_1 \le x \le a_{1n}$, we have

$$
P(\alpha_n(x) > 0) \le P(\alpha_n(x) - E_n\alpha_n(x) \ge -\frac{1}{2}\alpha(x)). \qquad (2.14)
$$

Together with (2.8) and (2.9) and (2.14), by Bernstein's inequality, it generates

$$
\begin{aligned}
&P(\alpha_n(x) > 0) \\
&\le 2\exp\left\{\frac{-n^2(-\alpha(x)/2)^2}{2Var(\sum_{i=1}^n R(x, X_i, h_n)) + 4(M_0/h_n + \theta_0 M_1/h_n^2)(-n\alpha(x)/2)/3}\right\} \\
&= 2\exp\left\{\frac{-n(\alpha(x))^2/8}{Var(R(x, X_i, h_n)) + (M_0/h_n + \theta_0 M_1/h_n^2)|\alpha(x)|/3}\right\} \\
&\le 2\exp\left\{-\frac{nh_n^3}{8} \times \frac{A_G^2(\beta(x))^2}{2ch_n^2 + 2c + (M_0h_n^2 + \theta_0 M_1 h_n)E(\theta^{-1})|\beta(A_1)|/3}\right\} \\
&\hat{=} 2\exp\left\{-nh_n^3 J(h_n)(\beta(x))^2\right\}, \quad A_1 \le x \le a_{1n}, \qquad (2.15)
\end{aligned}
$$

where $J(h_n) = A_G^2/[8(2ch_n^2 + 2c + (M_0h_n^2 + \theta_0 M_1 h_n)E(\theta^{-1})|\beta(A_1)|/3)]$.

Following from (2.10) and (2.15), we have

$$\begin{aligned} I_1 &\le -2\int_{A_1}^{a_{1n}} \exp\left\{-nh_n^3 J(h_n)(\beta(x))^2\right\}\beta(x)f(x)dx \\ &\le -2\sup_{A_1\le x\le A_2}\left[\frac{f(x)}{\beta^{(1)}(x)}\right]\int_{A_1}^{a_{1n}} \exp\left\{-nh_n^3 J(h_n)(\beta(x))^2\right\}\beta(x)\beta^{(1)}(x)dx \\ &= O(\frac{1}{nh_n^3}). \end{aligned} \tag{2.16}$$

Similar to I_1, we get

$$I_4 = O(\frac{1}{nh_n^3}). \tag{2.17}$$

Combining (2.10)-(2.12) with (2.16) and (2.17) and taking $h_n = n^{-\frac{1}{s+2}}$, we conclude that

$$0 \le R(\delta_n(x), G(\theta)) - R(\delta_G(x), G(\theta)) = O(n^{-\frac{s-1}{s+2}}). \tag{2.18}$$

Hence, we state the following Theorem.

THEOREM 1. Let Bayes test $\delta_G(x)$ and EB test $\delta_n(x)$ be defined in (1.10) and (2.3), respectively. If $G(\theta) \in \mathcal{F}$, which is defined before, and $E(\theta^{-(s+1)}) < \infty$, then choosing $h_n = n^{-\frac{1}{s+2}}$, where $s \ge 2$ is an arbitrary but fixed integer, we have

$$R(\delta_n(x), G(\theta)) - R(\delta_G(x), G(\theta)) = O(n^{-\frac{s-1}{s+2}}).$$

Remark 3. If we use the following linear loss function

$$L(\theta, d_m) = (1-m)(\theta-\theta_0)I_{[\theta>\theta_0]} + m(\theta_0-\theta)I_{[\theta_0\ge\theta]}, \quad m = 0, 1,$$

then it is not difficult to see

$$\begin{aligned} \alpha(x) &= \int_\Theta (\theta-\theta_0)f_\theta(x)dG(\theta) = \int_\Theta \exp(-\frac{x}{\theta})dG(\theta) - \theta_0 f(x) \\ &= \int_x^\infty f(x)dx - \theta_0 f(x). \end{aligned}$$

Thus, we only need to estimate $f(x)$. Following a proof analogous to the preceding discussion, we can improve the rate of convergence $O(n^{-\frac{s-1}{s+2}})$ to the best rate $O(n^{-\frac{s}{s+1}})$ for the testing the hypothesis (1.1).

To our best knowledge, the convergence rate of the order $o(n^{-1})$ can not be attained with any EB test in any non-discrete density. Therefore, it is very hard to improve the rate of convergence $O(n^{-\frac{s-1}{s+2}})$ under the weighted linear loss function (1.3) since it tends to be $O(n^{-1})$ as s gets larger.

3. The case when the data is contaminated

Suppose that the past samples $(X_1, \cdots, X_n)$ are contaminated due to measurement or the nature of environment, one can only observe

$$Y_j = X_j + \epsilon, \quad j = 1, \cdots, n, \tag{3.1}$$

where the error variable ϵ has a known distribution F_ϵ. Furthermore, assume that ϵ and X_j are independent. Problems with contaminating error exist in many different fields (e.g., biostatistics, electrophoresis) and has been widely studied. [10] and [11] discussed the EB estimation for the continuous one-parameter exponential family with errors in variables under the squared loss function, and they obtained asymptotic optimality and uniform rate of convergence for the proposed EB estimator over a class of prior distribution.

In this section, we simply discuss the asymptotic behavior of EB test under the assumption that $\epsilon \sim N(0, \sigma^2)$ with σ^2 known.

Similar to [1], using the deconvolution kernel method, we make the following assumptions on the kernel

(1) $k(x)$ is bounded, continuous, and $\int_{-\infty}^{\infty} |x|^s |k(x)| dx < \infty$.

(2) The Fourier transform $\phi_k(t)$ of $k(x)$ is a symmetric function satisfying $\phi_k(t) = 1 + O(|t|^s)$, as $t \to 0$.

(3) $\phi_k(t)$=0, for $|t| \geq 1$.

where $s \geq 2$ is an arbitrary but fixed integer and $\phi_k(t) = \int_{-\infty}^{\infty} \exp(itx) k(x) dx$.

By assumptions (1)-(3), we can easily get

$$\int k(x)dx = 1, \quad \int x^p k(x)dx = 0, \quad p = 1, \cdots, s-1, \int |x|^s |k(x)| < \infty.$$

Note that $f(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \exp(-itx) \phi_Y(t) \exp(\sigma^2 t^2/2) dt$, we define an estimator of $f^{(l)}(x)$ $(l = 0, 1)$ by

$$\hat{f}_n^{(l)}(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \exp(-itx)(-it)^l \phi_k(tb_n) \hat{\phi}_n(t) \exp(\sigma^2 t^2/2) dt, \tag{3.2}$$

where $0 < b_n \to 0$ as $n \to \infty$, and $\hat{\phi}_n(t) = \frac{1}{n}\sum_{j=1}^n \exp(itY_j)$ is an estimator of the characteristic function (c.f) $\phi_Y(t)$ of r.v. Y, which is called the empirical c.f. of Y.

Rewrite (3.2) as kernel type of estimate

$$\hat{f}_n^{(l)}(x) = \frac{1}{nb_n^{l+1}}\sum_{j=1}^n k_{nl}\left(\frac{x-Y_j}{b_n}\right), \quad l = 0, 1, \tag{3.3}$$

where

$$k_{nl}(x) = \frac{1}{2\pi}\int_{-\infty}^{\infty} \exp(-itx)(-it)^l \phi_k(t)\exp(\sigma^2 t^2/(2b_n^2))dt. \tag{3.4}$$

Hence, under the assumption that $G(\theta) \in \mathcal{F}$, we propose an EB test as follows

$$\hat{\delta}_n(x) = \begin{cases} 1 & x < A_1 \quad \text{or} \quad (A_1 \le x \le A_2 \quad \text{and} \quad \hat{\alpha}_n(x) \le 0) \\ 0 & x > A_2 \quad \text{or} \quad (A_1 \le x \le A_2 \quad \text{and} \quad \hat{\alpha}_n(x) > 0) \end{cases}, \tag{3.5}$$

where

$$\hat{\alpha}_n(x) = \hat{f}_n(x) + \theta_0 \hat{f}_n^{(1)}(x) \hat{=} \frac{1}{n}\sum_{j=1}^n V(x, Y_j, b_n) \tag{3.6}$$

with i. i. d. r.v.

$$V(x, Y_j, b_n) = \frac{1}{b_n}k_{n0}\left(\frac{x-Y_j}{b_n}\right) + \frac{\theta_0}{b_n^2}k_{n1}\left(\frac{x-Y_j}{b_n}\right). \tag{3.7}$$

LEMMA 2. Let $\hat{f}_n^{(l)}(x)$ $(l = 0, 1)$ be given in (3.2). If $G(\theta) \in \mathcal{F}$ and $E(\theta^{-(s+1)}) < \infty$, where $\mathcal{F}$ is the same as before and $s \ge 2$ is an arbitrary but fixed integer, then

(a) $|E_n \hat{f}_n^{(l)}(x) - f(x)| = O(b_n^{s-l}), \quad l = 0, 1;$
(b) $|E_n \hat{\alpha}_n(x) - \alpha(x)| = O(b_n^{s-1});$
(c) $|V(x, Y_j, b_n) - E_n V(x, Y_j, b_n)| \le 2(b_n^{-1} + \theta_0 b_n^{-2}) O(\exp(\sigma^2 b_n^{-2}/2));$
(d) $Var(V(x, Y_j, b_n)) = 2(b_n^{-2} + \theta_0^2 b_n^{-4}) O(\exp(\sigma^2 b_n^{-2})).$

where E_n denotes the expectation with respect to the joint distribution of $(Y_1, \cdots, Y_n)$.

Proof. (a) By the assumptions (1)-(3) on $k(x)$, we have

$$E_n\,\hat{f}_n^{(l)}(x) - f^{(l)}(x) = \int f^{(l)}(x - b_n y)k(y)dy - f^{(l)}(x)$$

$$= \int \left[f^{(l)}(x) + \cdots + \frac{f^{(s-1)}(x)(-b_n y)^{s-l-1}}{(s-l-1)!} + \frac{f^{(s)}(x-\xi_1 b_n y)(-b_n y)^{s-l}}{(s-l)!} \right]$$

$$\times k(y)dy$$

$$-f^{(l)}(x)$$

$$= \int \frac{f^{(s)}(x-\xi_1 b_n y)(-b_n y)^{s-l}}{(s-l)!} k(y)dy, \quad 0 < \xi_1 < 1. \tag{3.8}$$

Hence, (a) is held under the condition that $E(\theta^{-(s+1)}) < \infty$.

(b) is obvious.

(c) For $l = 0, 1$, by the Theorem 1 of [1], we know

$$|k_{nl}(x)|^2 \le \frac{1}{(2\pi)^2} \left(\int |\phi_k(t)t^l| \exp(\sigma^2 t^2/(2b_n^2))dt \right)^2 = O(\exp(\sigma^2 b_n^{-2})), \tag{3.9}$$

where we let $\beta = 2$ and $\beta_0 = 0$ which appeared in [1]. Thus, $|V(x, Y_j, b_n)| = b_n^{-1} O(\exp(\sigma^2 b_n^{-2}/2)) + \theta_0 b_n^{-2} O(\exp(\sigma^2 b_n^{-2}/2))$. (c) is proved.

(d) Note that

$$Var\left(k_{nl}\left(\frac{x - Y_j}{b_n}\right)\right) = n b_n^{2l+2} Var(\hat{f}_n(x)) = O(\exp(\sigma^2 b_n^{-2})), \tag{3.10}$$

we know that (d) is true.

Using LEMMA 2, by mimicking the steps in Section 2, we have

$$0 \le R(\hat{\delta}_n(x), G(\theta)) - R(\delta_G(x), G(\theta)) = \frac{1}{n b_n^4} O(\exp(\sigma^2 b_n^{-2})) + O(b_n^{s-1}). \tag{3.11}$$

Taking $b_n = \sqrt{2\sigma^2}(\log n)^{-1/2}$, we obtain $(n b_n^4)^{-1} O(\exp(\sigma^2 b_n^{-2})) = o(n^{-\tau})$, where $\tau > 0$ can be arbitrarily close to $1/2$. Therefore, we show that the EB test $\hat{\delta}_n(x)$ (3.5) is asymptotically optimal under the conditions that $G(\theta) \in \mathcal{F}$ and $E(\theta^{-(s+1)}) < \infty$.

Remark 4. As described in [1], it is extremely difficult to solve deconvolution problems when the error distributions are normal and Cauchy (called supersmooth distributions). Actually, if we consider the error as Gamma distribution, which belongs to ordinary smooth, then we can obtain a higher rate of convergence employing the idea of [1] which used by [10, 11]. However, we should put more attention to the normal error due to its importance.

Acknowledgement. The author is indebted to an anonymous referee for his valuable suggestions.

References

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

[2] M. V. Jr. Johns and J. R. Van Ryzin, *Convergence rates for empirical Bayes two-action problem I: Discrete case*, Ann. Math. Statist 42 (1971), 1521-1539.

[3] M. V. Jr. Johns and J. R. Van Ryzin, *Convergence rates for empirical Bayes two-action problem II: Continuous case*, Ann. Math. Statist 43 (1972), 934-937.

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

[5] T. Liang, *On empirical Bayes tests in a positive exponential family*, J. Statist. Plann. Inference 83 (2000), 169-181.

[6] H. Robbins, *An empirical Bayes approach to statistics*. Proc, Third Berkeley Symp. Math. Statist. Prob. University of California Press 1 (1955), 157-163.

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

[8] TH. Stijnen, *On the asymptotic behaviour of empirical Bayes tests for the continuous one-parameter exponential family*, Ann. Statist 13 (1985), 403-412.

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

[10] S. P. Zhang and R. J. Karunamuni, *Empirical Bayes estimation for the continuous one-parameter exponential family with error in variables*, Statistics & Decision 15 (1997), 261-279.

[11] S. P. Zhang and R. J. Karunamuni, *Bayes and empirical Bayes estimation with errors in variables*, Statistics & Probability Letters 33 (1997), 23-34.

Limburgs Universitair Centrum
Center for Statistics
B-3590, Diepenbeek
Belgium
E-mail: lichun.wang@luc.ac.be