Ann. Inst. Statist. Math.
Vol. 44, No. 3, 435–449 (1992)

EMPIRICAL BAYES WITH RATES AND BEST RATES OF CONVERGENCE IN $u(x)C(\theta)\exp(-x/\theta)$-FAMILY: ESTIMATION CASE* **

R. S. SINGH[1] AND LAISHENG WEI[2]

[1] *Department of Mathematics and Statistics, University of Guelph, Guelph, Ontario, Canada N1G 2W1*
[2] *Department of Mathematics, University of Science and Technology of China, Hefei, Anhui, P. R. China*

(Received May 30, 1990; revised June 24, 1991)

Abstract. Let $\{(X_i, \theta_i)\}$ be a sequence of independent random vectors where X_i, conditional on θ_i, has the probability density of the form $f(x \mid \theta_i) = u(x)C(\theta_i)\exp(-x/\theta_i)$ and the unobservable θ_i are i.i.d. according to an unknown G in some class $\mathcal{G}$ of prior distributions on Θ, a subset of $\{\theta > 0 \mid C(\theta) = \left(\int u(x)\exp(-x/\theta)dx\right)^{-1} > 0\}$. For a $\mathcal{S}(X_1, \dots, X_n, X_{n+1})$-measurable function ϕ_n, let $R_n = E(\phi_n - \theta_{n+1})^2$ denote the Bayes risk of ϕ_n and let $R(G)$ denote the infimum Bayes risk with respect to G. For each integer $s > 1$ we exhibit a class of $\mathcal{S}(X_1, \dots, X_n, X_{n+1})$-measurable functions ϕ_n such that for δ in $[s^{-1}, 1]$, $c_0 n^{-2s/(1+2s)} \le R_n(\phi_n, G) - R(G) \le c_1 n^{-2(s\delta-1)/(1+2s)}$ under certain conditions on u and G. No assumptions on the form or smoothness of u is made, however. Examples of functions u, including one with infinitely many discontinuities, are given for which our conditions reduce to some moment conditions on G. When Θ is bounded, for each integer $s > 1$ $\mathcal{S}(X_1, \dots, X_n, X_{n+1})$-measurable functions ϕ_n are exhibited such that for δ in $[2/s, 1]$, $c_0' n^{-2s/(1+2s)} \le R(\phi_n, G) - R(G) \le c_1' n^{-2s\delta/(1+2s)}$. Examples of functions u and class $\mathcal{G}$ are given where the above lower and upper bounds are achieved.

Key words and phrases: Exponential family, empirical Bayes, estimation, asymptotic optimality, rates and best rates.

* Part of the research was carried out during R. S. Singh's visit to the University of Science and Technology of China.

** Research supported in part by a Natural Sciences and Engineering Research Council of Canada Grant No. #A4631.

1. Introduction, notation and preliminaries

In empirical Bayes (EB) context (as introduced by Robbins (1955) and later developed in great detail by Johns (1957), Robbins (1963, 1964), Samuel (1963), Johns and Van Ryzin (1971, 1972), among others), one considers a sequence of statistical problems having the same generic structure being possessed by what is called the component problem. In the component problem there are a parameter space Θ, a measurable space $(\Theta, \mathcal{F})$, where $\mathcal{F}$ is the Borel σ-field of subsets of Θ, a set $\mathcal{G}$ of all prior distributions on $(\Theta, \mathcal{F})$, a family of probability measures $\mathcal{P} = \{P(\cdot \mid \theta) \mid \theta \in \Theta\}$ over a measurable space $(\mathcal{X}, \mathcal{B})$, where $\mathcal{B}$ is the Borel σ-field of subsets of $\mathcal{X}$, and there are an action space $\mathcal{A}$ and a loss function $L \geq 0$ on $\mathcal{A} \times \Theta$. A $(\mathcal{X}, \mathcal{B})$-measurable function ϕ into $\mathcal{A}$, which decides about a θ in Θ based on a random observation X from $P(\cdot \mid \theta)$, results in a $(\Theta, \mathcal{F})$-measurable expected loss (risk) function $R(\phi, \theta) = \int L(\phi(x), \theta) dP(x \mid \theta)$ and an overall risk (Bayes risk) against a G in $\mathcal{G}$, $R(\phi, G) = \int R(\phi, \theta) dG(\theta)$. For a G in $\mathcal{G}$, a procedure $\phi = \phi_G$ which achieve the *Bayes envelope* against G, $R(G) = \min_\phi R(\phi, G)$, whenever the latter exists, is called *Bayes optimal* procedure versus G.

In an EB context, G remains unknown at every stage in the sequence (and therefore the optimal procedure ϕ_G is not available for use at any stage), and at the $(n+1)$-st stage random observations $X_1, \ldots, X_n$ from the previous n stages and $X_{n+1} = X$ from the present stage, where $X_i \sim P(\cdot \mid \theta_i)$ and $\theta_1, \theta_2, \ldots$ i.i.d. $\sim G$, are used to exhibit a $\phi_n(X) = \phi_n(X_1, \ldots, X_n, X)$, where ϕ_n is a $(\mathcal{X}^{n+1}, \mathcal{B}^{n+1})$-measurable function into $\mathcal{A}$, such that ϕ_n is, in Bayes risk, as good as ϕ_G at least for large n. Such a function, called EB procedure, is said to be asymptotically optimal (a.o.) or a.o. with a rate $n^{-\gamma}$ for some $\gamma > 0$ whenever $R(\phi_n, G) - R(G) = o(1)$ or $O(n^{-\gamma})$ as $n \to \infty$. We assume that in the component problem there is a σ-finite measure μ on $(\mathcal{X}, \mathcal{B})$ which dominates each $P(\cdot \mid \theta)$, $\theta \in \Theta$, and is dominated by the Lebesgue measure on $(\mathcal{X}, \mathcal{B})$. Let $p(x \mid \theta)$ and $u(x)$ denote the Radon-Nikodym derivatives $dP(x \mid \theta)/d\mu(x)$ and $d\mu(x)/dx$ respectively.

Empirical Bayes estimation problem where Θ, $\mathcal{A}$ and $\mathcal{X}$ are subsets of the real line, $L(\theta, a) = (\theta - a)^2$ and $p(\cdot \mid \theta)$ is of the form $p(x \mid \theta) = C(\theta)\exp(\theta x)$, (i.e. the EB squared error loss estimation (SELE) for the usual Lebesgue exponential family $f(x \mid \theta) = u(x)C(\theta)\exp(\theta x)$), has been considered by Maritz (1969), Hannan and Macky (1971), Yu (1971), Maritz and Lwin (1975), Lin (1975), O'Bryan and Susarla (1976) and Singh (1976, 1979), among others. EB SELE involving some non-regular family of densities has been considered, among others, by Fox (1978), Prasad and Singh (1990) and Singh and Prasad (1991). In this paper we consider SELE in an equally important exponential family, but has received very little attention, if at all, in the EB context, namely the family, with $p(\cdot \mid \theta)$ given by $p(x \mid \theta) = C(\theta)\exp(-x/\theta)$ with Θ being a subset of the natural parameter space $\{\theta > 0 \mid \left(\int \exp(-x/\theta) d\mu(x)\right)^{-1} > 0\}$, and $\mathcal{X}$ and $\mathcal{A}$ being subsets of the real line. This family includes the usual simple scale exponential density $f(x \mid \theta) = u(x)p(x \mid \theta) = \theta^{-1}\exp(-x/\theta)I(x > 0)$, and the gamma density $f(x \mid \theta) = (\Gamma(\gamma))^{-1}x^{\gamma-1}\theta^{-\gamma}\exp(-x \mid \theta)I(x > 0)$ (most widely used for statistical modellings in engineering, medical sciences, demography, reliability theory and survival analysis). The $N(0, \theta/2)$-density with $X = Y^2$ where

$Y \sim N(0, \theta/2)$ also falls in this category. For integers $s > 0$ we exhibit EB estimators $\phi_n = \phi_n(X_1, \ldots, X_n, X)$ such that for δ in $[s^{-1}, 1]$,

$$cn^{-2s/(1+2s)} \leq R(\phi_n, G) - R(G) \leq c'n^{-2(\delta s-1)/(1+2s)}, \tag{1.1}$$

where c and c' are positive constants, the r.h.s. inequality holds for $n > 1$ uniformly in G satisfying certain conditions, and the l.h.s. inequality holds at each degenerate G and for all n sufficiently large. If Θ is bounded, then for integers $s > 0$ we have estimators ϕ_n such that for δ in $[s^{-1}, 1]$,

$$cn^{-2s/(1+2s)} \leq R(\phi_n, G) - R(G) \leq c'n^{-2\delta s/(1+2s)}. \tag{1.1$'$}$$

We have made no assumption on the form or on smoothness, of whatsoever nature, on u. Thus ϕ_n are exhibited with rates arbitrarily close to $O(n^{-1})$. *As noted in Singh* (1976, 1979), *a rate $O(n^{-1})$ has not yet been established in any EB problem involving a Lebesgue density whatever may be G.* However, a rate of the order $O(\exp(-cn))$ for some constant $c > 0$ is established by Liang (1988) for a two-action linear loss hypothesis testing problem involving *discrete* exponential family.

In Section 2 we give a class of EB estimators ϕ_n. In Sections 3 and 5 we obtain respectively the upper and lower bounds in (1.1). In Section 4 we give examples of some important families, including one with u having infinitely many discontinuities, where conditions for (1.1) and (1.1)$'$ are satisfied. We conclude the paper with a few remarks and some further results in Section 5.

2. Proposed class of EB estimators ϕ_n

2.1 *Introduction and reduction of problem*

Throughout the remainder of this paper, we assume that $\mathcal{X}$ and $\mathcal{A}$ are subsets of the nonnegative real line, Θ is a subset of $\{\theta > 0 \mid C(\theta) = \left(\int \exp(-x/\theta)d\mu \cdot (x)\right)^{-1} > 0\}$, $L(\theta, a) = (\theta - a)^2$, $p(x \mid \theta) = C(\theta)\exp(-x/\theta)$, $u(x) = d\mu(x)/dx$, $f(x \mid \theta) = u(x)p(x \mid \theta)$ and $\mathcal{G}$ is the class of all probability measures G on $(\Theta, \mathcal{F})$ for which ϕ_G exists and $R(G) < \infty$. Since in the EB context, $\theta_1, \ldots, \theta_n, \theta_{n+1} = \theta$ are i.i.d. with a common distribution G in $\mathcal{G}$, and $X_1, \ldots, X_n, X_{n+1} = X$ are i.i.d. with common marginal p.d.f. $f(x) = \int f(x \mid \theta)dG(\theta)$, the Bayes optimal estimator versus G in the EB context at $(n+1)$-st stage, $n \geq 1$, remains the same as in any one component problem, i.e. $\phi_G(X_1, \ldots, X_n, X) = E(\theta \mid X_1, \ldots, X_n; X) = E(\theta \mid X) = \phi_G(X)$ for $n \geq 1$. Further, with $p(\cdot) = \int p(\cdot \mid \theta)dG(\theta)$ and $f(\cdot) = \int f(\cdot \mid \theta)dG(\theta)$, it is easy to see that

$$\phi_G(X) = E(\theta \mid X) = \psi(X)/p(X) \tag{2.1}$$

with ψ defined by

$$\psi(x) = \int_x^\infty p, \tag{2.2}$$

where E stands for the expectation operator, (here and wherever possible) the argument of integration is indicated by omission and the ratio 0/0 is defined to be 0. Further, it is well known (e.g. see Singh (1979), Lemma 2.1) that if ϕ_G exists and $R(G) < \infty$, then for any $\mathcal{S}(X_1, \ldots, X_n, X)$-measurable mapping $\phi^* = \phi^*(X_1, \ldots, X_n, X)$ into $\mathcal{A}$,

$$R(\phi^*, G) - R(G) = E(\phi^* - \phi_G)^2. \tag{2.3}$$

It may be noted that $E(\theta^2) < \infty$ is sufficient for the finiteness of $R(G)$ and for (2.3), but not necessary (e.g. see Remark 2.1 of Singh (1979)).

Thus, in view of (2.3), the search for an a.o. estimator reduces to the search for a mean square consistent estimator of ϕ_G, which, in view of (2.1), can be achieved if p and ψ are estimated appropriately.

2.2 *Estimation of ψ and p*

For an event A, let $I(A)$ denote the indicator function of A. Since $\psi(x) = E[I(X \geq x, u(X) > 0)/u(X)]$, we propose to estimate ψ by its natural estimator

$$\psi_n(x) = n^{-1} \sum_{j=1}^{n} [I(X_j \geq x, u(X_j) > 0)/u(X_j)]. \tag{2.4}$$

To estimate the μ-density p, we take the estimators for a μ-density considered in Singh (1974, 1978*b*, 1981), which we re-introduce briefly for the convenience of our readers as well as to facilitate our proofs here. Let $s > 0$ be a fixed integer and let $\mathcal{K}_s$ be the class of all Borel-measurable real valued bounded function functions K vanishing off $(0, 1)$ such that $\int y^i K(y) = 1$ or 0 according as $i = 0$ or $1, \ldots, s-1$. (For examples of such functions, see Singh (1977*a*, 1979).) Then for K in $\mathcal{K}_s$, the estimators of p from Singh (1974, 1978*b*) are given by

$$p_n(x) = (nh)^{-1} \sum_{j=1}^{n} \left[K\left(\frac{X_j - x}{n}\right) I(u(X_j) > 0)/u(X_j) \right] \tag{2.5}$$

where $0 < h = h(n) \downarrow 0$ as $n \to \infty$ is the usual window-width function used in kernel type density estimators. It follows from Singh's general results applied to our special density p that p_n are mean square as well as strongly consistent estimators for p. Though speed of convergence for the mean square consistency of p_n can be obtained from Singh (1974, 1978*b*) but, since we are dealing with a very special density $p(x) = \int C(\theta) \exp(-x/\theta) dG(\theta)$, we can give a more specific bound for the mean square errors of p_n in Theorem 2.1 below. Let M denote the bound for K in p_n, and for an $\epsilon > 0$ let

$$\begin{aligned} &u_\epsilon(x) = \inf\{u(t) \mid x < t < x + \epsilon\}, \qquad \nu_\epsilon(x) = p(x)/u_\epsilon(x), \\ &p_\epsilon^{(s)}(x) = \sup\{p^{(s)}(t) \mid x < t < x + \epsilon\} \qquad \text{and} \\ &w(x) = \int_x^\infty (p(y)/u(y)) dy. \end{aligned} \tag{2.6}$$

Remark 2.1. Since $p(x)$ is a decreasing function of x, from Theorem 2.9 of Lehmann (1959), $p_\epsilon^{(s)}(x) \leq E(\theta^{-s} \mid X = x) p(x)$ for all $\epsilon > 0$.

THEOREM 2.1. *For every* $0 < t \leq 2$,

$$E|\psi_n - \psi|^t \leq (n^{-1}w)^{t/2} \quad \text{and} \tag{2.7}$$

$$E|p_n - p|^t \leq 2^{t-1}M^t[\{h^s p_h^{(s)}\}^t + \{n^{-1}h^{-1}v_h\}^{t/2}] \tag{2.8}$$

where the argument x *in* ψ_n, p_n, ψ, p, w, p_h *and* v_h *is indicated by omission.*

PROOF. Since ψ_n is an unbiased estimator of ψ and $\operatorname{var}(\psi_n) = n^{-1}\operatorname{var}[I(X_1 \geq x, u(X_1) > 0)/u(X_1)]$ is bounded by $n^{-1}\int_x^\infty (1/u^2(y))f(y)dy = n^{-1}w(x)$, (2.7) follows from Hölder inequality. Further from the arguments used for (3.6) of Singh (1977*a*) it follows that

$$|Ep_n - p| \leq Mh^s p_h^{(s)}. \tag{2.9}$$

Also, since $X_1, \ldots, X_n$ are i.i.d. with p.d.f. $f = up$,

$$\begin{aligned} \operatorname{var}(p_n) &= (nh^2)^{-1}\operatorname{var}\left[K\left(\frac{X_1 - x}{h}\right) I(u(X_1) > 0)/u(X_1)\right] \\ &\leq n^{-1}h^{-1}\int K^2(y)(p(x+hy)/u(x+hy))dy \\ &\leq n^{-1}h^{-1}M^2 v_h \end{aligned} \tag{2.10}$$

since $p(\cdot)$ is a decreasing function. The proof of (2.8) is complete by c_r-inequality followed by (2.9), Hölder inequality and (2.10). □

Taking $t = 2$ in Theorem 2.1, we see that ψ_n is mean square error (and hence in probability) consistent estimator of ψ, and if $h \to 0$ and $nh \to \infty$ as $n \to \infty$, then same holds with p_n as an estimator of p.

Throughout the remainder of this paper we take $s > 1$ a fixed integer, and for a positive constant $0 < c_0 \leq 1$, we take $h = c_0 n^{-1/(1+2s)}$.

2.3 *Empirical Bayes estimators* ϕ_n

Since ψ_n and p_n are consistent estimators of ψ and p respectively, from (2.1) it is natural to estimate ϕ_G by ψ_n/p_n. However, to ensure the existence of our estimator, our proposed EB estimator in the $(n+1)$-st component problem is given by

$$\phi_n(X) = [\psi_n(X)/p_n(X)]_{c_1 h^{-1}} \tag{2.11}$$

where $c_1 > 0$ is an arbitrary finite constant and $[a]_L$ is defined as $-L$, a or L depending on whether $a < -L$, $|a| \leq L$ or $a > L$.

3. An upper bound for $[R(\phi_n, G) - R(G)]$ and rates of asymptotic optimality

3.1 *The main result*

In Theorem 3.1 below we obtain an upper bound for $R(\phi_n, G) - R(G)$ which holds uniformly in $G \in \mathcal{G}$ satisfying certain conditions. First we state without proof the following lemma due to Singh (1974, 1977*b*), which facilitates the proof of Theorem 3.1.

LEMMA 3.1. *For every pair (Y, Y') of random variables and for real numbers $y \neq 0$, y', $0 < L < \infty$ and $0 < \gamma \leq 2$,*

$$E\left(\left|\frac{Y'}{Y} - \frac{y'}{y}\right| \Lambda L\right)^{\gamma} \leq 2|y|^{-\gamma}\left\{E|Y' - y'|^{\gamma} + \left(\left|\frac{y'}{y}\right| + L\right)^{\gamma} E|Y - y|^{\gamma}\right\}. \tag{3.1}$$

THEOREM 3.1. *For an integer $s \geq 1$ and for some λ in $[(2/s), 2]$ and a finite k_0, let $\mathcal{G}$ be the class of all probability measures G on $(\Theta, \mathcal{F})$ such that*

$$\int \theta^{\lambda s} dG(\theta) \leq k_0, \tag{A0}$$

$$\int \left(\int_x^{\infty} (p/u)\right)^{\lambda/2} p^{1-\lambda}(x) u(x) dx \leq k_0, \tag{A1}$$

$$\int p^{1-\lambda}(x)(p_{\epsilon}^{(s)}(x))^{\lambda} u(x) dx \leq k_0 \quad \textit{for some} \quad \epsilon > 0 \quad \textit{and} \tag{A2}$$

$$\int u_{\epsilon}^{-\lambda-2}(x) u(x) p^{1-(\lambda/2)}(x) dx \leq k_0 \quad \textit{for some} \quad \epsilon > 0, \tag{A3}$$

then there exists a finite constant c' such that

$$\sup_{G \in \mathcal{G}} \{R(\phi_n, G) - R(G)\} \leq c' n^{-(\lambda s - 2)/(1+2s)} \tag{3.2}$$

for each $n \geq 1$.

Remark 3.1. Since Theorem 3.1 deals with a general Lebesgue exponential family $u(x)C(\theta)\exp(-x/\theta)I(x > 0)$, with no assumptions whatsoever on the form or smoothness of u, it is difficult to see how (A1), (A2) or (A3) relate to any moment condition on G. We conjecture that no specific moment conditions on G can be provided underwhich results of the sort (3.2) hold for the general function u. We will see later that for some most widely used exponential families, (A0)–(A3) reduce simply to some moment conditions on G.

PROOF OF THEOREM 3.1. With c_1 as in (2.11), let $A_n = c_1 h^{-1}$. It is easy to see that (2.1) and (2.11) followed by (3.1), (2.7) and (2.8) give, for $0 < \lambda \leq 2$,

$$\begin{aligned} &E|\phi_n - \phi_G|^2 I(\phi_G \leq A_n) \\ &\leq E\left(\left|\frac{\phi_n}{p_n} - \frac{\psi}{p}\right| \Lambda A_n\right)^2 I(\psi \leq A_n p) \\ &\leq 2A_n^{2-\lambda} p^{-\lambda}[(n^{-1}w)^{\lambda/2} \\ &\qquad + (2A_n)^{\lambda} 2^{\lambda-1} M^{\lambda}\{(h^s p_h^{(s)})^{\lambda} + (n^{-1}h^{-1}v_h)^{\lambda/2}\}] \\ &= c_2 n^{-(\lambda s-2)/(1+2s)}[p^{-\lambda}\{w^{\lambda/2} + (p_h^{(s)})^{\lambda} + v_h^{\lambda/2}\}] \end{aligned} \tag{3.3}$$

since $A_n = c_1 h^{-1}$ and $h = c_0 n^{-1/(1+2s)}$. Consequently, for $0 < \lambda \le 2$,

$$E|\phi_n(X) - \phi_G(X)|^2 I(\phi_G(X) \le A_n) \\ = c_2 n^{-(\lambda s-2)/(1+2s)}[E\{p^{-\lambda}(X)w^{\lambda/2}(X)\} \\ + E\{p^{-\lambda}(X)(p_h^{(s)}(X))^{\lambda}\} + E\{p^{-\lambda}(X)(v_h(X))^{\lambda/2}\}]. \tag{3.4}$$

Now, since X has the unconditional p.d.f. $up = f$, it can be easily verified that the quantity in the square bracket on the r.h.s. of (3.4) reduces to sum of the three integrals in (A1), (A2) and (A3) above by choosing $c_0 < \epsilon$. Hence, from our hypotheses (A1)–(A3), the l.h.s. of (3.3) is no more than $c_3 n^{-(\lambda s-2)/(1+2s)}$ for a finite constant c_3.

Now, since $\phi_n \le A_n$, Hölder inequality followed by Markov inequality gives

$$E|\phi_n(X) - \phi_G(X)|^2 I(\phi_G(X) > A_n) \\ \le E\phi_G^2(X) I(\phi_G(X) > A_n) \\ \le (E(\phi_G(X))^{2t_1})^{1/t_1}\{E(\phi_G(X))^{(\lambda s-2)t_2}\}^{1/t_2} A_n^{-(\lambda s-2)} \tag{3.5}$$

where $t_1 > 1$, $t_2 > 1$ are such that $t_1^{-1} + t_2^{-1} = 1$. Taking $\lambda > (2/s)$ and $t_1 = \lambda s/2$ in the above inequality and noting that $A_n = c_1 h^{-1}$, we see from (A0) that the r.h.s. of (3.5) is no more than $c_3' n^{-(\lambda s-2)/(1+2s)}$ for some finite constant c_3'. Thus we concluded that

$$E|\phi_n(X) - \phi_G(X)|^2 \le c' n^{-(\lambda s-2)/(1+2s)} \tag{3.6}$$

for some finite constant c'. Now the proof of theorem is complete from the identity (2.3). □

3.2 *Examples and subtheorems*

We now give examples of some important exponential families where all conditions of Theorem 3.2 reduce to some moment conditions on G. We first prove the following lemma.

LEMMA 3.2. *For all $k > 0$ and for all $(1+k)^{-1} < t < 1$,*

$$\left(\int_0^\infty f^t\right) \le (EX^k)^t + 1. \tag{3.7}$$

PROOF. Writing $\int_0^\infty f^t$ as $I_1 + I_2$, where $I_1 = \int_0^1 f^t$ and $I_2 = \int_1^\infty f^t$, we see by Hölder inequality that $I_1 \le \left(\int_0^1 f\right)^t \le 1$ and

$$I_2 = \int_1^\infty x^{-kt}(x^k f)^t \le \left(\int_1^\infty x^k f\right)^t \le (EX^k)^t \text{ since } 1 > t > (1+k)^{-1}. \quad \square$$

3.2.1.

Example 3.1. Our first example is a simple, but quite widely applied, exponential family, namely, $f(x \mid \theta) = \theta^{-1}e^{-x/\theta}I(x > 0, \theta > 0)$ where the conditions of Theorem 3.1 simply reduce to two moment conditions on G.

SUBTHEOREM 3.1. *Let $\mathcal{G}$ be the class of all probability measures G on $(\Theta, \mathcal{F})$, where Θ is a subset of $(0, \infty)$, such that for some integer $s > 1$, $(2/s) < \lambda < 2(1 - s^{-1})$ and a finite k_0,*

$$\int \theta^{\lambda s} dG(\theta) \leq k_0 \quad \text{and} \quad \int \theta^{-2s-1} dG(\theta) \leq k_0. \tag{3.8}$$

If the σ-finite measure μ on $(\mathcal{X}, \mathcal{B})$, where $\mathcal{X} = (0, \infty)$, is itself the Lebesgue measure on $\mathcal{X}$, then

$$\sup_{G \in \mathcal{G}} (R(\phi_n, G) - R(G)) \leq k_1 n^{-(\lambda s - 2)/(1+2s)} \tag{3.9}$$

for some finite constant k_1.

PROOF. Notice that if μ is the Lebesgue measure on $(0, \infty)$, then $d\mu(x) = dx$, $u(x) \equiv 1$, $C(\theta) = \left(\int_0^\infty \exp(-x/\theta)dx\right)^{-1} = \theta^{-1}$ and for θ in Θ,

$$f(x \mid \theta) = p(x \mid \theta) = \theta^{-1} \exp(-x/\theta) I(x > 0). \tag{3.10}$$

Hence $f(x) = p(x) = \int p(x \mid \theta) dG(\theta)$. Now, since $\int_x^\infty p = \int \exp(-x/\theta) dG(\theta)$, by Hölder inequality, for $\alpha = \lambda s/(1 + \lambda s)$,

$$\int_x^\infty p \leq p^\alpha(x) \left(\int \theta^{\alpha/(1-\alpha)} \exp(-x/\theta) dG(\theta) \right)^{1-\alpha} \leq p^\alpha(x) (E\theta^{\lambda s})^{1-\alpha}.$$

Therefore, in view of (3.8) the integral in (A1) is bounded by a constant times

$$\int_0^\infty p^{1-\lambda(1-\alpha/2)} = \int_0^\infty p^t \tag{3.11}$$

where $t = 1 - \lambda(1 - \alpha/2) = 1 - \lambda(\lambda s + 2)(2 + 2\lambda s)^{-1}$. Since $1 < t^{-1} < (1 + \lambda s)$ from the condition on λ and the identity $EX^{\lambda s} = O(E\theta^{\lambda s})$, the integral in (3.11) is finite by Lemma 3.2. Also, since

$$p_\epsilon^{(s)}(x) = \sup_{x < t < x + \epsilon} |p^{(s)}(t)| = \int \theta^{-1-s} \exp(-x/\theta) dG(\theta),$$

by Hölder inequality

$$p_\epsilon^{(s)}(x) = \left(\int \theta^{-(2s+1)} dG(\theta) \right)^{1/2} f^{1/2}(x).$$

Therefore in view of (3.8) to see (A2) holds, it suffices to show that $\int p^{1-\lambda/2} < \infty$. But this as well as (A3) follows from Lemma 3.2 since $1 < \lambda < 2 - (2/s)$ and $EX^{\lambda s} = O(E\theta^{\lambda s})$. Hence by Theorem 3.1 we get (3.9). □

3.2.2.

Example 3.2. Here we give an example of a scale exponential family which has a variety of applications, particularly in survival analysis, reliability theory, quality control, and statistical modellings in demographic and medical sciences.

SUBTHEOREM 3.2. *Let the σ-finite measure μ on $(\mathcal{X}, \mathcal{B})$, where $\mathcal{X} = (0, \infty)$, be such that the Randon-Nikodym derivation of μ with respect to Lebesgue measure on $\mathcal{X}$ is $x^{\gamma-1}I(x > 0)$, for some $\gamma > 1$, let $\mathcal{G}$ be the class of all probability measures G on $(\Theta, \mathcal{F})$, where Θ is a subset of $(0, \infty)$, such that for some integer $s > 1$, $(2/s) < \lambda < 2(1 - s^{-1})$ and finite k_0,*

$$\int \theta^{\lambda s} dG(\theta) \le k_0, \qquad \int \theta^{-\gamma+1+\lambda s} dG(\theta) \le k_0$$

$$\int \theta^{-\gamma(s+1)} dG(\theta) \le k_0 \quad \textit{if} \;\; 1 < \gamma \le 2, \qquad \textit{and} \tag{3.12}$$

$$\int \theta^{-\gamma(1+s)+s} dG(\theta) \le k_0 \quad \textit{if} \;\; \gamma > 2.$$

Then

$$\sup_{G \in \mathcal{G}} (R(\phi_n, G) - R(G)) \le k_1 n^{-(\lambda s - 2)/(1+2s)} \qquad \textit{for some constant } k_1. \tag{3.13}$$

PROOF. Since $d\mu(x)/dx$ is $x^{\gamma-1}$ for some $\gamma > 1$, $u(x) = x^{\gamma-1}I(x > 0)$, $C(\theta) = \left(\int_0^\infty x^{\gamma-1}\exp(-x/\theta)dx\right)^{-1} = \theta^{-\gamma}/\Gamma(\gamma)$ and

$$f(x \mid \theta) = \frac{1}{\Gamma(\gamma)} x^{\gamma-1}\theta^{-\gamma}\exp(-x/\theta)I(x > 0) \qquad \text{for} \quad \theta \in \Theta. \tag{3.14}$$

Thus $u_\epsilon(x) = \inf_{x<t<x+\epsilon} u(t) = x^{\gamma-1}$ and, since $p(x) = \int p(x \mid \theta)dG(\theta)$, by Theorem 2.9 of Lehmann (1959),

$$\begin{aligned} p_\epsilon^{(s)}(x) &= \sup_{x<t<x+\epsilon} |p^{(s)}(t)| \\ &= \int \theta^{-\gamma-s}\exp(-x \mid \theta)dG(\theta) \\ &\le p^\alpha(x)E(\theta^{(-\gamma-s)/(1-\alpha)}), \qquad \text{for} \quad 0 < \alpha < 1, \end{aligned} \tag{3.15}$$

by Hölder inequality. Later we will take α in (3.15) equal to $(\gamma - 1)/\gamma$ or $(\gamma - 2)/(\gamma - 1)$ depending on whether $1 < \gamma \le 2$ or $\gamma > 2$.

Now consider

$$\int_x^\infty (p/u) = \int_x^\infty y^{1-\gamma}\left(\int \theta^{-\gamma}\exp(-y/\theta)dG(\theta)\right)dy. \tag{3.16}$$

Now for $1 < \gamma \le 2$, (3.16) is no more than

$$\begin{aligned} x^{1-\gamma}&\int \theta^{1-\gamma}e^{-x/\theta}dG(\theta) \\ &\le x^{1-\gamma}p^{\alpha'}(x)(E(\theta^{-\gamma+(1-\alpha')^{-1}}))^{1-\alpha'} \quad \text{for} \quad 0<\alpha'<1, \end{aligned}$$

by Hölder inequality. Therefore, from (3.12) with $\alpha' = \lambda s/(1+\lambda s)$, the integral in (A1) for $1<\gamma\le 2$ is a constant times

$$\int f^{1-\lambda(1-\alpha'/2)}u^{\lambda(1-\alpha')/2} \le I_1 + I_2, \tag{3.17}$$

where I_1 and I_2 are integrals over $\{x\le 1\}$ and $\{x>1\}$ respectively. From Lemma 3.2 $I_1 \le \int f^{1-\gamma+\lambda\alpha'/2} < \infty$ since with $t = 1-\lambda+(\lambda\alpha'/2) = 1-\lambda(\lambda s+2)(2+2\lambda s)^{-1}$, $1 < t^{-1} < 1+\lambda s$ from the condition on λ that $1<\lambda<2-(2/s)$, and $EX^{\lambda s} = O(E\theta^{\lambda s}) < \infty$. Since with this t and with $1<\gamma\le 2$, $(\gamma-1)\lambda(1-\alpha') \le 2\lambda/(2+2\lambda s) \le \lambda t$, $u^{\lambda(1-\alpha')/2} \le x^{\lambda t/2}$ for $x \ge 1$. Therefore, for $(2s)^{-1} \le \eta < \{t(\lambda s+1)-1\}/(t\lambda s)$,

$$\begin{aligned} I_2 &\le \int_1^\infty (x^{\lambda t/2}f^t) \le \int (x^{\lambda t/2}f^{\eta t}f^{t(1-\eta)}) \\ &\le (EX^{\lambda/2\eta})^{\eta t}\left(\int f^{t(1-\eta)/(1-\eta t)}\right)^{1-\eta t} \end{aligned}$$

which is finite by Lemma 3.2 since $(1+\lambda s)^{-1} < t(1-\eta)(1-\eta t)^{-1} < 1$ and $EX^{\lambda 2\eta} = O(E\theta^{\lambda/2\eta}) < \infty$ (by 3.12) since $(\lambda/2\eta) \le \lambda s$.

Now for $\gamma > 2$, $\int_x^\infty (p/u) \le \int_x^\infty y^{1-\gamma}\int\theta^{-\gamma}\exp(-y/\theta)dG(\theta)dy \le p(x)x^{2-\gamma}$, and the integral in (A1) is

$$\int f^{1-(\lambda/2)}(x)\cdot x^{\lambda/2}dx = I_1 + I_2 \tag{3.18}$$

where I_1 and I_2 are integrals over $\{x\le 1\}$ and $\{x>1\}$ respectively. It is easy to see that $I_1 \le 1$ and by Hölder inequality,

$$I_2 \le \left(\int_1^\infty x^{-(1+\xi)}\right)^{\lambda/2}\left(\int_1^\infty x^{(2+\xi)\lambda/(2-\lambda)}f(x)\right)^{1-\lambda/2}$$

which is finite from (3.12) for a choice of $\xi > 0$ such that $(2+\xi)\lambda/(2-\lambda) \le \lambda s$, since $EX^{\lambda s} < \infty$ whenever $E(\theta^{\lambda s}) < \infty$.

Now to show that (A2) holds, we notice from (3.15) and (3.12) that the integral in (A2) is no more than a constant times

$$\int p^{1-\lambda(1-\alpha)}u = \int f^{1-\lambda(1-\alpha)}u^{\lambda(1-\alpha)} = \int x^{\lambda(\gamma-1)(1-\alpha)}f^{1-\lambda(1-\alpha)}. \tag{3.19}$$

Taking $\alpha = (\gamma-1)/\gamma$ or $(\gamma-2)/(\gamma-1)$ depending on whether $1 < \gamma \le 2$ or $\gamma > 2$, the finiteness of the integral on the r.h.s. of (3.19) can be shown by arguments similar to those used to show the finiteness of (3.18). Also, since $u_\epsilon(x) = u(x)$, the integral in (A3) is $\int f^{1-\lambda/2}$, which is finite by Lemma 3.2. Hence by Theorem 3.1 we get (3.13). □

3.2.3.

Example 3.3. Our last but not the least example will be a noncontinuous scale exponential family. In fact, the family we are going to consider now is similar to the one considered in Singh (1979) and has infinitely many discontinuities. Consider the σ-finite measure μ on $(\mathcal{X}, \mathcal{B})$ with $\mathcal{X} = (0, \infty)$ whose Radon-Nikodym derivative u w.r.t. the Lebesgue measure on $\mathcal{X}$ is given by

$$u(x) = \sum_{i=0}^{\infty} (i+1) I(i < x \le i+1). \tag{3.20}$$

Notice that $C(\theta) = \left(\int_0^\infty u(x)\exp(-x/\theta)dx\right)^{-1}$ is $(1-\exp(-1/\theta))/\theta$, and the family of the conditional (on θ) p.d.f. in the component problem is

$$f(x|\theta) = (1 - \exp(-1/\theta))\theta^{-1} \left(\sum_{i=0}^{\infty}(i+1)I(i < x \le i+1)\right) \exp(-x/\theta). \tag{3.21}$$

SUBTHEOREM 3.3. *For an integer $s > 1$ and for $1 < \lambda < 2 - (2/s)$, let $\mathcal{G}$ be the class of all probability measures G on $(\Theta, \mathcal{F})$ with Θ a subset of $(0, \infty)$ such that $\int \theta^{\lambda s} dG(\theta) < \infty$ and $\int \theta^{-2s-1} dG(\theta) < \infty$. Then for the σ-finite measure μ on $\mathcal{X}$ with $d\mu(x)/dx$ given by* (3.20),

$$\sup_{G \in \mathcal{G}} (R(\phi_n, G) - R(G)) = O(n^{-(\lambda s - 2)/(1+2s)}).$$

PROOF. Proof follows by arguments similar to those given for the proof of Subtheorem 3.1. □

4. A lower bound for $[R(\phi_n, G) - R(G)]$ and the best possible rate of convergence

In this section we will show that the rate of asymptotic optimality of ϕ_n obtained in the previous section is not far away from the best possible rate that one can expect with ϕ_n, especially for large s. In fact, we will show that for large n, $R(\phi_n, G) - R(G) \ge c_1 n^{-2s/(1+2s)}$ at every G degenerate at a point in Θ.

THEOREM 4.1. *Let G be a degenerate distribution with total mass at an arbitrary but fixed point θ in Θ. Let ϕ_n be as given in Theorem 3.1 with $h = c_0 n^{-1/(1+2s)}$. Let there exist an $\eta > 0$ and $l > 0$ such that Lebesgue-*$\inf_{l<t<l+\eta} u(t)$

and Lebesgue-$\sup_{l<t<l+\eta} u(t)$ *are respectively positive and finite. Then for all* n *sufficiently large*

$$R(\phi_n, G) - R(G) > c' n^{-2s/(1+2s)} \tag{4.1}$$

for some constant c'.

PROOF. Throughout this proof let $c_1, c_2, \ldots$ stand for finite positive constants and denote $f(\cdot \mid \theta)$ and $p(\cdot \mid \theta)$ respectively by f and p. Since θ is in Θ,

$$0 < \inf_{l<t<l+\eta} p(t) \le \sup_{l<t<l+\eta} p(t) < \infty. \tag{4.2}$$

Since G is degenerate at θ, $\phi_G \equiv \theta$ and by (2.3)

$$\begin{aligned} &(R(\phi_n, G) - R(G))^{1/2} \\ &\ge E\{|\phi_n(X) - \theta| I(l < X < l + (\eta/2)) \\ &\ge E\{I(l < X < l + \eta/2)\} \\ &\cdot \int_0^{\beta-\theta} P_X[\psi_n(X) - \theta p_n(X) > t|p_n(X)|]dt\} \end{aligned} \tag{4.3}$$

for some (known or unknown finite) $\beta > \theta$, where P_X stands for the conditional probability on the Borel field generated by $\{X_1, \ldots, X_n\}$ given X.

Now for $0 < t < \beta - \theta$, $l < x < l + (\eta/2)$ and for $j = 1, \ldots, n$, define Y_j by

$$\begin{aligned} u(X_j)Y_j(x) = \Bigg[&I(X_j \ge x, u(X_j) > 0) \\ &- h^{-1}\left\{\theta K\left(\frac{X_j - x}{h}\right) + t\left|K\left(\frac{X_j - x}{h}\right)\right|\right\} I(u(X_j) > 0\Bigg], \end{aligned}$$

where K is the kernel appearing in the definition of p_n. Note that $Y_1, \ldots, Y_n$ are marginally i.i.d. since $X_1, \ldots, X_n$ are so. Let μ_0 and σ_0^2 denote respectively the mean and variance of Y_1. Then in view of (4.2) it can easily be checked (e.g. see pp. 79–80 of Singh (1974)) that for $l < x < l + (\eta/2)$ and for all n sufficiently large,

$$\begin{aligned} \mu_0 &= \left\{\int_x^\infty p(y) - \theta \int K(y)p(x + hy)dy - t \int |K(y)|p(x + hy)dy\right\} \\ &\ge -c_1(h^s + t) \end{aligned} \tag{4.4}$$

and

$$\sigma_0^2 \ge c_2 h^{-1}. \tag{4.5}$$

Now, note that $P_{X=x}[\psi_n(x) - \theta p_n(x) > t|p_n(x)|] \ge P_{X=x}\left[\sum_1^n Y_j > 0\right] \ge P_{X=x}\left[\sum_1^n (Y_j - \mu_0) \ge c_1(h^s + t)\right]$ from (4.4). Therefore, it follows from Lemma 3 on p. 47 of Lamperty (1966) that for all $\xi > 0$ and for all n sufficiently large

$$\begin{aligned} P_{X=x}[\psi_n(x) - \theta p_n(x) > t|p_n(x)|] &\ge \exp\left\{-\frac{nc_1^2(h^s + t)^2}{2\sigma_0^2}(1 + \xi)\right\} \\ &\ge \exp\{-c_3 nh(h^s + t)^2\}. \end{aligned} \tag{4.6}$$

Therefore, with the transformation $(c_3nh)^{1/2}(h^s+t)=v$ we get from (4.6),

$$(4.7)\qquad \int_0^{\beta-\theta} P_{X=x}[\psi_n(x)-\theta p_n(x)>t|p_n(x)|]dt \geq (c_3nh)^{-1/2}\int_a^{a'} e^{-v^2}dv$$

where $a=(c_3nh)^{1/2}h^s$ and $a'=(c_3nh)^{1/2}(h^s+\beta-\theta)$. Since from the definition of h, $a=c_5^{1/2}$ and $a'\to\infty$ as $n\to\infty$, the integral on the r.h.s. of (4.3) converges to a constant uniformly in $l<x<l+(\eta/2)$. Therefore, from (4.3) for all n sufficiently large

$$(4.8)\qquad (R(\phi_n,G)-R(G))^{1/2}\geq c_5(nh)^{-1/2}EI(l<X<l+(\eta/2)).$$

Since θ is in Θ, the proof of the theorem is now complete since the expectation on the r.h.s. of (4.8) is strictly positive in view of (4.2) and our choice of l and η. □

Remark 4.1. The assumptions of Theorem 4.1 are fairly mild in the sense that these are satisfied for every choice of l and η, as long as these are positive and finite, in a number of exponential families, including those given in Examples 3.1, 3.2 and 3.3

5. Concluding remarks and some further results

We have given a method to construct a class of EB estimators ϕ_n for the family $u(x)C(\theta)\exp(-x/\theta)$ which are a.o. and achieve the rate of convergence *arbitrarily close to their best possible rates* obtained in Section 4. In fact, we have exhibited EB estimators a.o. with rates arbitrarily close to $O(n^{-1})$. Section 4 shows that the rate $O(n^{-1})$ can not be achieved by our estimators. *In fact, a rate of the order $O(n^{-1})$ has not yet been established for any EB procedures, whatever may be the component problem, in any Lebesgue-exponential, nonexponential regular or irregular family.*

Since p_n in (2.5) consistently estimates p, another version of estimators of $\psi(x)=\int_x^\infty p(t)dt$, worth utilizing in estimation of $\phi_G=\psi/p$, are $\psi_n^*(x)=\int_x^\infty p_n(t)dt$ and $\psi_n^{**}=\int_x^{x+k_n}p_n(t)dt$ where $0<k_n\to 0$ as $n\to\infty$. Singh (1978*a*, 1980) has considered estimators ψ_n^{**} with $k_n\sim h^{-1}$ and has proved mean square and strong consistencies. It can be shown by arguments similar to those of Singh (1979) and of the proof of Theorem 3.1, that under conditions analogous to those of Theorem 3.1, the estimators ϕ_n^* and ϕ_n^{**}, resulting from the replacement of ψ_n in (2.11) by ψ_n^* and ψ_n^{**} (with $k_n\sim h^{-1}$) respectively, are also a.o. with the same rate as achieved in (3.2) by ϕ_n.

Instead of ϕ_n in (2.11), let us consider $\hat{\phi}_n=\psi_n/\hat{p}_n$ where $\hat{p}_n$ is p_n or δ_n for some positive $\delta_n(\to 0)$ depending on whether $|p_n|>\delta_n$ or $|p_n|\leq\delta_n$. Then from (2.1), $E(\hat{\phi}_n-\phi_G)^2\leq 2\delta_n^{-2}\{E(\psi_n-\psi)^2+\phi_G^2E(\hat{p}_n-p)^2\}$, which in turn is no more than $2\delta_n^{-2}\{E(\psi_n-\psi)^2+\phi_G^2E(p_n-p)^2\}+2P[|p_n|\leq\delta_n]$. In view of the consistency properties (Theorem 2.1) of ψ_n and p_n as estimators of ψ and p respectively, and in view of the identity (2.3), the following theorem can be easily proved.

THEOREM 5.1. *Let u_ϵ, v_ϵ and $p_\epsilon^{(s)}$ be as defined in (2.6). Let $\mathcal{G}$ be the set of all probability measures on $(\Theta, \mathcal{F})$ such that for prior distributions in $\mathcal{G}$, $Ew(X) < \infty$ and for some $\epsilon > 0$ and integer $s > 0$, $\int \{(u(t)/u_\epsilon(t)) + p_\epsilon^{(s)}(t)\}\psi^2(t)dt < \infty$. Then choosing $h \sim n^{-1/(1+2s)}$ and $\delta \sim h^s$, $(R(\hat{\phi}_n, G) - R(G)) = o(1)$ for all G in $\mathcal{G}$.*

The condition $\int p_\epsilon^{(s)}\psi^2$ in the above theorem is not needed if p is assumed to satisfy s-th order Lipschitz condition, because then, it can be shown that $E(p_n(x) - p(x))^2$ is a constant times $\max\{V_\epsilon(x), 1\}h^{2s}$. It can be checked by arguments similar to those used in Section 3 that the conditions of Theorem 5.1 are satisfied for the σ-finite measure μ and u in Examples 3.1, 3.2 and 3.3.

Finally, if Θ is assumed to be a subset of $(0, a)$ for some $0 < a < \infty$, then our suggested EB estimators ϕ_n, ϕ_n^* and ϕ_n^{**} would be respectively the restriction of ψ_n/p_n, ψ_n^*/p_n and ψ_n^{**}/p_n to the interval $(0, a)$ (i.e. take ϕ_n equal to 0, ψ_n/p_n or a depending on whether ψ_n/p_n is ≤ 0, between 0 to a or $\geq a$). Then using the identity (2.3) and inequality (3.1) with $L = a$, and (2.7) and (2.8) we get for $0 < \lambda \leq 2$,

$$\begin{aligned} R(\phi_n, G) - R(G) &\leq 2p^{-\lambda}\{a^{2-\lambda}E|\psi_n - \psi|^\lambda + 2a^2E|p_n - p|^\lambda\} \\ &= O(n^{-\lambda s/(1+2s)}) \\ &\quad \cdot \{Ep^{-\lambda}(X)w^{\lambda/2}(X) + Ep^{-\lambda}(X)(p_\epsilon^{(s)}(X))^\lambda + Ep^{-\lambda}v_\epsilon^{\lambda/2}(X)\} \end{aligned}$$

from (2.7). And by arguments used in the proof of Theorem 3.1 it follows that under (A1), (A2) and (A3) with λ in $[2/s, 2]$,

$$R(\phi_n, G) - R(G) = O(n^{-\lambda s/(1+2s)}). \tag{5.1}$$

Thus with bounded parameter space Θ, ϕ_n achieve almost the best possible rate $O(n^{-2s/(1+2s)})$ by taking λ equal to 2 or arbitrarily close to 2. It can be easily shown that for every family $\mathcal{G}$ of probability measures on $(\Theta, \mathcal{F})$ with Θ, a subset of $[\theta_0, \theta_1]$, $0 < \theta_0 < \theta_1 < \infty$, and for every σ-finite measure μ on $(\mathcal{X}, \mathcal{B})$ with $u(x) = d\mu(x)/dx$ vanishing off a finite interval, conditions for (5.1) are satisfied for $\lambda = 2$, *thus giving examples where the best possible rates are indeed achieved.*

REFERENCES

Fox, R. (1978). Solutions to empirical Bayes squared error loss estimation problems, *Ann. Statist.*, **6**, 846–853.

Hannan, J. and Macky, D. W. (1971). Empirical Bayes squared-error loss estimation of unbounded functionals in exponential families, RM-290, Department of Statistics and Probability, Michigan State University, East Lansing.

Johns, M. V., Jr. (1957). Nonparametric empirical Bayes procedures, *Ann. Math. Statist.*, **28**, 649–669.

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

Johns, M. V., Jr. and Van Ryzin, J. R. (1972). Convergence rates for empirical Bayes two-action problems II: continuous case, *Ann. Math. Statist.*, **43**, 934–947.

Lamperty, J. (1966). *Probability*, W. A. Benjamin, New York.

Lehmann, E. L. (1959). *Testing of Statistical Hypothesis*, Wiley, New York.

Liang, T. (1988). On the convergence rates of empirical Bayes rules for two-action problems: discrete case, *Ann. Statist.*, **16**, 1635–1642.

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

Maritz, J. S. (1969). Empirical Bayes estimation for continuous distributions, *Biometrika*, **56**, 349–359.

Maritz, J. S. and Lwin, T. (1975). Construction of simple empirical Bayes estimators, *J. Roy. Statist. Soc. Ser. B*, **75**, 421–425.

O'Bryan, T. E. and Susarla, V. (1976). Rates in the empirical Bayes estimation problem with nonidentical components: case of normal distributions, *Ann. Inst. Statist. Math.*, **28**, 389–397.

Prasad, B. and Singh, R. S. (1990). Estimation of prior distribution and empirical Bayes estimation in a nonexponential family, *J. Statist. Plann. Inference*, **24**, 81–86.

Robbins, H. (1955). An empirical Bayes approach to statistics, *Proc. Third Berkeley Symp. on Math. Statist. Prob.*, Vol. 1, 157–163, Univ. California Press, Berkeley.

Robbins, H. (1963). The empirical Bayes approach to the testing of statistical hypothesis, *Review International Statistical Institute*, **31**, 195–208.

Robbins, H. (1964). The empirical Bayes approach to statistical decision problems, *Ann. Math. Statist.*, **35**, 1–20.

Samuel, E. (1963). An empirical Bayes approach to the testing of certain parametric hypotheses, *Ann. Math. Statist.*, **34**, 1370–1385.

Singh, R. S. (1974). Estimation of derivatives of average of μ-densities and sequence compound estimation in exponential family, RM-318, Department of Statistics and Probability, Michigan State University, East Lansing.

Singh, R. S. (1976). Empirical Bayes estimation with convergence rates in non-continuous Lebesgue-exponential families, *Ann. Statist.*, **4**, 431–439.

Singh, R. S. (1977*a*). Improvement on some known nonparametric uniformly consistent estimators of derivatives of a density, *Ann. Statist.*, **5**, 394–400.

Singh, R. S. (1977*b*). Applications of estimators of a density and its derivatives to certain statistical problems, *J. Roy. Statist. Soc. Ser. B*, **39**, 357–363.

Singh, R. S. (1978*a*). Sequence-compound estimation in scale-exponential families and speed of convergence, *J. Statist. Plann. Inference*, **2**, 53–62.

Singh, R. S. (1978*b*). Nonparametric estimation of derivatives of average of μ-densities with convergence rates and applications, *SIAM J. Appl. Math.*, **35**, 637–649.

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

Singh, R. S. (1980). Estimation of regression curves when the conditional density of the predictor variable is in scale-exponential family, *Multivariate Statistical Analysis* (ed. R. P. Gupta), 199–205, North-Holland, Amsterdam.

Singh, R. S. (1981). Speed of convergence in nonparametric estimation of a multivariate μ-density and its mixed partial derivatives, *J. Statist. Plann. Inference*, **5**, 287–298.

Singh, R. S. and Prasad, B. (1991). Uniformly strongly consistent prior-distribution and empirical Bayes estimators with asymptotic optimality and rates in a nonexponential family, *Sankhyā Ser. A*, **53**, 334–342.

Yu, B. (1971). Rates of convergence in empirical Bayes two-action and estimation problems and in sequence-compound estimation problems, RM-279, Department of Statistics and Probability, Michigan State University, East Lansing.