Communications of the Korean Statistical Society Vol. 15, No. 2, 2008, pp. 213–228

Empirical Bayes Test for the Exponential Parameter with Censored Data†

Lichun Wang[1)]

Abstract

Using a linear loss function, this paper considers the one-sided testing problem for the exponential distribution via the empirical Bayes(EB) approach. Based on right censored data, we propose an EB test for the exponential parameter and obtain its convergence rate and asymptotic optimality, firstly, under the condition that the censoring distribution is known and secondly, that it is unknown.

Keywords: Asymptotic optimality; convergence rate; empirical Bayes; random censorship.

1. Introduction

In reliability and life testing studies, the exponential distribution plays an important role. It was the first lifetime model for which inference procedures are extensively developed. The probability density function (*pdf*) of an exponential distribution is

$$f(x|\lambda) = \lambda \exp(-\lambda x) I(x > 0),$$

where $\lambda > 0$ is the exponential parameter and $I(A)$ denotes the indicator function of the set A. To employ the empirical Bayes approach, we assume that the parameter λ has an unknown non-degenerate prior $G(\lambda)$ with support on $\Lambda = (0, \infty)$.

Consider the following problem of testing the hypotheses

$$H_0 : 0 < \lambda \leq \lambda_0 \longleftrightarrow H_1 : \lambda > \lambda_0,$$

under a linear loss function, defined as follows

$$L(\lambda, d_i) = (1 - i)(\lambda - \lambda_0) I(\lambda > \lambda_0) + i(\lambda_0 - \lambda) I(\lambda_0 \geq \lambda),$$

where $\mathcal{D} = \{d_0, d_1\}$ denotes the action space with d_i $(i = 0, 1)$ accepting H_i.

In this paper we assume that the true sample $X = x$ is not observable. Instead, we only observe $T = \min\{X, Y\}$ and $\Delta = I(X \leq Y)$, where Y denotes the censoring

† Supported by NNSF of China (No. 10771012).

1) Associate Professor, Department of Mathematics, Beijing Jiaotong University, Beijing 100044, China. Correspondence: wlc@amss.ac.cn

variable, which is nonnegative and independent of X and follows an absolutely continuous distribution W.

Obviously, given λ, T has the conditional *pdf*,

$$h(t|\lambda) = f(t|\lambda)\{1 - W(t)\} + w(t)\{1 - F(t|\lambda)\}, \quad t > 0,$$

where $w(t)$ denotes the *pdf* of Y and $F(t|\lambda) = \int_{-\infty}^{t} f(x|\lambda)dx$.

Set $p(t) = P\{\text{accepting } H_0|(T, \Delta) = (t, \delta)\}$. Let $R(p(t), G(\lambda))$ be the Bayes risk of the test $p(t)$ with respect to the prior $G(\lambda)$. Then, by Fubini's theorem and the fact that $\int_0^\infty h(t|\lambda)dt = 1$, we have

$$\begin{aligned} R(p(t), G(\lambda)) &= \int_0^\infty \int_\Lambda [L(\lambda, d_0)p(t) + L(\lambda, d_1)\{1 - p(t)\}]h(t|\lambda)dG(\lambda)dt \\ &= \int_0^\infty m(t)p(t)dt + \int_\Lambda L(\lambda, d_1)dG(\lambda), \end{aligned}$$

where

$$m(t) = \int_\Lambda (\lambda - \lambda_0)h(t|\lambda)dG(\lambda).$$

Since $\int_\Lambda L(\lambda, d_1)dG(\lambda)$ does not depend on t, it is easy to see that the Bayes test $p_G(t)$, which minimizes the risk $R(p(t), G(\lambda))$, would have form

$$p_G(t) = \begin{cases} 1, & \text{if } m(t) \le 0, \\ 0, & \text{if } m(t) > 0. \end{cases} \tag{1.1}$$

Its Bayes risk is

$$R(p_G(t), G(\lambda)) = \int_0^\infty m(t)p_G(t)dt + \int_\Lambda L(\lambda, d_1)dG(\lambda). \tag{1.2}$$

Furthermore, note that $m(t) \le 0 \iff m(t)/h(t) \le 0 \iff E(\lambda|T = t) \le \lambda_0$ with probability one, we have

$$p_G(t) = \begin{cases} 1, & \text{if } E(\lambda|T = t) \le \lambda_0, \\ 0, & \text{if } E(\lambda|T = t) > \lambda_0, \end{cases} \tag{1.3}$$

where $h(t) = \int_\Lambda h(t|\lambda)dG(\lambda) = f(t)\{1 - W(t)\}I(t > 0) + w(t)\{1 - F(t)\}I(t > 0)$ is the marginal *pdf* of T, $f(t) = \int_\Lambda f(t|\lambda)dG(\lambda)$ and $F(t)$ denote the marginal *pdf* and distribution of X, respectively and $E(\lambda|T = t) = \int_\Lambda \lambda h(t|\lambda)dG(\lambda)/h(t)$ is the posterior mean of λ given $T = t$.

From (1.3), we see that the Bayes test $p_G(t)$ is uniquely determined by the posterior mean of λ, given T. If $E(\lambda|T = t) \le \lambda_0$, we accept H_0; otherwise, we accept H_1.

However, since the prior $G(\theta)$ is unknown to us, the Bayes test $p_G(t)$ is unavailable to use. As an alternative we can employ the EB approach, which was introduced to statistical problems by Robbins (1955, 1964), to estimate $E(\lambda|T = t)$ so as to obtain an EB test.

Since Robbins' pioneering work, the EB approach has been applied in a wide range of paradigms and to numerous real-life problems. Some earlier results dealing with the EB estimation or testing problem in one-parameter exponential family can be found in Johns and Van Ryzin (1971, 1972), Van Houwelingen (1976), Singh (1979) and Stijnen (1985), *etc.* Recently, Liang (2000a) employed Bernstein's inequality and obtained a better order of convergence rate than that of Karunamuni and Yang (1995). Under the assumption that the critical point is within an unknown compact interval, Liang (2000a) claimed that the order of convergence rate is $O(n^{-s/(s+3)})$ for a positive exponential family. Similar work is also presented in Liang (2000b).

Differing from the above works, this paper considers the case that the true sample is censored. We assume that the same case had taken place n times in the past historical experiments. In the random censorship model, we observe only $T_i = \min\{X_i, Y_i\}$ and $\delta_i = I(X_i \leq Y_i)$, where $i = 1, 2, \ldots, n$. Hence, at the present stage, (T, Δ) are the present sample while $(T_1, \delta_1), \ldots, (T_n, \delta_n)$ denote the n past data. Actually, many statistical experiments result in incomplete samples, even under well-controlled situations. This is because individuals will experience some other competing events which cause them to be removed.

Based on (T_i, δ_i) $(1 \leq i \leq n)$ and (T, Δ), in Section 2 we construct a monotone EB test and obtain its convergence rate under the condition that the censoring distribution W is known. In Section 3, for the case that the censoring distribution W is unknown to us, an asymptotically optimal EB test is proposed. Finally, some remarks and conclusions are given in Section 4.

2. EB Test for the Case that W is Known

Note that

$$
\begin{aligned}
E(\lambda|T=t) &= \frac{\int_\Lambda \lambda h(t|\lambda)dG(\lambda)}{h(t)} \\
&= \frac{\{1-W(t)\}\int_\Lambda \lambda f(t|\lambda)dG(\lambda) + w(t)\int_\Lambda \lambda\{1-F(t|\lambda)\}dG(\lambda)}{f(t)\{1-W(t)\}+w(t)\{1-F(t)\}} \\
&= \frac{-f^{(1)}(t)\{1-W(t)\}+w(t)f(t)}{f(t)\{1-W(t)\}+w(t)\{1-F(t)\}},
\end{aligned}
\tag{2.1}
$$

where

$$\int_\Lambda \lambda f(t|\lambda)dG(\lambda) = \int_\Lambda \lambda^2 \exp(-\lambda t)I(t>0)dG(\lambda) = -f^{(1)}(t)$$

and

$$\int_\Lambda \lambda\{1-F(t|\lambda)\}dG(\lambda) = \int_\Lambda \lambda \exp(-\lambda t)I(t>0)dG(\lambda) = f(t).$$

Since $W(t)$ and its density $w(t)$ are known in this section, we only need to estimate $F(t)$, $f(t)$ and its derivative $f^{(1)}(t)$ in (2.1).

An estimator $\hat{F}(t)$ of $F(t)$ can be defined by (see, Kaplan and Meier, 1958)

$$1-\hat{F}(t)=\prod_{i=1}^{n}\left(\frac{n-i}{n-i+1}\right)^{I(T_{(i)}\leq t,\ \delta_{(i)}=1)},\quad t<T_{(n)}, \tag{2.2}$$

where $T_{(1)} \leq \cdots \leq T_{(n)}$ denote the ordered samples and $\delta_{(i)}$ is the concomitant of $T_{(i)}$.

Consequently, we can define the estimator for $f^{(i)}(t)$ $(i = 0, 1)$ as follows

$$\hat{f}^{(i)}(t)=\frac{1}{h_n^{1+i}}\int k_i\left(\frac{t-y}{h_n}\right)d\hat{F}(y), \tag{2.3}$$

where $0 < h_n \to 0$ (as $n \to \infty$) denotes the bandwidth and $k_i(x)$ $(i = 0, 1)$ are kernel functions. Denote $e(t) = E(\lambda|T = t)$. we define the following estimator for $e(t)$,

$$\hat{e}(t)=\left[\frac{-\hat{f}^{(1)}(t)\{1-W(t)\}+w(t)\hat{f}(t)}{\hat{f}(t)\{1-W(t)\}+w(t)\{1-\hat{F}(t)\}}\right]_{n^v}, \tag{2.4}$$

where $0 < v < 1$ is to be determined and $[b]_M = \begin{cases} b, & \text{if } |b| \leq M, \\ 0, & \text{if } |b| > M. \end{cases}$

Hence, an EB test is defined as

$$p_{EB}(t)=\begin{cases}1, & \text{if } \hat{e}(t)\leq\lambda_0,\\ 0, & \text{if } \hat{e}(t)>\lambda_0.\end{cases} \tag{2.5}$$

The overall Bayes risk of $p_{EB}(t)$ is

$$R(p_{EB}(t),G(\lambda))=\int_0^\infty m(t)E_n\{p_{EB}(t)\}dt+\int_\Lambda L(\lambda,d_1)dG(\lambda), \tag{2.6}$$

where E_n denotes the expectation with respect to the joint distribution of $(T_1, \ldots, T_n)$.

By definition, the EB test $p_{EB}(t)$ is said to be asymptotically optimal relative to the prior $G(\lambda)$ if $R(p_{EB}(t), G(\lambda)) - R(p_G(t), G(\lambda)) = o(1)$, where $o(1)$ denotes terms converging to 0 as $n \to \infty$. If for some $q > 0$, $R(p_{EB}(t), G(\lambda)) - R(p_G(t), G(\lambda)) = O(n^{-q}) = n^{-q}O(1)$ with $O(1)$ denoting terms bounded, then the convergence rate of $p_{EB}(t)$ is said to be the order $O(n^{-q})$.

Following from (1.2) and (2.6) and using Markov's inequality,

$$\begin{aligned}0 &\leq R(p_{EB}(t),G(\lambda))-R(p_G(t),G(\lambda))\\ &=\int_0^\infty m(t)\left[E_n\{p_{EB}(t)\}-p_G(t)\right]dt\\ &=\begin{cases}\displaystyle\int_0^\infty h(t)\{e(t)-\lambda_0\}[P\{\hat{e}(t)\leq\lambda_0\}-1]dt, & \text{if } e(t)\leq\lambda_0,\\ \displaystyle\int_0^\infty h(t)\{e(t)-\lambda_0\}P\{\hat{e}(t)\leq\lambda_0\}dt, & \text{if } e(t)>\lambda_0\end{cases}\end{aligned}$$

$$\leq \int_0^\infty h(t)|e(t)-\lambda_0|P\{|\hat{e}(t)-e(t)|>|e(t)-\lambda_0|\}$$
$$\leq \int_0^\infty h(t)E_n|\hat{e}(t)-e(t)|dt = E_T E_n|\hat{e}(T)-e(T)|, \tag{2.7}$$

where E_T stands for the expectation with respect to the variable T.

The main result in this section can be formulated in the following theorem.

Theorem 2.1 Let $p_G(t)$ and $p_{EB}(t)$ be defined in (1.1) and (2.5), and $R(p_G(t), G(\lambda))$ and $R(p_{EB}(t), G(\lambda))$ be defined in (1.2) and (2.6), respectively. If the following conditions hold:

$$(C1)\ E[h^{-2}(T)\{1-W(T)\}^2\{f^{(s)}(T)\}^2] < \infty,$$
$$(C2)\ E[h^{-2}(T)\{1-W(T)\}f(T)] < \infty,$$
$$(C3)\ E[h^{-2}(T)w^2(T)\{f^{(s)}(T)\}^2] < \infty,$$
$$(C4)\ E[h^{-2}(T)w^2(T)\{1-W(T)\}^{-1}f(T)] < \infty,$$
$$(C5)\ E\{h^{-2}(T)w^2(T)\} < \infty,$$
$$(C6)\ E(\lambda^{2s}) < \infty,$$

then, taking $h_n = n^{-1/(2s+1)}$ and $v = (2s+1)^{-1}$,

$$\sup_{t\leq T_0}\left[R(p_{EB}(t), G(\lambda)) - R(p_G(t), G(\lambda))\right] = O\left(n^{-\frac{s-1}{2s+1}}\right),$$

where T_0 is such that $1-H(T_0) > \epsilon$ with some $\epsilon > 0$ and $s \geq 2$ is an integer.

To prove Theorem 2.1, we need the following several lemmas.

Lemma 2.1 Let Y, Y' be random variables, let y, y' and $M > 0$ be real numbers, then for $0 \leq r \leq 2$

$$E\left|\left[\frac{Y'}{Y} - \frac{y'}{y}\right]_M\right|^r \leq 2|y|^{-r}\left\{E|Y'-y'|^r + \left(\left|\frac{y'}{y}\right| + M\right)^r E|Y-y|^r\right\}.$$

Proof: See Singh (1977). □

Lemma 2.2 Let T_0 be such that $1-H(T_0) > \epsilon$ with some $\epsilon > 0$. Then the process $\hat{F}(t) - F(t)$, $-\infty < t < \infty$, $1-H(t) > 0$, can be represented as

$$\hat{F}(t) - F(t) = \frac{1}{n}\sum_{j=1}^n \{1-F(t)\}M_j(t) + \frac{1}{n}R_n(t)$$

in such a way that

$$P\left(\sup_{t\leq T_0}|R_n(t)| > \frac{2C}{\epsilon}\log^2 n + x\log n\right) \leq 2K\exp(-\tau\epsilon^2 x), \quad x > 0,$$

where *i.i.d.* Gaussian processes $M_1(t), M_2(t), \ldots, EM_n(t) = 0$ with the covariance function

$$EM(s)M(t) = EM(s)^2 = \int_{-\infty}^{s} \frac{dF(t)}{\{1-W(t)\}\{1-F(t)\}^2}, \quad -\infty < s \le t < \infty.$$

Here $C > 0$, $K > 0$ and $\tau > 0$ are some universal constants.

Proof: See Major and Rejtö (1988).

Before stating next lemma, we assume the kernel function $k_i(x)$ $(i = 0, 1)$ satisfy:

(A1) $k_i(x)$ $(i = 0, 1)$ are continuously differentiable with support (0,1) and

$$(1) \int_0^1 x^j k_i(x)dx = \begin{cases} (-1)^j, & \text{if } j = i, \\ 0, & \text{if } j \ne i, \quad j = 0, 1, \ldots, s-1, \end{cases}$$

$$(2) \int_0^1 x^s k_i(x)dx \ne 0, \quad k_i(0) = k_i(1) = 0, \quad i = 0, 1,$$

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

Lemma 2.3 Let $\hat{f}^{(i)}(t)$ be defined in (2.3). Under the assumption (A1), choosing $h_n = n^{-1/(2s+1)}$, for $t \le T_0$, which is defined in Lemma 2.2, we have

$$E_n\{\hat{f}^{(i)}(t) - f^{(i)}(t)\}^2 \le \left[c_{1i}\{f^{(s)}(t)\}^2 + c_{2i} f(t)\{1-W(t)\}^{-1}\right] n^{-(2s-2i)/(2s+1)}, \quad i = 0, 1,$$

where c_{1i} and c_{2i} are positive constants that do not depend on n and $s \ge 2$.

Proof: The original idea maybe come from Lemdani and Ould-Saïd (2002). Integrating by parts, we have

$$\begin{aligned}
&\hat{f}^{(i)}(t) - f^{(i)}(t) \\
&= \frac{1}{h_n^{1+i}} \int k_i\left(\frac{t-y}{h_n}\right) d\{\hat{F}(y) - F(y)\} + \frac{1}{h_n^{1+i}} \int k_i\left(\frac{t-y}{h_n}\right) dF(y) - f^{(i)}(t) \\
&= \frac{1}{h_n^{2+i}} \int \left\{\hat{F}(y) - F(y)\right\} k_i^{(1)}\left(\frac{t-y}{h_n}\right) dy + \left\{\frac{1}{h_n^i} \int k_i(u) f(t - uh_n)du - f^{(i)}(t)\right\} \\
&\hat{=} I_1 + I_2.
\end{aligned} \tag{2.8}$$

By Lemma 2.2, we have

$$\begin{aligned}
I_1 &= \frac{1}{nh_n^{1+i}} \int_0^1 \sum_{j=1}^{n} \{1 - F(t-uh_n)\} M_j(t-uh_n) k_i^{(1)}(u)du \\
&\quad + \frac{1}{nh_n^{1+i}} \int_0^1 R_n(t-uh_n) k_i^{(1)}(u)du \\
&\hat{=} I_{11} + I_{12}
\end{aligned} \tag{2.9}$$

and

$$I_{11}^2 = \frac{1}{n^2h_n^{2+2i}} \int_0^1 \int_0^1 \bar{F}(t-uh_n)\bar{F}(t-vh_n) \sum_{j=1}^n M_j(t-uh_n) \times \sum_{l=1}^n M_l(t-vh_n) k_i^{(1)}(u) k_i^{(1)}(v) du dv, \tag{2.10}$$

where $\bar{F}(t) = 1 - F(t)$.

Firstly, expanding $d(t-uh_n)$, $\bar{F}(t-uh_n)$ and $\bar{F}(t-uh_n)$ at point t, one obtains

$$\begin{aligned} E_n I_{11}^2 &= \frac{1}{nh_n^{2+2i}} \int_0^1 \int_0^1 E\{M_1(t-uh_n)M_1(t-vh_n)\} \\ &\quad \times \bar{F}(t-uh_n)\bar{F}(t-vh_n)k_i^{(1)}(u)k_i^{(1)}(v)dudv \\ &= \frac{1}{nh_n^{2+2i}} \int_0^1 \int_0^u d(t-uh_n)\bar{F}(t-uh_n)\bar{F}(t-vh_n)k_i^{(1)}(v)dvk_i^{(1)}(u)du \\ &\quad + \frac{1}{nh_n^{2+2i}} \int_0^1 \int_u^1 d(t-vh_n)\bar{F}(t-uh_n)\bar{F}(t-vh_n)k_i^{(1)}(v)dvk_i^{(1)}(u)du \\ &\hat{=} \frac{1}{nh_n^{2+2i}} \int_0^1 \{Q_1(u)+Q_2(u)\}\, k_i^{(1)}(u)du, \end{aligned} \tag{2.11}$$

with

$$Q_1(u) = d(t)\bar{F}^2(t) \int_0^u \left[1 + h_n \left\{u\frac{f(t)}{\bar{F}(t)} + v\frac{f(t)}{\bar{F}(t)} - u\frac{d^{(1)}(t)}{d(t)}\right\} + o(h_n)\right] k_i^{(1)}(v)dv$$

and

$$Q_2(u) = d(t)\bar{F}^2(t) \int_u^1 \left[1 + h_n \left\{u\frac{f(t)}{\bar{F}(t)} + v\frac{f(t)}{\bar{F}(t)} - v\frac{d^{(1)}(t)}{d(t)}\right\} + o(h_n)\right] k_i^{(1)}(v)dv,$$

where $d^{(1)}(t)$ denotes the derivative of

$$d(t) = \int_{-\infty}^t \frac{dF(t)}{\{1-W(t)\}\{1-F(t)\}^2}. \tag{2.12}$$

Together with the assumption (A1), we have

$$E_n I_{11}^2 = \frac{f(t)}{\{1-W(t)\}nh_n^{2i+1}} \int_0^1 k_i^2(u)du + o\left(\frac{1}{nh_n^{2i+1}}\right). \tag{2.13}$$

Secondly, also by Lemma 2.2,

$$\begin{aligned} E_n I_{12}^2 &\le c_{1i} \left(\frac{1}{nh_n^{1+i}}\right)^2 E\left\{\sup_{t\le T_0} |R_n(t)|\right\}^2 \\ &= c_{1i} \left(\frac{1}{nh_n^{1+i}}\right)^2 \int_0^\infty tP\left\{\sup_{t\le T_0} |R_n(t)| \ge t\right\} dt \\ &= O\left(\frac{\log^4 n}{n^2h_n^{2+2i}}\right). \end{aligned} \tag{2.14}$$

On the other hand, expanding $f(t-uh_n)$ and using the assumption (A1), we know

$$\begin{aligned} I_2 &= \frac{1}{h_n^i}\int_0^1 k_i(u)\left\{f(t)+\sum_{k=1}^{s-1}\frac{f^{(k)}(t)(-uh_n)^k}{k!}+\frac{f^{(s)}(t^*)(-uh_n)^s}{s!}\right\}du - f^{(i)}(t) \\ &= c_{2i}h_n^{s-i}f^{(s)}(t)+o(h_n^{s-i}), \quad t^*\in(t-uh_n,t). \end{aligned} \tag{2.15}$$

Following from (2.8)–(2.15) and taking $h_n = n^{-1/(2s+1)}$, we conclude that the Lemma 2.3 is true. □

Lemma 2.4 Let $\hat{F}(t)$ be given in (2.2). Then $E_n[\hat{F}(t)-F(t)]^2 = O(\log n/n)$.

Proof: Using the Theorem 3.2 of Földes and Rejtö (1981), we know

$$\sup_{t\le T_0}|\hat{F}(t)-F(t)| = O\left(\frac{\sqrt{\log n}}{\sqrt{n}}\right), \tag{2.16}$$

where T_0 is such that $1-H(T_0)>\epsilon$ with some $\epsilon>0$. Hence, the conclusion of Lemma 2.4 is obvious. □

Proof: **The proof of Theorem 2.1** Denote

$$\begin{aligned} E_n\{\hat{e}(T)-e(T)\}^2 &= E_n\{\hat{e}(T)-e(T)\}^2 I\{e(T)\le n^v\} \\ &\quad + E_n\{\hat{e}(T)-e(T)\}^2 I\{e(T)>n^v\} \\ &\hat{=}\ Q_1+Q_2. \end{aligned} \tag{2.17}$$

First, by Lemma 2.1, we have

$$\begin{aligned} Q_1 &= E_n\{\hat{e}(T)-e(T)\}^2 I\{e(T)\le n^v\} \\ &\le E_n|\{\hat{e}(T)-e(T)\}_{2n^v}|^2 \\ &= E\left|\left[\frac{-\hat{f}^{(1)}(T)\{1-W(T)\}+w(T)\hat{f}(T)}{\hat{f}(T)\{1-W(T)\}+w(T)\{1-\hat{F}(T)\}}-\frac{-f^{(1)}(T)\{1-W(T)\}+w(T)f(T)}{f(T)\{1-W(T)\}+w(T)\{1-F(T)\}}\right]_{2n^v}\right|^2 \\ &\le 4h^{-2}(T)[1W(T)]^2E_n[\hat{f}^{(1)}(T)-f^{(1)}(T)]^2+4h^{-2}(T)w^2(T)E_n[\hat{f}(T)-f(T)]^2 \\ &\quad +36n^{2v}h^{-2}(T)\{1-W(T)\}^2E_n\{\hat{f}(T)-f(T)\}^2 \\ &\quad +36n^{2v}h^{-2}(T)w^2(T)E_n\{\hat{F}(T)-F(T)\}^2. \end{aligned} \tag{2.18}$$

Second, using Lemma 2.3 and Lemma 2.4 and Theorem 2.1's conditions, we can easily obtain

$$E_TQ_1 \le c_4n^{-\frac{2s-2}{2s+1}}+c_5n^{-\left(\frac{2s}{2s+1}-2v\right)}. \tag{2.19}$$

On the other hand, note that $e(T)=E(\lambda|T)$, by Hölder's inequality, we have

$$\begin{aligned} E_TQ_2 &= E_T[E_n\{\hat{e}(T)-e(T)\}^2I\{e(T)>n^v\}] \\ &\le 4E_T[e^2(T)I\{e(T)>n^v\}] \\ &\le 4E_T[E(\lambda^2|T)I\{E(\lambda|T)>n^v\}] \\ &\le 4(E\lambda^{2s})^{\frac{1}{s}}[E\{E\lambda|T\}^{2s}n^{-2vs}]^{1-\frac{1}{s}} \\ &= 4n^{-2v(s-1)}E\lambda^{2s}. \end{aligned} \tag{2.20}$$

Together with (2.7) and taking $v = (2s+1)^{-1}$, we have

$$\begin{aligned} 0 &\le \sup_{t\le T_0}\{R(p_{EB}(t), G(\lambda)) - R(p_G(t), G(\lambda))\} \\ &\le E_T E_n|\hat{e}(T) - e(T)| \\ &\le E_T[E_n\{\hat{e}(T) - e(T)\}^2]^{\frac{1}{2}} \\ &\le (E_T Q_1 + E_T Q_2)^{\frac{1}{2}} \\ &= O\left(n^{-\frac{s-1}{2s+1}}\right). \end{aligned} \tag{2.21}$$

The proof of Theorem 2.1 is finished. □

Remark 2.1 If there is no censorship, the Theorem 2.1's conditions and conclusion will become $E[f^{-2}(X)\{f^{(s)}(X)\}^2] < \infty$, $E\lambda^{2s} < \infty$ and $\sup\{R(p_{EB}(x), G(\lambda)) - R(p_G(x), G(\lambda))\} = O\{n^{-(s-1)/(2s+1)}\}$. Obviously, due to the existence of censorship, the same problem gets very complicated.

3. EB Test for the Case that W is Unknown

Since $W(t)$ is unknown in the expression of $e(t)$, we estimate it by its product limit estimator $\hat{W}(t)$ (see, Kaplan and Meier, 1958) given by

$$1 - \hat{W}(t) = \prod_{i=1}^{n}\left(\frac{n-i}{n-i+1}\right)^{I(T_{(i)}\le t,\ \delta_{(i)}=0)}, \quad \text{if } t < T_{(n)}. \tag{3.1}$$

Accordingly, we propose an estimator of $w(t)$ as

$$\hat{w}(t) = \frac{1}{h_n}\int k_2\left(\frac{t-y}{h_n}\right) d\hat{W}(y), \tag{3.2}$$

where $k_2(x)$ is a kernel function.

Note that $\{1 - W(t)\}f(t) + w(t)\{1 - F(t)\} = h(t)$ as $t > 0$, by $(T_1, T_2, \ldots, T_n)$ and T, we define the kernel estimation of $h(t)$ as

$$\hat{h}(t) = \frac{1}{nh_n}\sum_{i=1}^{n} k_3\left(\frac{T_i - t}{h_n}\right), \tag{3.3}$$

where $k_3(x)$ is a kernel function.

Together with Section 2, finally we propose the following EB estimator for $e(t)$,

$$\bar{e}(t) = \frac{-\hat{f}^{(1)}(t)\{1 - \hat{W}(t)\} + \hat{w}(t)\hat{f}(t)}{A_n(t)}, \tag{3.4}$$

with

$$A_n(t) = \begin{cases} \hat{h}(t), & \text{if } \hat{h}(t) \ge \theta_n, \\ \theta_n, & \text{if } \hat{h}(t) < \theta_n, \end{cases} \tag{3.5}$$

where $0 < \theta_n \to 0$ as $n \to \infty$.

Hence, in this case the EB test is defined by

$$\bar{p}_{EB}(t) = \begin{cases} 1, & \text{if } \bar{e}(t) \le \lambda_0, \\ 0, & \text{if } \bar{e}(t) > \lambda_0. \end{cases} \tag{3.6}$$

Its overall Bayes risk is

$$R(\bar{p}_{EB}(t), G(\lambda)) = \int_0^\infty m(t) E_n\{\bar{p}_{EB}(t)\} dt + \int_\Lambda L(\lambda, d_1) dG(\lambda). \tag{3.7}$$

Obviously, similar to (2.7), we have

$$0 \le R(\bar{p}_{EB}(t), G(\lambda)) - R(p_G(t), G(\lambda)) \le E_T E_n |\bar{e}(T) - e(T)|. \tag{3.8}$$

In what follows, let $c, c_0, c_1, \ldots,$ denote positive constants that do not depend on n and they can take different values while appearing even within the same expression.

Lemma 3.1 Let $\hat{f}^{(i)}(t)$ be defined in (2.3). Under the assumption (A1) with $s = 2$,

$$E_n\{\hat{f}^{(i)}(t) - f^{(i)}(t)\}^2 \le c_1 h_n^{4-2i}\{f^{(2)}(t)\}^2 + c_2 \frac{f(t)}{\{1 - W(t)\} n h_n^{2i+1}} + c_3 \frac{\log^4 n}{n^2 h_n^{2+2i}},$$
$$i = 0, 1.$$

Mimicking the proof of Lemma 2.3.

Note that $F(t)$ and $W(t)$ are opposite in the censorship model, we have Lemma 3.2.

Lemma 3.2 Let $\hat{w}(t)$ be defined in (3.2). Assume that $w(t)$ has the second order derivative. For a suitable kernel $k_2(x)$, which has the same properties as $k_0(x)$, then for $t \le T_0$,

$$E_n\{\hat{w}(t) - w(t)\}^2 \le c_1 h_n^4 \{w^{(2)}(t)\}^2 + c_2 w(t)\{1 - F(t)\}^{-1}(n h_n)^{-1} + c_3 (n^2 h_n^2)^{-1} \log^4 n.$$

Proof: Replace $F(t)$ by $W(t)$ in Lemma 2.2 and follow a similar proof to that of Lemma 3.1.

Assume that $k_3(x)$ is a non-negative bounded Borel measurable function satisfying

$$(1)\ k_3(x) = 0, x \notin (0,1), \quad (2) \int_0^1 k_3(x) dx = 1.$$

Then by (3.3), we have the following lemma. □

Lemma 3.3 Let $\hat{h}(t)$ be given by (3.3) with the kernel function $k_3(x)$ satisfying the above assumption. If $\sup_t w^{(1)}(t) < \infty$ and $E\lambda^2 < \infty$, then

$$E_n\{\hat{h}(t) - h(t)\}^2 \le \frac{c_1}{n h_n} + c_2 h_n^2,$$

where $w^{(1)}(t)$ denotes the derivative of $w(t)$.

Proof: Note that

$$E_n\{\hat{h}(t)-h(t)\}^2 \le 2E_n\{\hat{h}(t)-E_n\hat{h}(t)\}^2+2\{E_n\hat{h}(t)-h(t)\}^2 \hat{=} 2(J_1+J_2). \tag{3.9}$$

First, one has

$$\begin{aligned} J_1 &= \mathrm{Var}[\hat{h}(t)] \\ &\le \frac{1}{nh_n^2}E\left[k_3^2\left(\frac{T_1-t}{h_n}\right)\right] \\ &= \frac{1}{nh_n}\int_0^1 k_3^2(u)h(t+h_nu)du. \end{aligned} \tag{3.10}$$

By $h(t+h_nu)=\{1-W(t+h_nu)\}f(t+h_nu)+w(t+h_nu)\{1-F(t+h_nu)\}$ and the monotonicity of $f(t)$ and $E\lambda^2<\infty$, we have

$$J_1 \le \frac{c_1}{nh_n}. \tag{3.11}$$

Secondly, since

$$E_n\hat{h}(t)=\int_0^1 K_3(u)h(t+h_nu)du, \tag{3.12}$$

expanding $h(t+h_nu)=h(t)+[\{1-W(\xi)\}f^{(1)}(\xi)+w^{(1)}(\xi)[\{1-F(\xi)]\}-2w(\xi)f(\xi)]h_nu$, where $t<\xi<t+h_nu$ and using $\sup_t w^{(1)}(t)<\infty$, one has $|E_n\hat{h}(t)-h(t)|<c_2h_n$.

Thus

$$J_2 \le c_2h_n^2. \tag{3.13}$$

Therefore, by (3.11) and (3.13), Lemma 3.3 is proved. Now we state the main result of this section. □

Theorem 3.1 Let $R(p_G(t),G(\lambda))$ and $R(\bar{p}_{EB}(t),G(\lambda))$ be defined by (1.2) and (3.7), respectively. If $E\lambda^2<\infty$ and the following conditions are satisfied:

$$\text{(a) } E\left[\frac{f(T)}{1-W(T)}\right]<\infty, \quad E\left[\frac{w(T)}{1-F(T)}\right]<\infty,$$

$$\text{(b) } E\left[\bar{d}(T)(f^{(1)}(T))^2\right]<\infty,$$

$$\text{(c) } E\left[w^{(2)}(T)\right]^2<\infty, \quad \sup_t w^{(1)}(t)<\infty.$$

Then, choosing $nh_n^6\to\infty$, $nh_n\theta_n^2\to\infty$ and $h_n=o(\theta_n)$ as $n\to\infty$, for $t\le T_0$, which is such that $1-H(T_0)>\epsilon$ with some $\epsilon>0$, we have

$$\lim_{n\to\infty} R(\bar{p}_{EB}(t),G(\lambda))=R(p_G(t),G(\lambda)),$$

where $\bar{d}(T)=\int_0^T dW(y)/[\{1-F(y)\}\{1-W(y)\}^2]$.

Proof: From (3.8), we have

$$0 \le \lim_{n\to\infty} R(\bar{p}_{EB}(t), G(\lambda)) - R(p_G(t), G(\lambda))$$
$$\le \sqrt{\lim_{n\to\infty} E_T E_n\{\bar{e}(T) - e(T)\}^2}. \tag{3.14}$$

To obtain the Theorem's conclusion, by the dominated convergence theorem, we need prove the following two steps:

(i) $E_n\{\bar{e}(T) - e(T)\}^2 \le M(T)$ and $E_T M(T) < \infty$,

(ii) $\lim_{n\to\infty} E_n\{\bar{e}(T) - e(T)\}^2 = 0$, for any fixed T.

In order to prove (ii), by (2.1) and (3.4), one has

$$E_n\{\bar{e}(T) - e(T)\}^2$$
$$\le 2E_n\left[\frac{-\hat{f}^{(1)}(T)\{1-\hat{W}(T)\} + \hat{w}(T)\hat{f}(T) + \{1-W(T)\}f^{(1)}(T) - w(T)f(T)}{A_n(T)}\right]^2$$
$$+ 2e^2(T)E_n\left[\frac{\{1-W(T)\}f(T) + w(T)\{1-F(T)\}}{A_n(T)} - 1\right]^2$$
$$\hat{=} 2\left\{Q^{(1)} + e^2(T)Q^{(2)}\right\}. \tag{3.15}$$

Since $A_n(T) \ge \theta_n$, together with the fact that

$$E_n\Big[-\hat{f}^{(1)}(T)\left\{1-\hat{W}(T)\right\} + \hat{w}(T)\hat{f}(T) + \left\{1-W(T)\right\}f^{(1)}(T) - w(T)f(T)\Big]^2$$
$$\le 12E_n\left\{f^{(1)}(T) - \hat{f}^{(1)}(T)\right\}^2 + 8E_n\left[\{\hat{W}(T) - W(T)\}f^{(1)}(T)\right]^2$$
$$+4E_n\left[\{\hat{w}(T) - w(T)\}\hat{f}(T)\right]^2 + 4E_n\left[\{\hat{f}(T) - f(T)\}w(T)\right]^2, \tag{3.16}$$

where we use the fact that $\hat{W}(T) \le 1$, we have

$$Q^{(1)} \le \theta_n^{-2}\Big[12E_n\left\{f^{(1)}(T) - \hat{f}^{(1)}(T)\right\}^2 + 8E_n\left[\{\hat{W}(T) - W(T)\}f^{(1)}(T)\right]^2$$
$$+4E_n\left[\{\hat{w}(T) - w(T)\}\hat{f}(T)\right]^2 + 4E_n\left[\{\hat{f}(T) - f(T)\}w(T)\right]^2\Big]. \tag{3.17}$$

Firstly, by Lemma 3.1, as $nh_n^6 \to \infty$, we have

$$E_n\left\{f^{(1)}(T) - \hat{f}^{(1)}(T)\right\}^2 \le c_1\left[\frac{f(T)}{1-W(T)} + \{f^{(2)}(T)\}^2\right]h_n^2. \tag{3.18}$$

Secondly, following Lemma 2.4, similar to the above,

$$E_n\left[\{\hat{W}(T) - W(T)\}f^{(1)}(T)\right]^2 \le c_2\left[\frac{\bar{d}(T)}{n}\{f^{(1)}(T)\}^2 + \frac{\log^4 n}{n^2}\{f^{(1)}(T)\}^2\right]. \tag{3.19}$$

Thirdly, by Lemma 3.2 and the fact that $\hat{f}(t) < ch_n^{-1}$, we know

$$E_n\left[\{\hat{w}(T) - w(T)\}\hat{f}(T)\right]^2 \le c_3\left[\frac{w(T)}{1-F(T)} + \{w^{(2)}(T)\}^2\right]h_n^2, \tag{3.20}$$

as $nh_n^6 \to \infty$.

Finally, under the condition that $\sup_t w(t) < \infty$, one has

$$E_n\left[\{\hat{f}(T) - f(T)\}w(T)\right]^2 \le c_4\left[\frac{f(T)}{1-W(T)} + \{f^{(2)}(T)\}^2\right]h_n^4, \tag{3.21}$$

as $nh_n^6 \to \infty$.

Following from (3.18)–(3.21) and choosing h_n and θ_n such that:

$$nh_n^6 \to \infty, \quad n\theta_n^2 \to \infty, \quad h_n = o(\theta_n),$$

we obtain

$$\lim_{n\to\infty} Q^{(1)} = 0. \tag{3.22}$$

On the other hand, we write

$$\begin{aligned} Q^{(2)} &= E_n\left\{\frac{h(T)-A_n(T)}{A_n(T)}\right\}^2 I(h(T) \ge \theta_n + E_n\left\{\frac{h(T)-A_n(T)}{A_n(T)}\right\}^2 I(h(T) < \theta_n) \\ &\triangleq Q^{(21)} + Q^{(22)}. \end{aligned} \tag{3.23}$$

Also by $A_n(T) \ge \theta_n$ and the fact that

$$h(T) - A_n(T) = \begin{cases} h(T) - \hat{h}(T), & \text{if } \hat{h}(T) \ge \theta_n, \\ h(T) - \theta_n < h(T) - \hat{h}(T), & \text{if } \hat{h}(T) < \theta_n, \end{cases} \tag{3.24}$$

we have

$$Q^{(21)} \le \theta_n^{-2} E_n\left\{h(T) - \hat{h}(T)\right\}^2. \tag{3.25}$$

It follows from Lemma 3.3,

$$Q^{(21)} \le \theta_n^{-2}\left\{\frac{c_1}{nh_n} + c_2h_n^2\right\}. \tag{3.26}$$

Moreover, note that

$$Q^{(22)} \le 4I(h(T) < \theta_n). \tag{3.27}$$

Choosing $nh_n\theta_n^2 \to \infty$ and $h_n = o(\theta_n)$, it follows from (3.23), (3.26) and (3.27)

$$\lim_{n\to\infty} e^2(T)Q^{(2)} = 0. \tag{3.28}$$

Hence, combining (3.15) and (3.22) with (3.28), we know the step (ii) is true. The task to remain is to prove (i).

Note that $f^{(1)}(T) < E\lambda^2$, we have

$$Q^{(1)} \leq c_1 \left[\frac{f(T)}{1-W(T)} + \{f^{(2)}(T)\}^2\right] + c_2 \left[\frac{w(T)}{1-F(T)} + \{w^{(2)}(T)\}^2\right] \\ +c_3\bar{d}(T)\left\{f^{(1)}(T)\right\}^2. \tag{3.29}$$

Obviously

$$e^2(T)Q^{(2)} \leq c_4 e^2(T). \tag{3.30}$$

Then by (3.15), we obtain

$$E_n\{\bar{e}(T) - e(T)\}^2 \leq c_1 \left[\frac{f(T)}{1-W(T)} + \{f^{(2)}(T)\}^2\right] + c_2 \left[\frac{w(T)}{1-F(T)} + \{w^{(2)}(T)\}^2\right] \\ +c_3\bar{d}(T)\left\{f^{(1)}(T)\right\}^2 + c_4 e^2(T) \\ \hat{=} M(T). \tag{3.31}$$

Under the conditions of the Theorem 3.1, we know

$$E_T\left[M(T)\right] < \infty. \tag{3.32}$$

We have proved the step (i). The proof of the Theorem 3.1 is complete. □

Remark 3.1 If there is no censorship, in order to obtain the asymptotic optimality of the proposed EB test, we only need the conditions that $E\lambda^2 < \infty$ and $E[f(X)] < \infty$.

4. Concluding Remarks

Although the conditions of Theorem 2.1 and Theorem 3.1 are a little complicated, it is easy to verify that for a gamma prior $g(\lambda) = 1/\Gamma(l)\lambda^{l-1}\exp(-\lambda)I(\lambda > 0, l > 0)$ and a censoring distribution $W(y) = 1 - (y+1)^{-\epsilon_0}$, where $y > 0$ and $\epsilon_0 > l$, all the conditions of Theorem 2.1 and Theorem 3.1 are satisfied.

There are few papers dealing with the EB testing problem of hypothesis under censorship. In this paper, when the historical samples $X_1, X_2, \ldots, X_n$ and the present sample X are both randomly censored from the right by a sequence $Y_1, Y_2, \ldots, Y_n$ and Y with a distribution function W, we propose an EB test firstly, under the condition that the censoring distribution W is known and secondly, that it is unknown.

Under some suitable conditions, for the case that W is known, the convergence rate of the proposed EB test can be very close to $O(n^{-1/2})$. Liang (2000a) claimed that the order of convergence rate is $O(n^{-s/(s+3)})$ for a positive exponential family under the assumption that the critical point, corresponding to the prior, is within an unknown compact interval. Similar result also appears in Liang (2000b). It is not difficult to find that the Bernstein's inequality plays an important role in these two papers. However, Bernstein's inequality cannot work in our case since the data are censored.

It is necessary to note that in this paper the proposed EB tests are not monotone. If we make some assumptions on the critical point and take the monotonicity of $e(t)$ into account, then we can construct a monotone EB test which can improve the order $O(n^{-1/2})$ to be close to $O(n^{-1})$ under some conditions.

Acknowledgement

The author would like to thank two anonymous referees and the editor for their helpful suggestions and comments which improved the paper.

References

Földes, A. and Rejtö, L. (1981). Strong uniform consistency for nonparametric survival curve estimators from randomly censored data. *The Annals Statistics*, **19**, 122–129.

Johns, M. V. Jr. and Van Ryzin, J. R. (1971). Convergence rates for empirical Bayes two-action problem I: Discrete case. *The Annals of Mathematical Statistics*, **42**, 1521–1539.

Johns, M. V. Jr. and Van Ryzin, J. R. (1972). Convergence rates for empirical Bayes two-action problem II: Continuous case. *The Annals of Mathematical Statistics*, **43**, 934–947.

Kaplan, E. L. and Meier, P. (1958). Nonparametric estimation from incomplete observations. *Journal of the American Statistical Association*, **53**, 457–481.

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

Lemdani, M. and Ould-Saïd, E. (2002). Exact asymptotic $\mathcal{L}_1$-error of a kernel density estimator under censored data. *Statistics & Probability Letters*, **60**, 59–68.

Liang, T. C. (2000a). On empirical Bayes tests in a positive exponential family. *Journal of Statistical Planning and Inference*, **83**, 169–181.

Liang, T. C. (2000b). On an empirical Bayes test for a normal mean. *The Annals of Statistics*, **28**, 648–655.

Major, P. and Rejtö, L. (1988). Strong embedding of the estimator of the distribution function under random censorship. *The Annals of Statistics*, **16**, 1113–1132.

Robbins, H. (1956). An empirical Bayes approach to statistics. In *Proceedings of the Third Berkeley Symposium on Mathematical Statistics and Probability*, University of California Press, **1**, 157–163.

Robbins, H. (1964). The empirical Bayes approach to statistical decision problem. *The Annals of Mathematical Statistics*, **35**, 1–20.

Singh, R. S. (1977). Applications of estimators of a density and its derivatives to certian statistical problems. *Journal of the Royal Statistical Society*, Ser. B, **39**, 357–363.

Singh, R. S. (1979). Empirical Bayes estimation in lebesgue-exponential families with rates near the best possible rate. *The Annals of Statistics*, **7**, 890–902.

Stijnen, T. (1985). On the asymptotic behaviour of empirical Bayes tests for the continuous one-parameter exponential family. *The Annals of Statistics*, **13**, 403–412.

Van Houwelingen, J. C. (1976). Monotone empirical Bayes test for the continuous one-parameter exponential family. *The Annals of Statistics*, **4**, 981–989.

[Received November 2007, Accepted January 2008]