Sankhyā : The Indian Journal of Statistics
2015, Volume 77-A, Part 2, pp. 364-379

Empirical Bayes Test Problem in Continuous One-parameter Exponential Families under Dependent Samples

Qingzhu Lei and Yongsong Qin
Guangxi Normal University, Guilin, China

Abstract

In this paper, we study the empirical Byes (EB) test problem in the continuous one-parameter exponential family under associated samples and strong mixing samples. Under mild regularity conditions, it is shown that the convergence rates of proposed EB test rules under associated or strong mixing samples are the same as that of EB test rules under independent observations.

AMS (2000) *subject classification.* Primary 62C12; Secondary 62F12.
Keywords and phrases. Dependent sample, Continuous one-parameter exponential family, Empirical Bayes test rule, Convergence rate

1 Introduction

Empirical Bayes (EB) method is a procedure for statistical inference proposed by Robbins (1955). For the continuous one-parameter exponential family, a basic and important distribution family, Lin (1975) and Singh (1979) investigated the convergence rates of EB estimators. Johns and Van Ryzin (1972) and Wei (1989) studied the EB test problem for this family. However, these results only focus on independent samples. Since dependent data appear quite often in applications, statistical inference for dependent data is naturally necessary and important. For negatively associated (NA) observations, Chen and Wei (2006) obtained the rates of convergence of EB estimators in the continuous one-parameter exponential family. Wang, Zhang and Guo 2008 generalized the results in Chen and Wei (2006) to positively associated (PA) observations. Chen and Wei (2004) also studied the EB test problem in this family under NA samples.

The NA conception was introduced by Block, Savits and Sharked 1982 and Joag-Dev and Proschan (1983), and the PA concept was introduced by Esary, Proschan and Walkup 1967. Either PA or NA random variables are called associated random variables. A brief review of the applications of NA and PA random variables can be found in Roussas (2000). In this paper, apart from EB test under associated samples, EB test under α-mixing samples is also investigated. The dependence described by α-mixing introduced by Rosenblatt (1956) is the weakest among well-known mixing structures, as it is implied by other types of mixing such as the widely used ϕ, ρ and β-mixings; see Doukhan (1987) for a comprehensive discussion on mixing.

In this paper, we study the EB test problem in the continuous one-parameter exponential family under associated samples and strong mixing samples. Under certain regularity conditions, it is shown that the convergence rates of proposed EB test rules under associated or mixing samples are the same as that of EB test rules under independent observations. We note here that the EB test problem for the continuous one-parameter exponential family has been studied by Johns and Van Ryzin (1972) and Wei (1989) under independent observations and by Chen and Wei (2004) under NA settings and that the EB test problem for this distribution family under other dependent structures has not be investigated yet. On the other hand, we also show that the convergence rates of the EB test under NA settings are better than those in Chen and Wei (2004) (e. g. Remark 4 below).

We now state the definitions of NA, PA and α-mixing random variable sequences.

Random variables $\{\eta_i, 1 \leq i \leq n\}$ are said to be negatively associated (NA), if for every pair of disjoint subsets A_1, A_2 of $\{1, 2, \cdots, n\}$ and any real-valued coordinatewise increasing functions f and g,

$$Cov\{f(\eta_i, i \in A_1), g(\eta_j, j \in A_2)\} \leq 0,$$

provided $Ef^2(\eta_i, i \in A_1) < \infty, Eg^2(\eta_j, j \in A_2) < \infty$. Random variables $\{\eta_i, 1 \leq i \leq n\}$ are said to be positively associated (PA), if for any real-valued coordinatewise nondecreasing functions f and g,

$$Cov(f(\eta_1, \eta_2, \cdots, \eta_n), g(\eta_1, \eta_2, \cdots, \eta_n)) \geq 0,$$

whenever this covariance exists. Infinitely many r.v.s. are said to be PA(NA), if any finite subset of them is a set of PA(NA) r.v.s.

Let $\{\eta_i, 1 \geq 1\}$ be a random variable sequence and $\mathcal{F}_s^t$ denote the σ-algebra generated by $\{\eta_i, s \leq i \leq t\}$ for $s \leq t$. Random variables $\{\eta_i, 1 \geq 1\}$ are said to be strongly mixing or α-mixing (e. g. Rosenblatt, 1956), if

$$\alpha(n) = \sup_{k \geq 1} \sup_{A \in \mathcal{F}_1^k, B \in \mathcal{F}_{k+n}^\infty} |P(AB) - P(A)P(B)| \to 0$$

as $n \to \infty$, where $\alpha(n)$ is called α-mixing coefficient.

The rest of the paper is organized as follows. The main results of this paper are presented in Section 2. Some lemmas to prove the main results are given in Section 3. The proofs of the main results are presented in Section 4.

2 Main Results

2.1. Bayes Test Function. Let $(X, \theta) \in R^2$ be a random vector with parameter θ and $f(x|\theta)$ be the conditional probability density function (p.d.f.) of X given θ. The continuous one-parameter exponential family has the following form:

$$f(x|\theta) = c(\theta)u(x)\exp(\theta x), \tag{1}$$

where $x \in \mathcal{X} = (a, b), -\infty \leq a < b \leq \infty$, $\theta \in \Theta = \{\theta : \int_{\mathcal{X}} u(x)\exp(\theta x) < \infty\}$, $u(x) > 0$ as $x \in \mathcal{X}$ and $c(\theta) \geq 0$. $\mathcal{X}$ and Θ are called sample space and parameter space, respectively. Let $G(\theta)$ be the (unknown) prior distribution function of θ.

We consider the following two hypotheses:

$$\text{(I) } H_0 : \theta \leq \theta_0 \ \text{ vs. } \ H_1 : \theta > \theta_0$$

and

$$\text{(II) } H_0 : \theta_1 \leq \theta \leq \theta_2 \ \text{ vs. } \ H_1 : \theta < \theta_1 \text{ or } \theta > \theta_2,$$

where θ_0, θ_1 and θ_2 are all known constants.

Let $D = \{d_0, d_1\}$ be the action space, where d_0 stands for accepting H_0 and d_1 stands for rejecting H_0. For Hypothesis (I), we use the following loss function

$$L(\theta, i) = (1 - i)a(\theta - \theta_0)I(\theta > \theta_0) + ia(\theta - \theta_0)I(\theta \leq \theta_0), i = 0, 1,$$

where a is a given constant, $I(A)$ is the indicator function of a set A, $i = 0$ stands for taking action d_0 and $i = 1$ stands for taking action d_1.

Let

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

be a randomized decision rule for the above test problem. The Bayes risk (Johns and Van Ryzin, 1972) of $\delta(x)$ is

$$R(\delta, G) = a\int_{\mathcal{X}} \delta(x)\alpha(x)dx + C_G.$$

where

$$C_G = \int_{\Theta} L_1(\theta, d_1)dG(\theta),$$

$$\alpha(x) = \int_\Theta \theta f(x|\theta)dG(\theta) - \theta_0 f(x) = f^{(1)}(x) - \{u^{(1)}(x)/u(x) + \theta_0\}f(x),$$

where $f(\cdot)$ is the marginal p.d.f. of X, $f^{(1)}(x) = df(x)/dx$ and $u^{(1)}(x) = du(x)/dx$. We assume that both $f^{(1)}(x)$ and $u^{(1)}(x)$ exist. Then the Bayes test rule (Johns and Van Ryzin, 1972), or Bayes test function, is

$$\delta_G(x) = \begin{cases} 1, & \text{if } \alpha(x) \leq 0, \\ 0, & \text{if } \alpha(x) > 0, \end{cases}$$

which has the following Bayes risk

$$R(G) = R(\delta_G, G) = a\int_{\mathcal{X}} \delta_G(x)\alpha(x)dx + C_G.$$

Let $a_0 = (\theta_1 + \theta_2)/2$, $\mu_0 = (\theta_2 - \theta_1)/2$ and $D = \{d_0, d_1\}$ be the action space, where d_0 stands for accepting H_0 and d_1 stands for rejecting H_0. Then hypothesis (II) is equivalent to

$$\text{(III) } H_0^*: |\theta - a_0| \leq \mu_0 \text{ vs. } H_1^*: |\theta - a_0| > \mu_0.$$

For hypothesis (III) and $i = 0, 1$, we use the following loss function

$$\begin{aligned} L^*(\theta, i) &= (1-i)a\{(\theta - a_0)^2 - \mu_0^2\}I(|\theta - a_0| > \mu_0) \\ &\quad + ia\{\mu_0^2 - (\theta - a_0)^2\}I(|\theta - a_0| \leq \mu_0), i = 1, 2, \end{aligned}$$

where a is a given constant, $I(A)$ is the indicator function of a set A, $i = 0$ stands for taking action d_0 and $i = 1$ stands for taking action d_1.

Let

$$\delta^*(x) = P(\text{ accepting } H_0^*|X = x))$$

be a randomized decision rule for the above test problem. The Bayes risk (Wei, 1989) of $\delta^*(x)$ is

$$R(\delta^*, G) = a\int_{\mathcal{X}} \delta^*(x)\alpha^*(x)dx + C_G^*.$$

where

$$C_G^* = \int_\Theta L_1^*(\theta, d_1)dG(\theta),$$

$$\begin{aligned} \alpha^*(x) &= f^{(2)}(x) - 2\left\{u^{(1)}(x)/u(x) + a_0\right\}f^{(1)}(x) \\ &\quad + \left[2\left\{u^{(1)}(x)/u(x)\right\}^2 + 2a_0u^{(1)}(x)/u(x) - u^{(2)}(x)/u(x) + a_0^2 - \mu_0^2\right]f(x), \end{aligned}$$

where $f(\cdot)$ is the marginal p.d.f. of X, $f^{(j)}(x)$ and $u^{(j)}(x)$ are the j-th order derivatives of f and u, respectively. We assume that both $f^{(2)}(x)$ and $u^{(2)}(x)$ exist. Then the Bayes test rule (Wei, 1989), or Bayes test function, is

$$\delta_G^*(x) = \begin{cases} 1, & \text{if } \alpha^*(x) \le 0, \\ 0, & \text{if } \alpha^*(x) > 0, \end{cases}$$

which has the following Bayes risk

$$R^*(G) = R(\delta_G^*, G) = a \int_{\mathcal{X}} \delta_G^*(x)\alpha^*(x)dx + C_G^*.$$

2.2. Empirical Bayes Test Function. Suppose that we have a sample $X_1, X_2, \cdots, X_n$ from the population X. Let $f^{(0)}(x) = f(x)$. To obtain an EB test rule, or EB test function, we need to use the sample to estimate $f^{(j)}(x)$ for $j = 0, 1, 2$. From Rao (1983), the kernel estimators of $f^{(j)}(x)$ are

$$f_n^{(j)}(x) = \frac{1}{nh^{j+1}} \sum_{i=1}^{n} K_j\left(\frac{x - X_i}{h}\right), j = 0, 1, 2,$$

where $K_j (j = 0, 1, 2)$ are Borel measurable functions and $h = h_n (0 < h \to 0)$ are bandwidths.

Let

$$\alpha_n(x) = f_n^{(1)}(x) - \{u^{(1)}(x)/u(x) + \theta_0\} f_n^{(0)}(x).$$

The EB test rule for hypothesis (I) is defined as

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

Let

$$\begin{aligned} \alpha_n^*(x) &= f_n^{(2)}(x) - 2\left\{u^{(1)}(x)/u(x) + a_0\right\} f_n^{(1)}(x) \\ &\quad + \left[2\left\{u^{(1)}(x)/u(x)\right\}^2 + 2a_0 u^{(1)}(x)/u(x) - u^{(2)}(x)/u(x) + a_0^2 - \mu_0^2\right] f_n^{(0)}(x), \end{aligned}$$

The EB test rule for hypothesis (II) is defined as

$$\delta_n^*(x) = \begin{cases} 1, & \text{if } \alpha_n^*(x) \le 0, \\ 0, & \text{if } \alpha_n^*(x) > 0, \end{cases}$$

Throughout this paper, we use E_n to denote the expectation under the distribution of $(X_1, X_2, \cdots, X_n)$. The Bayes risks of $\delta_n(x)$ and $\delta_n^*(x)$ are respectively

$$R_n = a\int_{\mathcal{X}} E_n\{\delta_n(x)\}\alpha(x)dx + C_G \tag{2}$$

and

$$R_n^* = a\int_{\mathcal{X}} E_n\{\delta_n^*(x)\}\alpha(x)dx + C_G^*. \tag{3}$$

If $\lim_{n\to\infty} R_n = R(G)$, $\delta_n(x)$ is called an asymptotically optimal EB test function. If $R_n - R(G) = O(n^{-t})$, n^{-t} is called the convergence rate of $\delta_n(x)$, where $t > 0$. Similarly, one can define the asymptotical optimality and convergence rate of $\delta_n^*(x)$.

2.3. Assumptions. Throughput this paper, we use C to denote a positive constant independent of n, which may take a different value for each appearance. Before we state the main results of this paper, we list some assumptions.

Assumptions

(A1) (i) The r.v.s. $X_1, X_2, \cdots, X_n$ form a stationary sequence with finite variance.

(ii) The p.d.f. $f(\cdot)$ of X_1 exists. There exists $s \geq 2$ such that $f^{(s)}(x)$ exists for all $x \in R$ and $f^{(s)}(\cdot)$ is a bounded function.

(A2) The functions $K_j(j = 0, 1, 2)$ are bounded, the derivatives $|K_j^{(1)}| \leq C$, $K_j(j = 0, 1, 2)$ have compact supports and satisfy

$$\int_R u^t K_j(u)du = \begin{cases} 1, & \text{if } t = j, \\ 0, & \text{if } t \neq j, t \in \{0, 1, 2, \cdots, s-1\}, \end{cases}$$

where t is a nonnegative integer with s appearing in Assumption A1 (ii).

(A3) (i) The r.v.s. $X_1, X_2, \cdots, X_n$ form a NA sequence.

(A4) (i) The r.v.s. $X_1, X_2, \cdots, X_n$ form a PA sequence.

(ii) If $f_{1,j}$ is the joint p.d.f. of the r.v.s X_1, X_{j+1}, then $|f_{1,j}(u, v) - f(u)f(v)| \leq C$ for all $u, v \in R$ and $j > 1$.

(iii) Let $p = p(n)$ and $q = q(n)$ be positive integers satisfying $p+q \leq n$, and $k = [n/(p+q)]$, where $[t]$ denotes the integral part of t. p, q and h satisfy $q \leq p$, $ph \leq C$, $\frac{1}{h^3}\sum_{r=q}^{\infty} |Cov(X_1, X_{r+1})| \leq C$.

(A5) (i) The r.v.s. $\{X_i, 1 \leq i \leq n\}$ is a α-mixing sequence with mixing coefficient $\alpha(\cdot)$.

(ii) If $f_{1,j}$ is the joint p.d.f. of the r.v.s X_1, X_{j+1}, then $|f_{1,j}(u,v) - f(u)f(v)| \leq C$ for all $u, v \in R$ and $j > 1$.

(iii) Let $p = p(n)$ and $q = q(n)$ be positive integers satisfying $p+q \leq n$, and $k = [n/(p+q)]$, where $[t]$ denotes the integral part of t. p, q and h satisfy $q \leq p$, $ph \leq C$, $(1/h)\sum_{t=q}^{\infty} \alpha(t) \to 0$.

REMARK 1. We will use the small-block and large-block arguments to obtain the convergence rates of the EB test under dependent samples. These arguments are widely used in deriving the asymptotic properties for the sums of dependent random variables (e.g. Roussas, 2000). To use this method, some restrictions on the sizes of blocks such as p and q are needed. Similar conditions of A4 (iii) and A5(iii) can be found in Roussas (2000).

REMARK 2. In Condition A4(iii), if we assume that $Cov(X_1, X_{r+1}) = O(r^{-d}), d > 1$, then $\sum_{r=q}^{\infty} |Cov(X_1, X_{r+1})| \leq Cq^{-(d-1)}$. If we let $p = [n^{a_0}], q = [n^{b_0}], h = n^{-c_0}$, then A4(iii) is satisfied if

$$0 < b_0 \leq a_0 < 1, 0 < a_0 \leq c_0 < 1, 3c_0 \leq b_0(d-1). \tag{4}$$

The value a_0 in (4) can be attained if we choose b_0 and c_0 such that

$$0 < b_0 < 1, 0 < c_0 < 1, 3c_0 < b_0(d-1). \tag{5}$$

The value b_0 in (5) can be attained if we choose c_0 such that

$$0 < c_0 < 1, c_0 < (d-1)/3. \tag{6}$$

The value c_0 in (6) can be attained if $d > 1$. In our main results, we choose $h = n^{-1/(2s+1)}$. To ensure that this h is achievable, we need assume that $d > 1 + 3/(2s+1)$.

REMARK 3. In Condition A5(iii), if we assume that $\alpha(r) = O(r^{-d}), d > 1$, then $\sum_{r=q}^{\infty} \alpha(r) \leq Cq^{-(d-1)}$. If we let $p = [n^{a_0}], q = [n^{b_0}], h = n^{-c_0}$, then A5(iii) is satisfied if

$$0 < b_0 \leq a_0 < 1, 0 < a_0 \leq c_0 < 1, c_0 \leq b_0(d-1). \tag{7}$$

The value a in (7) can be attained if we choose b_0 and c_0 such that

$$0 < b_0 < 1, 0 < c_0 < 1, c_0 < b_0(d-1). \tag{8}$$

The value b_0 in (8) can be attained if we choose c_0 such that

$$0 < c_0 < 1, c_0 < d - 1. \tag{9}$$

The value c_0 in (9) can be attained if $d > 1$. In our main results, we choose $h = n^{-1/(2s+1)}$. To ensure that this h is achievable, we need assume that $d > 1 + 1/(2s + 1)$.

2.4. Main Results. We now state the main results in this paper.

THEOREM 1. *Suppose that Assumption (A1, A2, A3), Assumption (A1, A2, A4) or Assumption (A1, A2, A5) is satisfied,* $(X_1, X_2, \cdots, X_n)$ *and* X *are independent,* $h = n^{-1/(2s+1)}$, *and that there exists* $0 < \lambda < 2$ *such that* $\int_{\mathcal{X}} |\alpha(x)|^{1-\lambda} dx < \infty$ *and*

$$\int_{\mathcal{X}} |\alpha(x)|^{1-\lambda} |u^{(1)}(x)/u(x)|^{\lambda} dx < \infty.$$

Then, as $n \to \infty$,

$$R_n - R(G) = O(n^{-\lambda(s-1)/(2s+1)}). \tag{10}$$

THEOREM 2. *Suppose that Assumption (A1, A2, A3), Assumption (A1, A2, A4) or Assumption (A1, A2, A5) is satisfied,* $(X_1, X_2, \cdots, X_n)$ *and* X *are independent,* $h = n^{-1/(2s+1)}$, *and that there exists* $0 < \lambda < 2$ *such that* $\int_{\mathcal{X}} |\alpha(x)|^{1-\lambda} dx < \infty$, $\int_{\mathcal{X}} |\alpha(x)|^{1-\lambda} \times |u^{(1)}(x)/u(x)|^{2\lambda} dx < \infty$ *and* $\int_{\mathcal{X}} |\alpha(x)|^{1-\lambda} |u^{(2)}(x)/u(x)|^{\lambda} dx < \infty$. *Then, as* $n \to \infty$,

$$R_n^* - R^*(G) = O(n^{-\lambda(s-2)/(2s+1)}). \tag{11}$$

REMARK 4. From these results, we can see that the convergence rates of the EB test rules under an associated or strong mixing sample are the same as that of the EB test rules under an independent sample. See Johns and Van Ryzin (1972) and Wei (1989). Chen and Wei (2004) only studied the case of NA samples. We note that, the convergence rates in (10) and (11) are all better than the corresponding rates, $R_n - R(G) = O(n^{-\lambda(s-1)/\{2(s+2)\}})$ and $R_n^* - R^*(G) = O(n^{-\lambda(s-2)/\{2(s+2)\}})$, in Chen and Wei (2004). In the case of NA settings, Chen and Wei (2004) also used the additional condition $\sum_{j=1}^{\infty} |Cov(X_1, X_j)| < \infty$, which is not used in this paper.

REMARK 5. Examples which satisfy the conditions in Theorems 1 and 2 can be found in Wei (1989) and Chen and Wei (2004). For example, let $f(x|\theta) = -\theta e^{\theta x}, c(\theta) = -\theta, u(x) \equiv 1, \mathcal{X} = (0, \infty), \Theta = (-\infty, 0)$. The p.d.f. of $G(\theta)$ is chosen as

$$g(\theta) = \frac{\beta^{b+1}}{\Gamma(b+1)} (-\theta)^b e^{\beta\theta}, \beta > 0, b > 0, -\infty < \theta < 0,$$

where β and b are non-random parameters. By appropriately selecting β and b, one can show that the conditions related to the population in Theorems 1 and 2 are satisfied.

3 Lemmas

We need some lemmas to prove the main results.

Lemma 1. *Suppose that $\{\eta_1, \eta_2\}$ is a pair of NA or PA random variables with zero mean and $E\eta_j^2 < \infty, j = 1, 2,$ and $g_j(\cdot), j = 1, 2,$ have bounded derivatives. Then*

$$|Cov(g_1(\eta_1), g_2(\eta_2))| \le \sup_x |g_1'(x)| \sup_y |g_2'(y)| |Cov(\eta_1, \eta_2)|,$$

where $g_j'(\cdot)$ is the derivative of $g_j(\cdot)$, $j = 1, 2$.

Proof. See Lemma 3.1 in Birkel (1988).

Lemma 2. *Let $\{\eta_i, i \ge 1\}$ be a α-mixing process and $\mathcal{F}_s^t$ denote the σ-algebra generated by $\{\eta_i, s \le i \le t\}$ for $s \le t$. Suppose that ξ and η are random variables which are $\mathcal{F}_1^k$ and $\mathcal{F}_{k+m}^\infty$ measurable, respectively, and that $||\xi||_{p_1} < \infty, ||\eta||_{q_1} < \infty$, where $p_1, q_1 > 1, p_1^{-1} + q_1^{-1} < 1$. Then,*

$$|E(\xi\eta) - E(\xi)E(\eta)| \le 10\{\alpha(m)\}^{1-p_1^{-1}-q_1^{-1}} ||\xi||_{p_1} ||\eta||_{q_1}.$$

If ξ and η are bounded random variables, then

$$|E(\xi\eta) - E(\xi)E(\eta)| \le C\alpha(m), \tag{12}$$

were C is a positive constant.

Proof. See Lemma 1 in Deo (1973).

Lemma 3. *Let $\{\eta_j : 1 \le j \le n\}$ be NA random variables with $E\eta_j = 0, E|\eta_j|^r < \infty (r > 1)$ and $\{a_j : j \ge 1\}$ be a sequence of real constants. Then there exists a positive constant C which only depends on the given number r such that*

$$E\left|\sum_{j=1}^n a_j\eta_j\right|^r \le C\sum_{j=1}^n E|a_j\eta_j|^r, \text{ for } 1 < r \le 2$$

and

$$E\left|\sum_{j=1}^n a_j\eta_j\right|^r \le C\left\{\sum_{j=1}^n E|a_j\eta_j|^r + \left(\sum_{j=1}^n E(a_j\eta_j)^2\right)^{r/2}\right\}, \text{ for } r > 2.$$

PROOF. See Su et al. (1997).

Lemma 4. *Suppose that Assumption (A1, A2, A3), Assumption (A1, A2, A4) or Assumption (A1, A2, A5) is satisfied. Then for any* $0 < \lambda \leq 2$ *and* $j = 0, 1, 2,$

$$E_n \left| f_n^{(j)}(x) - f^{(j)}(x) \right|^\lambda \leq C \left\{ h^{\lambda(s-j)} + (nh^{1+2j})^{-\lambda/2} \right\}, \tag{13}$$

where C *is a constant independent of* n *and* x. *The optimal bandwidth in* (13) *is* $h = Cn^{-1/(2s+1)}$, *and under the optimal bandwidth,* (13) *becomes*

$$E_n|f_n^{(j)}(x) - f^{(j)}(x)|^\lambda \leq Cn^{-\lambda(s-j)/(2s+1)}. \tag{14}$$

PROOF. Equation 14 is obvious. We only need to prove (13). By C_r-inequality and Jensen's inequality $E(|Y|^{\lambda/2}) \leq (E|Y|)^{\lambda/2}$ for any random variable Y and $0 < \lambda \leq 2$, we have

$$E_n|f_n^{(j)}(x) - f^{(j)}(x)|^\lambda \leq C|E_n f_n^{(j)}(x) - f^{(j)}(x)|^\lambda + C\left\{Var_n(f_n^{(j)}(x))\right\}^{\lambda/2},$$

where $Var_n(f_n^{(j)}(x)) = E_n(f_n^{(j)}(x) - E_n f_n^{(j)}(x))^2$.

It can be shown that

$$E_n f_n^{(j)}(x) = h^{-j} \int K_j(u) f(x - hu) du.$$

On the other hand, by Taylor expansion, we have

$$f(x-hu) = f(x) + \frac{1}{1!} f^{(1)}(x)(-hu) + \cdots + \frac{f^{(s-1)}(x)(-hu)^{s-1}}{(s-1)!} + \frac{f^{(s-1)}(\xi)(-hu)^s}{s!},$$

where ξ is between x and $x - hu$. These results and Assumptions A1(ii) and A2 lead to

$$\left| E_n f_n^{(j)}(x) - f^{(j)}(x) \right|^\lambda \leq Ch^{\lambda(s-j)}. \tag{15}$$

To prove (13), it suffices to show that

$$Var_n(f_n^{(j)}(x)) \leq C(nh^{1+2j})^{-1}. \tag{16}$$

Next we will prove (16) under NA, PA and strong mixing samples separately.

NA case Note that K_j is of bounded variation on any finite interval from $|K_j^{(1)}| \le C$. It follows that K_j can be written as $K_j = K_{j+} - K_{j-}$, where both K_{j+} and K_{j-} are nondecreasing functions. It follows that both $\left\{K_{j+}\left(\frac{x-X_i}{h}\right), 1 \le i \le n\right\}$ and $\left\{K_{j-}\left(\frac{x-X_i}{h}\right), 1 \le i \le n\right\}$ are NA sequences. Let

$$f_{n+}^{(j)}(x) = \frac{1}{nh^{j+1}} \sum_{i=1}^{n} K_{j+}\left(\frac{x-X_i}{h}\right), f_{n-}^{(j)}(x) = \frac{1}{nh^{j+1}} \sum_{i=1}^{n} K_{j-}\left(\frac{x-X_i}{h}\right).$$

From Lemma 3, we have $Var_n(f_{n+}^{(j)}(x)) \le Cn(nh^{j+1})^{-2} Var_n\left(K_{j+}\left(\frac{x-X_1}{h}\right)\right)$ $\le C(nh^{1+2j})^{-1}$ and $Var_n(f_{n-}^{(j)}(x)) \le C(nh^{1+2j})^{-1}$. Therefore, (16) follows under NA samples.

PA case To prove (16), it suffices to show that

$$Var_n\left\{\frac{1}{nh}\sum_{i=1}^{n} K_j\left(\frac{x-X_i}{h}\right)\right\} \le C(nh)^{-1}. \tag{17}$$

In the rest of the proof, we will use E, Var and Cov to denote E_n, Var_n and Cov_n, respectively. We employ the small-block and large-block arguments to prove (17), where p, q and k are given in Assumption A4(iii). Let $Z_{ni} = K_j((x-X_i)/h) - EK_j((x-X_i)/h), 1 \le i \le n$ and $S_n = \frac{1}{\sqrt{nh}}\sum_{i=1}^{n} Z_{ni}$. Partition S_n as

$$S_n = S_n' + S_n'' + S_n''',$$

where $S_n' = \sum_{m=1}^{k} e_{nm}$, $S_n'' = \sum_{m=1}^{k} e_{nm}'$ and $S_n''' = e_{n,k+1}'$, with $e_{nm} = \frac{1}{\sqrt{nh}}\sum_{i=r_m}^{r_m+p-1} Z_{ni}, e_{nm}' = \frac{1}{\sqrt{nh}}\sum_{i=l_m}^{l_m+q-1} Z_{ni}, e_{n,k+1}' = \frac{1}{\sqrt{nh}}\sum_{i=k(p+q)+1}^{n} Z_{ni}$, $r_m = (m-1)(p+q)+1, l_m = (m-1)(p+q)+p+1, m = 1, \cdots, k$.

Equation 17 follows if we can verify

$$Var(S_n') \le C, \tag{18}$$

$$Var(S_n'') \le C, \tag{19}$$

and

$$Var(S_n''') \le C. \tag{20}$$

It is clear, as $f(x) \le C$, that

$$h^{-1}Var(Z_{n1}) = \int K_j^2(u)f(x-hu)du - h\left\{\int K_j(u)f(x-hu)du\right\}^2 \le C.$$

Furthermore, for $1 \le i < t \le n$, from stationarity and Assumption A4 (ii), we have

$$\begin{aligned}&|Cov(Z_{ni}, Z_{nt})| = |Cov(Z_{n1}, Z_{n,t-i+1})|\\ &= h^2 \left| \int_{R^2} K_j(u)K_j(v)\{f_{1,t-i+1}(x-hu, x-hv) - f(x-hu)f(x-hv)\}dudv \right|\\ &\le Ch^2\end{aligned}$$

Using $kp/n \le 1$, assumption $ph \le C$ and above derivations, it follows that

$$\frac{k}{nh} \sum_{1 \le i < t \le p} |Cov(Z_{ni}, Z_{nt})| \le \frac{Ckp^2h}{n} \le C$$

and

$$\sum_{m=1}^{k} Var(e_{nm}) = \frac{kp}{nh} Var(Z_{n1}) + \frac{2k}{nh} \sum_{1 \le i < t \le p} Cov(Z_{ni}, Z_{nt}) \le C. \qquad (21)$$

By stationarity, it can be shown that

$$\begin{aligned}&\sum_{1 \le i < t \le k} |Cov(e_{ni}, e_{nt})| = \sum_{l=1}^{k-1} (k-l)|Cov(e_{n1}, e_{n,l+1})|\\ &\le k \sum_{l=1}^{k-1} |Cov(e_{n1}, e_{n,l+1})| \le \frac{kp}{nh} \sum_{l=1}^{k-1} \sum_{t=l(p+q)-p}^{l(p+q)+p} |Cov(Z_{n1}, Z_{n,t+1})|.\end{aligned}$$

It follows, from Lemma 1, using conditions $\frac{1}{h^3} \sum_{t=q}^{\infty} |Cov(X_1, X_{t+1})| \le C$, that

$$\sum_{1 \le i < t \le k} Cov(e_{ni}, e_{nt}) \le \frac{Ckp}{nh^3} \sum_{t=q}^{\infty} |Cov(X_1, X_{t+1})| \le C. \qquad (22)$$

Equation 18 follows from (21) and (22).

Similarly, using $q \le p$, we have

$$\begin{aligned}\sum_{m=1}^{k} Var(e'_{nm}) &= \frac{kq}{nh} Var(Z_{n1}) + \frac{2k}{nh} \sum_{1 \le i < t \le q} Cov(Z_{ni}, Z_{nt})\\ &\le \frac{Ckq}{n} + \frac{Ckq^2h}{n} \le C,\end{aligned}$$

and

$$\sum_{1\le i<t\le k}|Cov(e'_{ni},e'_{nt})| = \sum_{l=1}^{k-1}(k-l)|Cov(e'_{n1},e'_{n,l+1})|$$
$$\le k\sum_{l=1}^{k-1}|Cov(e'_{n1},e'_{n,l+1})| \le \frac{kq}{nh}\sum_{l=1}^{k-1}\sum_{r=l(p+q)-(q-1)}^{l(p+q)+(q-1)}|Cov(Z_{n1},Z_{n,r+1})|.$$

It follows that

$$E(S''_n)^2 \quad \le \quad \frac{Ckq}{n}+\frac{Ckq^2h}{n}+\frac{Ckq}{nh^3}\sum_{t=p}^{\infty}|Cov(X_1,X_{t+1})| \le C,$$

which implies (19).

We can also show that

$$E(S'''_n)^2 \le \frac{C\{n-k(p+q)\}}{n}+\frac{C\{n-k(p+q)\}^2h}{n} \le \frac{Cp}{n}+\frac{Cp^2h}{n} \le C,$$

which leads to (20). The proof of (17) is thus complete.

Strong mixing case Using the same derivations before (22) and using (12), under conditions $\frac{1}{h}\sum_{t=q}^{\infty}\alpha(t)\le C$, we can show that

$$\sum_{1\le i<t\le k}Cov(e_{ni},e_{nt}) \le \frac{Ckp}{nh}\sum_{t=q}^{\infty}\alpha(t)\le C. \tag{23}$$

Equation 18 follows from (21) and (23).

Using the same derivations as the PA case, we have

$$\sum_{m=1}^{k}Var(e'_{nm}) \le \frac{Ckq}{n}+\frac{Ckq^2h}{n}\le C,$$

and

$$\sum_{1\le i<t\le k}|Cov(e'_{ni},e'_{nt})| \le \frac{kq}{nh}\sum_{l=1}^{k-1}\sum_{r=l(p+q)-(q-1)}^{l(p+q)+(q-1)}|Cov(Z_{n1},Z_{n,r+1})|.$$

It follows, by (12), that

$$E(S''_n)^2 \le \frac{Ckq}{n}+\frac{Ckq^2h}{n}+\frac{Ckq}{nh}\sum_{t=p}^{\infty}\alpha(t)\le C,$$

which implies (19).

One can also show that

$$E(S_n''')^2 \le \frac{C\{n-k(p+q)\}}{n} + \frac{C\{n-k(p+q)\}^2 h}{n} \le \frac{Cp}{n} + \frac{Cp^2 h}{n} \le C,$$

which leads to (20). We thus have (17) from (18), (19) and (20).

Lemma 5. *For $\alpha(\cdot)$ and $\alpha_n(\cdot)$ defined as before,*

$$0 \le R_n - R(G) \le a \int_{\mathcal{X}} |\alpha(x)| P(|\alpha_n(x) - \alpha(x)| \ge |\alpha(x)|) dx.$$

PROOF. See Lemma 1 in Johns and Van Ryzin (1972).

4 Proofs of Main Results

PROOF OF THEOREM 1. For $0 < \lambda < 2$, from Lemma 5, Markov's inequality and (14) in Lemma 4, we have

$$\begin{aligned} 0 \le R_n - R(G) &\le a \int_{\mathcal{X}} |\alpha(x)| P(|\alpha_n(x) - \alpha(x)| \ge |\alpha(x)|) dx \\ &\le a \int_{\mathcal{X}} |\alpha(x)|^{1-\lambda} E_n\{|\alpha_n(x) - \alpha(x)|^\lambda\} dx \\ &\le C \int_{\mathcal{X}} |\alpha(x)|^{1-\lambda} E_n\{|f_n^{(1)}(x) - f^{(1)}(x)|^\lambda\} dx \\ &\quad + C \int_{\mathcal{X}} |\alpha(x)|^{1-\lambda} |u^{(1)}(x)/u(x) + \theta_0|^\lambda E_n\{|f_n^{(0)}(x) - f^{(0)}(x)|^\lambda\} dx \\ &\le C n^{-\lambda(s-1)/(2s+1)}, \end{aligned} \tag{24}$$

which completes the proof of Theorem 1.

PROOF OF THEOREM 2. For $0 < \lambda < 2$, similar to the proof of Theorem 1, we have

$$\begin{aligned} R_n^* - R^*(G) &\le a \int_{\mathcal{X}} |\alpha^*(x)| P(|\alpha_n^*(x) - \alpha^*(x)| \ge |\alpha^*(x)|) dx \\ &\le a \int_{\mathcal{X}} |\alpha^*(x)|^{1-\lambda} E_n\{|\alpha_n^*(x) - \alpha^*(x)|^\lambda\} dx \\ &\le C \int_{\mathcal{X}} |\alpha^*(x)|^{1-\lambda} E_n\{|f_n^{(2)}(x) - f^{(2)}(x)|^\lambda\} dx \\ &\quad + C \int_{\mathcal{X}} |\alpha(x)|^{1-\lambda} |u^{(1)}(x)/u(x) + a_0|^\lambda E_n\{|f_n^{(1)}(x) - f^{(1)}(x)|^\lambda\} dx \\ &\quad + C \int_{\mathcal{X}} |\alpha(x)|^{1-\lambda} |v(x)|^\lambda\} E_n\{|f_n^{(0)}(x) - f^{(0)}(x)|^\lambda\} dx \end{aligned}$$

$$\leq \quad Cn^{-\lambda(s-2)/(2s+1)}, \tag{25}$$

which completes the proof of Theorem 2, where $v(x) = 2\{u^{(1)}(x)/u(x)\}^2 + 2a_0 u^{(1)}(x)/u(x) - u^{(2)}(x)/u(x) + a_0^2 - \mu_0^2$.

Acknowledgment. This work was partially supported by the National Natural Science Foundation of China (11271088, 11361011) and the Natural Science Foundation of Guangxi (2013GXNSFAA019004, 2013GXNSFAA019007, 2013GXNSFBA019001). The authors are thankful to the referees for constructive suggestions.

References

BIRKEL, T. (1988). On the convergence rate in the central limit theorem for associated processes. *Ann. Probab.,* **16**, 1685–1698.

BLOCK, H.W., SAVITS, T.H. and SHARKED, M. (1982). Some concepts of negative dependence. *Ann. Probab.,* **10**, 765–772.

CHEN, L. and WEI, L.S. (2004). The EB test problems for continuous one-parameter exponential family: in the case of NA samples (in Chinese). *Mathematica Applicata.,* **39**, 44–50.

CHEN, L. and WEI, L.S. (2006). The EB estimation problem for continuous one-parameter exponential family: in the case of NA samples (in Chinese). *J. Math. Study.,* **39**, 44–50.

DEO, C.M. (1973). A note on empirical processes of strong-mixing sequences. *Ann. Probab.,* **1**, 870–875.

DOUKHAN, P. (1987). *Mixing: properties and examples.* Springer, New York.

ESARY, J.D., PROSCHAN, F. and WALKUP, D.W. (1967). Association of random variables, with applications. *Ann. Math. Statist.,* **38**, 1466–1474.

JOAG-DEV, K. and PROSCHAN, F. (1983). Negative association of random variables with applications. *Ann. Statist.,* **11**, 286–295.

JOHNS, JR., M.V. and VAN RYZIN, J.R. (1972). Convergence rates for empirical Bayes two-action problems II: continuous case. *Ann. Math. Statist.,* **43**, 934–947.

LIN, P.E. (1975). Rates of convergence in empirical Bayes estimation problems: continuous case. *Ann. Statist.,* **3**, 155–164.

RAO, B.L.S.P. (1983). *Nonparametric functional estimation.* Academic Press, New York.

ROBBINS, H. (1955) An empirical Bayes approach to statistics. *Proceedings of the 3rd Berkeley Syrup. on Math. Statist. Prob.*, Vol. 1, pp. 157–163. University California Press, Berkeley.

ROSENBLATT, M. (1956). A central limit theorem and a strong mixing condition. *Proc. Natl. Acad. Sci. U.S.A.,* **42**, 43–47.

ROUSSAS, G.G. (2000). Asymptotic normality of the kernel estimate of a probability density function under association. *Statist. Probab. Lett.,* **50**, 1–12.

SINGH, R.S. (1979). Empirical Bayes estimation in Lebesgue-exponential families with rates near the best possible rate. *Ann. Statist.,* **7**, 890–902.

SU, C., ZHAO, L.C. and WANG, Y.B. (1997). Moment inequalities and weak convergence for negatively associated sequences. *Sci. in China (Ser. A),* **40**, 172–182.

wang, h.y., zhang, t. and guo, p.j. (2008). The EB estimators of the parameter for the continuous one-parameter exponential family under PA samples (in Chinese). *J. Northwest University,* **38**, 14–18.

wei, l.s. (1989). The empirical Bayes test problems for continuous one-parameter family. *Systems Science and Mathematical Sciences,* **2**, 369–384.

Qingzhu Lei
Yongsong Qin
Department of Mathematics,
Guangxi Normal University,
Guilin, Guangxi 541004, China
E-mail: ysqin@gxnu.edu.cn

Paper received: 24 September 2013.