Empirical Bayes Test for the Parameter of Rayleigh Distribution with Error of Measurement*

HUANG JUAN

(School of Science, Guangdong Ocean University, Zhanjiang, Guangdong, 524088)

Communicated by Wang De-hui

Abstract: For the data with error of measurement in historical samples, the empirical Bayes test rule for the parameter of Rayleigh distribution is constructed, and the asymptotically optimal property is obtained. It is shown that the convergence rate of the proposed EB test rule can be arbitrarily close to $O(n^{-\frac{1}{2}})$ under suitable conditions.

Key words: error of measurement, empirical Bayes, asymptotic optimality, convergence rate

2000 MR subject classification: 62C12, 62F12 Document code: A Article ID: 1674-5647(2011)01-0???-07

Introduction 1

Empirical Bayes (EB) approach has been studied extensively by the researchers, and the readers are referred to literature [1]-[8].

Data with error of measurement take place in many fields, including biology, ecology, geology and medicine (see [9]-[10]). Up to now, empirical Bayes test problem for the parameter of distribution with error of measurement has not been studied by any researcher. Rayleigh distribution plays an important role in reliability analysis. In this paper, we discuss the empirical Bayes test for the parameter of Rayleigh distribution with error data of measurement.

Let X have a conditional density function

$$f(x \mid \theta) = \frac{x}{\theta^2} e^{-\frac{x^2}{2\theta^2}},$$
(1.1)

where θ is an unknown parameter. Denote the sample space by $x \in \Omega = \{x \mid x > 0\}$ and

^{*}Received date: July 29, 2009. Foundation item: The NSF (1012138, 0612163) of Guangdong Ocean Unversity and the Scientific and Technological Project (2010C3112006) of Zhanjiang.

parameter space by $\Theta = \{\theta \mid \theta > 0\}$. In this paper, we discuss the one-sided test problem $H_0: \theta \le \theta_0 \iff H_1: \theta > \theta_0,$ (1.2)

where θ_0 is a given positive constant.

To construct EB test function, we have firstly loss functions

$$L_{0}(\theta, d_{0}) = \begin{cases} 0, & \theta \leq \theta_{0}; \\ a \left[1 - \left(\frac{\theta_{0}}{\theta}\right)^{2} \right], & \theta > \theta_{0}, \end{cases} \qquad L_{1}(\theta, d_{1}) = \begin{cases} a \left[\left(\frac{\theta_{0}}{\theta}\right)^{2} - 1 \right], & \theta \leq \theta_{0}; \\ 0, & \theta > \theta_{0}, \end{cases}$$
ore $a > 0, d = \{d_{0}, d_{0}\}$ is action space, d_{0} and d_{0} imply accordance and rejection of

where a > 0, $d = \{d_0, d_1\}$ is action space, d_0 and d_1 imply acceptance and rejection of H_0 . Assume that the prior distribution $G(\theta)$ of θ is unknown. Then we have the randomized

decision function

$$\delta(x) = P(\text{accept } \mathbf{H}_0 \mid X = x). \tag{1.3}$$

And the risk function of $\delta(x)$ is shown by

$$R(\delta(x), G(\theta)) = \int_{\Theta} \int_{\Omega} [L_0(\theta, d_0) f(x \mid \theta) \delta(x) + L_1(\theta, d_1) f(x \mid \theta) (1 - \delta(x))] dx dG(\theta)$$

= $a \int_{\Omega} \beta(x) \delta(x) dx + C_G,$ (1.4)

where

$$C_G = \int_{\Theta} L_1(\theta, d_1) \mathrm{d}G(\theta), \qquad \beta(x) = \int_{\Theta} \left[1 - \left(\frac{\theta_0}{\theta}\right)^2 \right] f(x \mid \theta) \mathrm{d}G(\theta). \tag{1.5}$$

The marginal density function of X is given by

$$f_G(x) = \int_{\Theta} f(x \mid \theta) \mathrm{d}G(\theta) = \int_{\Theta} \frac{x}{\theta^2} \mathrm{e}^{-\frac{x^2}{2\theta^2}} \mathrm{d}G(\theta).$$

By (1.5) and simple calculations, we have

$$\beta(x) = u(x)f_G(x) + v(x)f_G^{(1)}(x), \qquad (1.6)$$

where $f_G^{(1)}(x)$ is the first order derivative of $f_G(x)$, and

$$u(x) = 1 - \frac{1}{4}\theta_0 x^{-2}, \qquad v(x) = \frac{1}{2}\theta_0 x^{-1}.$$

Using (1.4), we obtain the Bayes test function as follows:

$$\delta_G(x) = \begin{cases} 1, & \beta(x) \le 0; \\ 0, & \beta(x) > 0. \end{cases}$$
(1.7)

Further, we can get the minimum Bayes risk

$$R(G) = \inf_{\delta} R(\delta, G) = R(\delta_G, G) = a \int_{\Omega} \beta(x) \delta_G(x) dx + C_G.$$
(1.8)

When the prior distribution of $G(\theta)$ is known and $\delta(x) = \delta_G(x)$, R(G) can be obtained. However, when $G(\theta)$ is unknown, so that $\delta_G(x)$ cannot be made use of, we need to introduce EB method.

2 Construction of EB Test Function

Under the following assumptions, we are to construct the EB test function. Let (X_1, θ_1) , $(X_2, \theta_2), \dots, (X_n, \theta_n)$ and $(X_{n+1}, \theta_{n+1}) \cong (X, \theta)$ be independent random vectors, where θ_i $(i = 1, \dots, n)$ and θ are independent value distributed (i.i.d.) and have common prior distribution $G(\theta)$. Let X_1, X_2, \dots, X_n , X be sequence of mutually independent random

variables, where X_1, X_2, \dots, X_n are historical samples and X is the present sample. Due to some factors, historical samples X_1, X_2, \dots, X_n cannot be observed, which are sufferred from interruption of random error variables $\varepsilon_1, \varepsilon_2, \dots, \varepsilon_n$. Hence, we can only observe the data with error of measurement X'_1, X'_2, \dots, X'_n , and X'_i is determined by

$$X_i = X'_i + \varepsilon_i, \qquad 1 \le i \le n_i$$

where ε_i $(i = 1, 2, \dots, n)$ are mutually independent random variables having the identical normal distribution $N(0, \sigma^2)$, and the variance σ^2 is known.

Assume that X'_1, X'_2, \dots, X'_n and $\varepsilon_1, \varepsilon_2, \dots, \varepsilon_n$ are mutually independent. Let X_1, X_2, \dots, X_n have the common marginal density function $f_G(x)$, and X'_1, X'_2, \dots, X'_n have the common marginal density function $f_{G'}(x)$. Assume $f_G(x) \in C_{s,\alpha}, x \in \mathbb{R}^1$, where

$$C_{s,\alpha} = \{g(x) \mid g(x) \text{ is the probability density function and has continuous}$$

s-th order derivative
$$g^{(s)}(x)$$
 with $|g^{(s)}(x)| \le \alpha, \ s \ge 2, \ \alpha > 0$.

First we construct the estimator of $\beta(x)$.

Let $K_r(x)$ be a Borel measurable bounded function vanishing off (0,1) and such that

(A1)
$$\frac{1}{t!} \int_0^1 y^t K_r(y) dy = \begin{cases} 1, & t = r; \\ 0, & t \neq r, & t = 0, 1, 2, \cdots, s - 1 \end{cases}$$

By the convolution formula, we get

$$f_G(x) = \int_{-\infty}^{+\infty} f_{G'}(x-y) \frac{1}{\sqrt{2\pi\sigma}} e^{-\frac{y^2}{2\sigma^2}} dy.$$
 (2.1)

The kernel estimation of $f_G^{(r)}(x)$ is defined by

$$f_{G}^{(r)}(x) = \int_{-\infty}^{+\infty} f_{G'}^{(r)}(x-y) \frac{1}{\sqrt{2\pi\sigma}} e^{-\frac{y^2}{2\sigma^2}} dy.$$
 (2.2)

Denote

$$f_G^{(0)}(x) = f_G(x).$$

Since $f_{G'}^{(r)}(x)$ is unknown, the kernel-type density estimation (see [5]) of $f_{G'}^{(r)}(x)$ is defined by

$$\widehat{f}_{G'}^{(r)}(x) = \frac{1}{nh_n^{(1+r)}} \sum_{i=1}^n K_r\left(\frac{X'_i - x}{h_n}\right).$$
(2.3)

Substituting (2.3) into (2.2), we get

$$\hat{f}_{G}^{(r)}(x) = \frac{1}{nh_{n}^{(1+r)}} \sum_{i=1}^{n} \int_{-\infty}^{+\infty} K_{r} \Big(\frac{X_{i}' - x + y}{h_{n}}\Big) \frac{1}{\sqrt{2\pi\sigma}} e^{-\frac{y^{2}}{2\sigma^{2}}} dy,$$
(2.4)

where $r = 0, 1, h_n$ is the smoothing bandwidth, $h_n > 0$, and $\lim_{n \to \infty} h_n = 0$. Write

$$f_{G'}^{(0)}(x) = f_{G'}(x)$$

Then, the estimator of $\beta(x)$ is obtained from

$$\beta_n(x) = u(x)\hat{f}_G(x) + v(x)\hat{f}_G^{(1)}(x).$$
(2.5)

Hence, the EB test function is defined by

$$\delta_n(x) = \begin{cases} 1, & \beta_n(x) \le 0; \\ 0, & \beta_n(x) > 0. \end{cases}$$
(2.6)

Let E_n denote the mathematical expectation with respect to the joint distribution of X'_1, X'_2, \dots, X'_n . We get the overall Bayes risk of $\delta_n(x)$ as

$$R(\delta_n, G) = a \int_{\Omega} \beta(x) E_n[\delta_n(x)] dx + C_G.$$
(2.7)

If

 $\lim_{n \to \infty} R(\delta_n, G) = R(\delta_G, G),$

then $\{\delta_n(x)\}$ is the asymptotic optimality of EB test function; if

$$R(\delta_n, G) - R(\delta_G, G) = O(n^{-q}), \qquad q > 0,$$

then $O(n^{-q})$ is the asymptotic optimality convergence rate of EB test function of $\{\delta_n(x)\}$.

We give two lemmas in the following.

Let c, c_1, c_2, c_3, c_4 be constants which can be different in different cases even in the same expression.

Lemma 2.1 Let $\hat{f}_{G}^{(r)}(x)$ be defined by (2.4). Assume that (A1) holds, and $x \in \Omega$. (I) If $f_{G}^{(r)}(x)$ is a continuous function,

$$\lim_{n \to \infty} h_n = 0, \qquad \lim_{n \to \infty} n h_n^{2r+2} = \infty,$$

then

$$\lim_{n \to \infty} E|\widehat{f}_G^{(r)}(x) - f_G^{(r)}(x)|^2 = 0$$

(II) If $f_G(x) \in C_{s,a}$, taking $h_n = n^{-\frac{1}{2+2s}}$, then, for $0 < \lambda \le 1$, we have $E|\widehat{f}_G^{(r)}(x) - f_G^{(r)}(x)|^{2\lambda} \le c \cdot n^{-\frac{\lambda(s-r)}{1+s}}.$

Proof. (I) By C_r inequality, we get

$$E|\widehat{f}_{G}^{(r)}(x) - f_{G}^{(r)}(x)|^{2} \le 2|E\widehat{f}_{G}^{(r)}(x) - f_{G}^{(r)}(x)|^{2} + 2\operatorname{Var}(\widehat{f}_{G}^{(r)}(x)) := 2(I_{1}^{2} + I_{2}).$$
(2.8)

Then

$$E\widehat{f}_{G}^{(r)}(x) = \frac{1}{h_{n}^{(1+r)}} \int_{-\infty}^{+\infty} \left\{ \left[\int_{-\infty}^{+\infty} K_{r} \left(\frac{s - x + y}{h_{n}} \right) f_{G'}(s) \mathrm{d}s \right] \frac{1}{\sqrt{2\pi\sigma}} \mathrm{e}^{-\frac{y^{2}}{2\sigma^{2}}} \right\} \mathrm{d}y$$
$$= \frac{1}{h_{n}^{r}} \left\{ \int_{-\infty}^{+\infty} \left[\int_{-\infty}^{+\infty} K_{r}(u) f_{G'}(x - y + h_{n}u) \mathrm{d}u \right] \frac{1}{\sqrt{2\pi\sigma}} \mathrm{e}^{-\frac{y^{2}}{2\sigma^{2}}} \right\} \mathrm{d}y$$
$$= \frac{1}{h_{n}^{r}} \int_{-\infty}^{+\infty} I_{1}(x, y) \frac{1}{\sqrt{2\pi\sigma}} \mathrm{e}^{-\frac{y^{2}}{2\sigma^{2}}} \mathrm{d}y,$$

where

$$I_1(x,y) = \int_{-\infty}^{+\infty} K_r(u) f_{G'}(x-y+h_n u) \mathrm{d}u.$$

Since $f_G(x) \in C_{s,\alpha}$, it is easy to see that $f_{G'}(x) \in C_{s,\alpha}$. We obtain the following Taylor expansion of $f_{G'}(x-y+h_n u)$ in x-y:

$$f_{G'}(x-y+h_nu) - f_{G'}(x-y)$$

= $\frac{f'_{G'}(x-y)}{1!}h_nu + \frac{f''_{G'}(x-y)}{2!}(h_nu)^2 + \dots + \frac{f^{(s)}_{G'}(\xi^*)}{s!}(h_nu)^s,$

where $\xi^* \in (x - y, x + y + h_n u)$.

Due to (A1) and $f_G(x) \in C_{s,\alpha}$, it is easy to see that

$$I_1(x,y) = f_{G'}^{(r)}(x-y) + o(h_n^{s-r}).$$

$$E\widehat{f}_{G}^{(r)}(x) = \int_{-\infty}^{+\infty} f_{G'}^{(r)}(x-y) \frac{1}{\sqrt{2\pi\sigma}} e^{-\frac{y^2}{2\sigma^2}} dy + o(h_n^{s-r})$$

= $f_{G}^{(r)}(x) + o(h_n^{s-r}).$

Furthermore, we hav

$$I_1 = |E\hat{f}_G^{(r)}(x) - f_G^{(r)}(x)|^2 \le ch_n^{2(s-r)}.$$
(2.9)

When $h_n \to 0$, we get

$$\lim_{n \to \infty} I_1^2 = \lim_{n \to \infty} |E\hat{f}_G^{(r)}(x) - f_G^{(r)}(x)|^2 = 0.$$
(2.10)

It follows that

$$I_{2} = n^{-1}h_{n}^{-2(r+2)}\operatorname{Var}\left[\int_{-\infty}^{+\infty} K_{r}\left(\frac{X_{i}'-x+y}{h_{n}}\right)\frac{1}{\sqrt{2\pi\sigma}}\mathrm{e}^{-\frac{y^{2}}{2\sigma^{2}}}\mathrm{d}y\right]$$
$$\leq n^{-1}h_{n}^{-2(r+2)}E\left\{\int_{-\infty}^{+\infty}\left[K_{r}\left(\frac{X_{1}'-x+y}{h_{n}}\right)\right]\frac{1}{\sqrt{2\pi\sigma}}\mathrm{e}^{-\frac{y^{2}}{2\sigma^{2}}}\mathrm{d}y\right\}^{2}.$$
s a Borel measurable bounded function, we get

Since $K_r(x)$ is

$$\int_{-\infty}^{+\infty} \left| K_r \left(\frac{X_1' - x + y}{h_n} \right) \right| \frac{1}{\sqrt{2\pi\sigma}} \mathrm{e}^{-\frac{y^2}{2\sigma^2}} \mathrm{d}y < c.$$

Hence

$$I_2 \le cn^{-1}h_n^{-(2r+2)}.$$
(2.11)

When $h_n \to 0$, $nh_n^{2r+2} \to \infty$, we get

$$\lim_{n \to \infty} I_2 = \lim_{n \to \infty} \operatorname{Var}(\widehat{f}_G^{(r)}(x)) = 0.$$
(2.12)

By substituting (2.10), (2.12) into (2.8), the proof of (I) is completed.

(II) Similar to (2.8), we can show that

$$E|\hat{f}_{G}^{(r)}(x) - f_{G}^{(r)}(x)|^{2\lambda} \le 2[E\hat{f}_{G}^{(r)}(x) - f_{G}^{(r)}(x)]^{2\lambda} + 2[\operatorname{Var}(\hat{f}_{G}^{(r)}(x))]^{\lambda}$$

$$:= 2(J_{1}^{2\lambda} + J_{2}^{\lambda}).$$
(2.13)

By (2.9), when
$$h_n = n^{-\frac{1}{2+2s}}$$
, we get
 $J_1^{2\lambda} = |E\hat{f}_G^{(r)}(x)|$

$$J_1^{2\lambda} = |E\hat{f}_G^{(r)}(x) - f_G^{(r)}(x)|^{2\lambda} \le c \cdot n^{-\frac{\lambda(s-r)}{s+1}}.$$
(2.14)
By (2.11), when $h_n = n^{-\frac{1}{2+2s}}$, we have

$$J_2^{\lambda} \le [c_1(nh_n^{2r+2})^{-1}]^{\lambda} \le c \cdot n^{-\frac{\lambda(s-r)}{1+s}}.$$
(2.15)

By substituting (2.14), (2.15) into (2.13), the proof of (II) is completed.

Lemma 2.2^[2] Let
$$R(\delta_G, G)$$
 and $R(\delta_n, G)$ be defined by (1.8) and (2.7). Then
 $0 \le R(\delta_n, G) - R(\delta_G, G) \le a \int_{\Omega} |\beta(x)| P(|\beta_n(x) - \beta(x)| \ge |\beta(x)|) dx.$

Asymptotic Optimality and Convergence Rates 3

Let $\widehat{f}_{G}^{(r)}(x)$ be defined by (2.4). Assume that (A1) and the following regu-Theorem 3.1 larity conditions hold:

(i) $h_n > 0$, $\lim_{n \to \infty} h_n = 0$, $\lim_{n \to \infty} nh_n^4 = \infty$; (ii) $\int_{\Theta} \theta^{-2} dG(\theta) < +\infty$;

(iii) $f_G^{(1)}(x)$ is a continuous function. Then we have

$$\lim_{n \to \infty} R(\delta_n, G) = R(\delta_G, G).$$

Proof. By Lemma 2.2 we have

$$0 \le R(\delta_n, G) - R(\delta_G, G) \le a \int_{\Omega} |\beta(x)| P(|\beta_n(x) - \beta(x)| \ge |\beta(x)|) \mathrm{d}x.$$

Write

$$Q_n(x) = |\beta(x)| P(|\beta_n(x) - \beta(x)| \ge |\beta(x)|).$$

Then

$$Q_n(x) \le |\beta(x)|.$$

Again by (1.6) and the Fubini theorem, we can get

$$\int_{\Omega} |\beta(x)| dx = \int_{\Omega} \int_{\Theta} \left[1 - \left(\frac{\theta_0}{\theta}\right)^2 \right] f(x \mid \theta) dG(\theta) dx$$
$$= 1 + \theta_0^2 \int_{\Omega} \int_{\Theta} \theta^{-2} f(x \mid \theta) dG(\theta) dx$$
$$= 1 + \theta_0^2 \int_{\Theta} \theta^{-2} dG(\theta)$$
$$< +\infty.$$

Applying the dominant convergence theorem, we have

$$0 \le \lim_{n \to \infty} R(\delta_n, G) - R(\delta_G, G) \le \int_{\Omega} [\lim_{n \to \infty} Q_n(x)] \mathrm{d}x.$$
(3.1)

To prove that Theorem 3.1 holds, we only need to prove

$$Q_n(x) = 0 \qquad a.s.x.$$

 $\lim_{n \to \infty} Q_n(x) = 0$ By Markov's and Jensen's inequalities, one has $Q_n(x)) \le E|\beta_n(x) - \beta(x)|$

$$\leq |u(x)|E|\widehat{f}_G(x) - f_G(x)| + |v(x)|E|\widehat{f}_G^{(1)}(x) - f_G^{(1)}(x)|$$

$$\leq |u(x)|[E|\widehat{f}_G(x) - f_G(x)|^2]^{1/2} + |v(x)|[E|\widehat{f}_G^{(1)}(x) - f_G^{(1)}(x)|^2]^{1/2}.$$

Again by Lemma 2.1(I), for fixed $x \in \Omega$, when r = 0, 1, we have

$$0 \leq \lim_{n \to \infty} Q_n(x)$$

$$\leq |u(x)| [\lim_{n \to \infty} E |\widehat{f}_G(x) - f_G(x)|^2]^{1/2} + |v(x)| [\lim_{n \to \infty} E |\widehat{f}_G^{(1)}(x) - f_G^{(1)}(x)|^2]^{1/2}$$

$$= 0.$$
(3.2)

By substituting (3.2) into (3.1), the proof of Theorem 3.1 is completed.

Let $\widehat{f}_{G}^{(r)}(x)$ be defined by (2.4). Assume (A1) and the following regularity Theorem 3.2 conditions hold:

(B1)
$$f_G(x) \in C_{s,\alpha};$$

(B2) $\int_{\Omega} |x|^{-m\lambda} |\beta(x)|^{1-\lambda} dx < +\infty, \text{ where } 0 < \lambda \leq 1, m = 0, 1, 2.$
Then, if $h_n = n^{-\frac{1}{2s+2}}$, we have
 $R(\delta_n, G) - R(\delta_G, G) = O\left(n^{-\frac{\lambda(s-1)}{2(s+1)}}\right),$
where $s \geq 2$

where $s \geq 2$.

Proof. By Lemma 2.2 and Markov's inequality, we have

$$0 \leq R(\delta_{n}, G) - R(\delta_{G}, G)$$

$$\leq \int_{\Omega} |\beta(x)|^{1-\lambda} E|\beta_{n}(x) - \beta(x)|^{\lambda} dx$$

$$\leq c_{1} \int_{\Omega} |\beta(x)|^{1-\lambda} |u(x)| E|\widehat{f}_{G}(x) - f_{G}(x)| dx$$

$$+ c_{2} \int_{\Omega} |\beta(x)|^{1-\lambda} |v(x)| E|\widehat{f}_{G}^{(1)}(x) - f_{G}^{(1)}(x)| dx$$

$$= A_{n} + B_{n}.$$
(3.3)

By Lemma 2.1(II) and condition (B2), we get

$$A_n \le c_1 n^{-\frac{\lambda s}{2s+2}} \int_{\Omega} |\beta(x)|^{1-\lambda} |u(x)|^{\lambda} \mathrm{d}x \le c_3 n^{-\frac{\lambda s}{2s+2}}, \tag{3.4}$$

$$B_n \le c_2 n^{-\frac{\lambda(s-1)}{2s+2}} \int_{\Omega} |\beta(x)|^{1-\lambda} |v(x)|^{\lambda}(x) \mathrm{d}x \le c_4 n^{-\frac{\lambda(s-1)}{2s+2}}.$$
(3.5)

Substituting (3.4) and (3.5) into (3.3), we get

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

The proof of Theorem 3.2 is completed.

Remark 3.1 When $\lambda \to 1$ and $s \to \infty$, $O\left(n^{-\frac{\lambda(s-1)}{2(s+1)}}\right)$ is arbitrarily close to $O(n^{-\frac{1}{2}})$.

References

- Johns, M. VJr. and Van Ryzin, J., Convergence rates in empirical Bayes two-action problems 1: Discrete case, Ann. Math. Statist., 42(1971), 1521–1539.
- Johns, M. VJr. and Van Ryzin, J., Convergence rates in empirical Bayes two-action problems 2: Continuous case, Ann. Math. Statist., 43(1972), 934–947.
- [3] Liang, Tachen, Empirical Bayes testing for double exponential distributions, Comm. Statist.-Theory Methods, 36(2007), 1543–1553.
- [4] Chen, L. S., On empirical Bayes two-tail tests for double exponential distributions, J. Nonparametr. Statist., 21(2009), 1037–1049.
- [5] Dunson, D. B., Empirical Bayes density regression, Statist. Sinica, 17(2007), 481–504.
- [6] Liang, T., Empirical Bayes testing for a normal mean: variance unknown case, J. Statist. Planning Inference, 136(2006), 1376–1393.
- [7] Chen, Lee, Shen., Empirical Bayes testing for a nonexponential family distribution, Comm. Statist.-Theory Methods, 36(2007), 2061–2074.
- [8] Karunamuni, R. J., Liang, T. and Wu, J., Robust empirical Bayes tests for discrete distributions, J. Nonparametr. Statist., 20(2008), 101–113.
- [9] Davis, D. J., An analysis of some failure data, J. Amer. Statist. Assoc., 47(1952), 113–150.
- [10] Fuller, W. A., Measurement Error Model, Wiley, New York, 1987.