*The Annals of Statistics*
1975, Vol. 3, No. 1, 155-164

# RATES OF CONVERGENCE IN EMPIRICAL BAYES ESTIMATION PROBLEMS: CONTINUOUS CASE[1]

BY PI-ERH LIN

*Florida State University*

In this paper we construct sequences of estimators for a density function and its derivatives, which are not assumed to be uniformly bounded, using classes of kernel functions. Utilizing these estimators, a sequence of empirical Bayes estimators is proposed. It is found that this sequence is asymptotically optimal in the sense of Robbins (*Ann. Math. Statist.* **35** (1964) 1-20). The rates of convergence of the Bayes risks associated with the proposed empirical Bayes estimators are obtained. It is noted that the exact rate is $n^{-q}$ with $q \leqq \frac{1}{3}$. An example is given and an explicit kernel function is indicated.

**1. Introduction.** Let $f(x \mid \theta)$ be a density function of $x \in R$, given $\theta \in \Omega$, where $R$ is an interval of the real line and $\Omega$ is the parameter space. Let $\lambda(\theta)$ be a real-valued measurable function of $\theta$ and $d(x)$ be a decision function when $x$ is observed. We wish to estimate $\lambda(\theta)$ with respect to the squared error $\{d(x) - \lambda(\theta)\}^2$. In the Bayes framework, it is assumed that $\theta$ has an a priori distribution $G(\theta)$ on the $\sigma$-field of subsets of $\Omega$. The Bayes estimate of $\lambda(\theta)$ relative to $G(\theta)$ is given by

$$d_G(x) = \int_\Omega \lambda(\theta) f(x \mid \theta)\, dG(\theta)/f(x),$$

where

$$f(x) = \int_\Omega f(x \mid \theta)\, dG(\theta).$$

The Bayes risk associated with $d_G(x)$ is given by

$$B(G) = \int_\Omega \int_R \{d_G(x) - \lambda(\theta)\}^2 f(x \mid \theta)\, dx\, dG(\theta). \tag{1.1}$$

In practice, $G(\theta)$ is usually unknown. This leads to the consideration of the empirical Bayes procedure first suggested by Robbins (1955) and later developed by Johns (1957), Johns and Van Ryzin (1971, 1972), Krutchkoff (1967), Lin (1972), Maritz (1970), Robbins (1963, 1964) and Samuel (1963), among others.

In the empirical Bayes framework, we make the following assumptions: Let $(x_1, \theta_1), \cdots, (x_n, \theta_n), \cdots$ be a sequence of independent random vectors, the $\theta_n$ having a common a priori distribution $G(\theta)$ and the conditional density of $x_n$ given $\theta_n = \theta$ being $f(x \mid \theta)$. At the $(n + 1)$st stage, when the decision is to be made about $\lambda(\theta_{n+1})$, we will have observed $x_1, \cdots, x_{n+1}$, although the values of

Received April 1972; revised January 1974.

[1] This research was supported in part by NSF Grant GP-7456 (Columbia University), in part by the Florida State University Grant 011204-043 and in part by the National Institute of General Medical Sciences through Training Grant 5T01 GM 913.

*AMS* 1970 *subject classifications.* Primary 62F15; Secondary 62F10.

*Key words and phrases.* Squared error, empirical Bayes estimation, kernel function, rate of convergence, asymptotically optimal.

$\theta_1, \cdots, \theta_n$ remain unknown. From this knowledge, we wish to construct the empirical Bayes decision function

$$d_n(x) = d_n(x_1, \cdots, x_n; x),$$

depending on $x_1, \cdots, x_n, x$, and thereby incur the loss $\{d_n(x) - \lambda(\theta_{n+1})\}^2$. The Bayes risk associated with $d_n(x)$ is given by

$$(1.2) \qquad B_n = \int_\Omega \overbrace{\int_R \cdots \int_R}^{n+1} \{d_n(x) - \lambda(\theta)\}^2 \prod_{i=1}^n f(x_i) f(x \mid \theta)\, dx_1 \cdots dx_n \cdot dx\, dG(\theta).$$

The sequence of estimators $\{d_n(x)\}$ is called asymptotically optimal if $B_n \to B(G)$ as $n \to \infty$.

Recently, Johns and Van Ryzin (1971, 1972) obtained the convergence rates for two-action problems. Lin (1972) obtained the rates of convergence in estimation problems for the discrete case. In this paper, we investigate the rates of convergence in estimation problems for the continuous case.

Consider the family of exponential distributions having a density of the form

$$f(x \mid \theta) = h(x)\beta(\theta)e^{-\theta x}, \qquad x > a,\ \theta \in \Omega = [0, \infty),$$

where $a$ may be finite or infinite. Assume that $G(\theta)$ is a prior distribution on the natural parameter space such that

$$\int_0^\infty \lambda^2(\theta)\, dG(\theta) < \infty,$$

and that $g(x) = d_G(x) f(x)$ may be written in the following form

$$(1.3) \qquad g(x) = \sum_{t=0}^m w_t(x) f^{(t)}(x),$$

where we have defined

$$f^{(t)}(x) = \frac{d^t}{dx^t} f(x), \qquad t = 0, \cdots, m,$$

$$f^{(0)}(x) = f(x),$$

and $w_t(x)$ is a function of $h^{(s)}(x)$, $s = 0, \cdots, t$. In the following section we propose a sequence of empirical Bayes estimators for $\lambda(\theta)$ by constructing estimators of $f^{(t)}(x)$, $t = 0, \cdots, m$. The exact rates of convergence are obtained in Section 3. An example is given in Section 4.

**2. Empirical Bayes estimations.** If $G(\theta)$ is unknown and if $d_G(x) = g(x)/f(x)$ where $g(x)$ takes the form (1.3), an empirical Bayes estimator of $\lambda(\theta)$ can be constructed using the estimators of $f^{(t)}(x)$ for $t = 0, \cdots, m$. Schuster (1969) obtains estimators of a density function and its derivatives using a known differentiable density, e.g., the standard normal, Cauchy, etc. In the same paper, the convergence rates are obtained by assuming that the first $m + 1$ derivatives of the density are uniformly bounded. In the paper by Johns and Van Ryzin (1972) as well as in the present paper the authors relax the uniform boundedness assumption and use the classes of kernel functions defined as follows: Let $\mathscr{K}_t$ be a class of measurable functions $K_t(u)$, $t = 0, \cdots, m$, called kernel functions, satisfying the conditions

$$K_t(u) = 0\,, \qquad u \leqq 0 \text{ or } u \geqq u_t,\ u_t > 0\,, \tag{2.1}$$

$$\textstyle\int_0^{u_t} u^t K_t(u)\,du = 1\,, \qquad \text{and} \tag{2.2}$$

$$\sup_u |K_t(u)| < \infty\,. \tag{2.3}$$

Using the kernel functions $K_t(u)$, we estimate $f^{(t)}(x)$ for each $t$ by

$$\begin{aligned} f_n^{(t)}(x) &= f_n^{(t)}(x_1, \cdots, x_n; x) \\ &= \frac{1}{na_n^{t+1}} \textstyle\sum_{j=1}^n \sum_{s=0}^t (-1)^s \binom{t}{s} \frac{1}{t-s+1} K_t\left\{\frac{x_j - x}{(t-s+1)a_n}\right\}, \end{aligned} \tag{2.4}$$

(and denote $f_n(x) = f_n^{(0)}(x)$) where $\{a_n\}$ is a sequence of positive numbers such that

$$a_n = O(n^{-\alpha}) \qquad \text{with } 0 < \alpha < \frac{1}{2(t+1)}\,. \tag{2.5}$$

It will be shown in Corollary 3.1.2 that for any $x > a$,

$$f_n^{(t)}(x) \to f^{(t)}(x)\,, \quad \text{in probability,} \qquad \text{as } n \to \infty\,. \tag{2.6}$$

Utilizing $f_n^{(t)}(x)$, we propose a sequence of empirical Bayes estimators $\{d_n(x)\}$ for $\lambda(\theta)$ as follows: Let

$$d_n(x) = d_n(x_1, \cdots, x_n; x) = \frac{g_n(x)}{f_n^*(x)} = \frac{\sum_{t=0}^m w_t(x) f_n^{(t)}(x)}{f_n^*(x)}\,, \tag{2.7}$$

where

$$\begin{aligned} f_n^*(x) &= f_n(x) \qquad \text{if } |f_n(x)| > \delta_n \\ &= \delta_n \qquad \text{if } |f_n(x)| \leqq \delta_n\,, \end{aligned} \tag{2.8}$$

with $\{\delta_n\}$ being a sequence of positive numbers such that

$$b_1 n^{-\gamma} \leqq \delta_n \leqq b_2 n^{-\gamma}\,, \qquad 0 < b_1 \leqq b_2 < \infty,\ \gamma > 0\,. \tag{2.9}$$

It will be seen in the next section that the sequence $\{d_n(x)\}$ given by (2.7) is asymptotically optimal.

**3. Rates of convergence.** In order to establish the convergence rates for the sequence of empirical Bayes estimators $\{d_n(x)\}$, we need further assumptions on the kernel function $K_t(u)$. Assume that there exists a positive integer $r$ such that $f^{(r)}(x)$ exists and is continuous for all $x > a$, and that for any $K_t(u) \in \mathscr{K}_t$, $t = 0, \cdots, m$,

$$\begin{aligned} \textstyle\int_0^{u_t} u^k K_t(u)\,du &= 0 \qquad \text{if } k = t+1, \cdots, r-1\,, \\ &= d_t \qquad \text{if } k = r\,, \end{aligned} \tag{3.1}$$

and

$$\textstyle\int_0^{u_t} u^r |K_t(u)|\,du < \infty\,, \tag{3.2}$$

where $d_t$ is finite depending only on $t$ and $r$.

Johns and Van Ryzin (1972, Lemma 2) have shown that if $G(\theta)$ is any prior distribution on the natural parameter space, then the existence and continuity of $h^{(r)}(x)$ imply the existence and continuity of $f^{(r)}(x)$ for all $x > a$. We have the following convergence rate for the estimator of the $t$th ($t \geqq 0$) derivative of $f(x)$.

THEOREM 3.1. *If $h^{(r)}(x)$ exists and is continuous for all $x > a$, $r$ a positive integer, and if conditions* (2.1), (2.2), (2.3), (3.1), *and* (3.2) *are satisfied, then for $t = 0, \cdots, m$, and $\varepsilon > 0$,*

$$|Ef_n^{(t)}(x) - f^{(t)}(x)| = O(n^{-\alpha(r-t)}) \cdot f_\varepsilon^{(r)}(x)\,, \tag{3.3}$$

$$E\{f_n^{(t)}(x) - f^{(t)}(x)\}^2 = O(n^{-2\alpha(r-t)}) \cdot \{f_\varepsilon^{(r)}(x)\}^2\,, \tag{3.4}$$

*where $\alpha = [2(r+1)]^{-1}$ and*

$$f_\varepsilon^{(r)}(x) = \sup_{0 \leqq \xi \leqq \varepsilon} |f^{(r)}(x+\xi)|\,. \tag{3.5}$$

In (3.3) and (3.4) $E$ denotes the expectation with respect to the joint distribution of $x_1, \cdots, x_n$, holding $x$ fixed.

PROOF. For a given $x > a$, the expectation of $f_n^{(t)}(x)$ is

$$\begin{aligned} Ef_n^{(t)}(x) &= \frac{1}{a_n^{t+1}} \textstyle\sum_{s=0}^t (-1)^s \binom{t}{s} \frac{1}{t-s+1} \int_a^\infty K_t\left\{\frac{y-x}{(t-s+1)a_n}\right\} f(y)\,dy \\ &= \frac{1}{a_n^t} \textstyle\sum_{s=0}^t (-1)^s \binom{t}{s} \int_0^{u_t} K_t(u) f\{x + (t-s+1)a_n u\}\,du \\ &= \frac{1}{a_n^t} \textstyle\sum_{s=0}^t (-1)^s \binom{t}{s} \sum_{k=0}^{r-1} f^{(k)}(x) \frac{(t-s+1)^k a_n^k}{k!} \int_0^{u_t} u^k K_t(u)\,du \\ &\quad + a_n^{r-t} R_t(r, x, n) \\ &= f^{(t)}(x) + a_n^{r-t} R_t(r, x, n)\,, \end{aligned} \tag{3.6}$$

where the remainder term is

$$R_t(r, x, n) = \textstyle\sum_{s=0}^t (-1)^s \binom{t}{s} \frac{(t-s+1)^r}{r!} \int_0^{u_t} u^r K_t(u) f^{(r)}(x+\xi)\,du$$

with $0 \leqq \xi \leqq (t-s+1)a_n u$. The second equality of (3.6) is obtained by the change of variable $u = (y-x)/\{(t-s+1)a_n\}$; the third and fourth equalities follow by applying a Taylor's expansion of $f\{x + (t-s+1)a_n u\}$ about $x$ and by using the identity (see Feller (1957), page 63)

$$\begin{aligned} \textstyle\sum_{s=0}^t (-1)^s \binom{t}{s} (t-s+1)^k &= 0 \qquad \text{if } k = 0, \cdots, t-1 \\ &= t! \qquad \text{if } k = t\,. \end{aligned} \tag{3.7}$$

The continuity of $f^{(r)}(x)$ for all $x > a$ and assumption (3.2) ensure that

$$|R_t(r, x, n)| < M_t \cdot f_\varepsilon^{(r)}(x)\,, \qquad 0 < M_t < \infty\,.$$

This proves the assertion (3.3). As for (3.4) we note that for a given $x > a$,

$$E\{f_n^{(t)}(x) - f^{(t)}(x)\}^2 = \operatorname{Var}\{f_n^{(t)}(x)\} + \{Ef_n^{(t)}(x) - f^{(t)}(x)\}^2\,. \tag{3.8}$$

The variance of $f_n^{(t)}(x)$ is

$$\text{(3.9)}\qquad \operatorname{Var}\{f_n^{(t)}(x)\} = \frac{1}{na_n^{2(t+1)}} \operatorname{Var}\left[\sum_{s=0}^{t} (-1)^s \binom{t}{s} \frac{1}{t-s+1} K_t\left\{\frac{y-x}{(t-s+1)a_n}\right\}\right].$$

Using the $c_\delta$-inequality (Loève (1960), page 155) with $\delta = 2$, we have

$$\begin{aligned} &\operatorname{Var}\left[\sum_{s=0}^{t} (-1)^s \binom{t}{s} \frac{1}{t-s+1} K_t\left\{\frac{y-x}{(t-s+1)a_n}\right\}\right] \\ \text{(3.10)}\qquad &\leqq \int_a^\infty \left[\sum_{s=0}^{t} (-1)^s \binom{t}{s} \frac{1}{t-s+1} K_t\left\{\frac{y-x}{(t-s+1)a_n}\right\}\right]^2 f(y)\,dy \\ &\leqq \sum_{s=0}^{t} c_2 \binom{t}{s}^2 \frac{1}{(t-s+1)^2} \int_a^\infty K_t^2\left\{\frac{y-x}{(t-s+1)a_n}\right\} f(y)\,dy \\ &\leqq \sum_{s=0}^{t} c_2 \binom{t}{s}^2 \frac{1}{(t-s+1)^2} \sup_u |K_t(u)|^2, \qquad (c_2 = 2^{t+1}), \end{aligned}$$

which is finite by condition (2.3). Thus (3.8) is bounded by

$$\text{(3.11)}\qquad O(n^{-1}a_n^{-2(t+1)}) + O(n^{-2\alpha(r-t)}) \cdot \{f_\varepsilon^{(r)}(x)\}^2$$

which is equal to $O(n^{-((r-t)/(r+1))}) \cdot \{f_\varepsilon^{(r)}(x)\}^2$ by choosing $\alpha = [2(r+1)]^{-1}$ such that both terms of (3.11) have the same convergence rate. This completes the proof of the theorem.

The results (3.3) and (3.4) both depend on $f_\varepsilon^{(r)}(x)$. If we impose the uniform boundedness on $f^{(r)}(x)$, the following corollary to Theorem 3.1 is easily obtained. The proof is omitted.

COROLLARY 3.1.1. *If the conditions of Theorem* 3.1 *hold and if for every* $\varepsilon > 0$,

$$\text{(3.12)}\qquad \sup_x f_\varepsilon^{(r)}(x) < \infty,$$

*then*

$$\sup_x |Ef_n^{(t)}(x) - f^{(t)}(x)| = O(n^{-\alpha(r-t)}),$$
$$\sup_x E\{f_n^{(t)}(x) - f^{(t)}(x)\}^2 = O(n^{-2\alpha(r-t)}),$$

*with* $\alpha = [2(r+1)]^{-1}$.

If condition (3.1) does not hold, the convergence rate for $f_n^{(t)}(x)$ may not be obtained. However, the asymptotic unbiasedness and the convergence property of $f_n^{(t)}(x)$ can always be established.

COROLLARY 3.1.2. *Suppose that* $h^{(t)}(x)$ *exists and is continuous for* $x > a$, $t$ *a positive integer, and that condition* (3.12) *holds. If* $f_n^{(t)}(x)$ *is given by* (2.4) *using the kernel function* $K_t(u)$ *satisfying conditions* (2.1), (2.2), *and* (2.3), *then* (2.6) *holds.*

PROOF. We wish to show that (3.8) converges to 0 as $n \to \infty$. From (3.9) and (3.10) it is clear that $\operatorname{Var}\{f_n^{(t)}(x)\} \to 0$ as $n \to \infty$. It remains to show that the bias of $f_n^{(t)}(x)$ also converges to 0. Similar to (3.6) we have

$$(3.13)\qquad Ef_n^{(t)}(x) = \frac{1}{a_n^t} \sum_{s=0}^{t} (-1)^s \binom{t}{s} \sum_{k=0}^{t-1} f^{(k)}(x) \frac{(t-s+1)^k a_n^k}{k!} \int_0^{u_t} u^k K_t(u)\,du + \sum_{s=0}^{t} (-1)^s \binom{t}{s} \frac{(t-s+1)^t}{t!} \int_0^{u_t} u^t K_t(u) f^{(t)}(x+\xi)\,du\,,$$

where $0 \leqq \xi \leqq (t-s+1)a_n u$. Note that the first term of the RHS of (3.13) vanishes and the second term converges to $f^{(t)}(x)$ as $n \to \infty$, by the bounded convergence theorem using conditions (3.2), (3.7), (2.2), (3.12), and the continuity of $f^{(t)}(x)$. This completes the proof of the corollary.

The following lemma is useful for establishing the convergence rate of $B_n$, as $n \to \infty$. The proof, here omitted, may be found, for example, in Maritz (1970, page 46).

LEMMA 3.1. *Let $B(G)$ and $B_n$ be given by* (1.1) *and* (1.2), *respectively. Then*

$$(3.14)\qquad 0 \leqq B_n - B(G) = \int_a^\infty f(x) E\{d_n(x) - d_G(x)\}^2\,dx\,,$$

*where $E$ denotes the expectation with respect to $x_1, \cdots, x_n$, given $x$.*

The main theorem of this paper is stated below:

THEOREM 3.2. *Let $B(G)$ and $B_n$ be given by* (1.1) *and* (1.2), *respectively. If*

(i) *the conditions of Theorem* 3.1 *are satisfied,*
(ii) *for every $\varepsilon > 0$,* (3.12) *holds,*
(iii) $\int_a^\infty \{w_t(x)\}^2 f(x)\,dx < \infty$, *for* $t = 0, \cdots, m$ $(m < r)$,
(iv) $\int_a^\infty \{d_G(x)\}^2 f(x)\,dx < \infty$,
(v) *for some $h$* $(0 < h \leqq 1)$ *and $b$* $(0 < b < \infty)$,

$$(3.15)\qquad \int_a^\infty \{d_G(x)\}^2 P\{|f_n(x)| \leqq \delta_n \mid x\} f(x)\,dx \leqq b\delta_n^h\,,$$

*where $d_G(x) = g(x)/f(x)$ and $\{\delta_n\}$ satisfies* (2.9). *Then*

$$(3.16)\qquad B_n - B(G) = O(n^{-q})$$

*with*

$$(3.17)\qquad q = \frac{h(r-m)}{(h+2)(r+1)}\,.$$

PROOF. To simplify notation, we suppress the argument $(x)$. Notice that by applying the $c_\delta$-inequality with $\delta = 2$ repeatedly the integrand of (3.14) is bounded above by

$$\frac{2f}{\delta_n^2} E(g_n - g)^2 + \frac{2g^2}{\delta_n^2 f} E(f_n^* - f)^2$$
$$\leqq \frac{2f}{\delta_n^2} E[\textstyle\sum_{t=0}^m w_t\{f_n^{(t)} - f^{(t)}\}]^2 + \frac{4g^2}{\delta_n^2 f}\{E(f_n - f)^2 + E(f_n^* - f_n)^2\}$$
$$\leqq \frac{f}{\delta_n^2} \textstyle\sum_{t=0}^m 2^{t+2} w_t^2 E\{f_n^{(t)} - f^{(t)}\}^2 + \frac{4g^2}{\delta_n^2 f}\{E(f_n - f)^2 + E(f_n^* - f_n)^2\}\,,$$

where we have set $g_n = g_n(x) = \sum_{t=0}^{m} w_t(x) f_n^{(t)}(x)$. From the above inequality in $x$, with the aid of Lemma 3.1, $B_n - B(G)$ is bounded above by

$$\begin{aligned}(3.18)\qquad & \frac{1}{\delta_n^2} \sum_{t=0}^{m} 2^{t+2} \int_a^\infty w_t^2 E[f_n^{(t)} - f^{(t)}]^2 f\,dx + \frac{4}{\delta_n^2} \int_a^\infty \frac{g^2}{f} E(f_n - f)^2\,dx \\ & + \frac{4}{\delta_n^2} \int_a^\infty \frac{g^2}{f} E[(\delta_n - f_n)^2 \,|\, |f_n| \leqq \delta_n] \cdot P[|f_n| \leqq \delta_n \,|\, x]\,dx \\ & = O(n^{2\gamma-(r-m)/(r+1)}) + O(n^{2\gamma - r/(r+1)}) + O(n^{-h\gamma}) \\ & = O(n^{-q}) .\end{aligned}$$

The first equality of (3.18) is obtained by using the second result of Corollary 3.1.1 and Conditions (iv) and (v). The second equality follows by choosing $\gamma = (r-m)/\{(h+2)(r+1)\}$. Finally, (3.16) is established by letting $q = h\gamma$.

Condition (v) of Theorem 3.2, which depends on $P\{|f_n(x)| \leqq \delta_n \,|\, x\}$, does not look appealing. If we impose a somewhat stronger condition on the Bayes estimator $d_G(x)$ than that given by (iv), then Condition (v) may be replaced by a condition depending only on $P\{f(x) \leqq \delta_n\}$. The following lemma provides a relationship between the unconditional probabilities of $\{|f_n(x)| \leqq \delta_n\}$ and $\{f(x) \leqq \delta_n\}$. Using this lemma other convergence rate results may be obtained.

LEMMA 3.2. *For any* $\delta_n > 0$,

$$(3.19)\qquad P\{|f_n(x)| \leqq \delta_n\} \leqq \frac{1}{\delta_n^2} \int_a^\infty E\{f_n(x) - f(x)\}^2 f(x)\,dx + P\{f(x) \leqq 2\delta_n\} ,$$

*where E is as defined in Lemma* 3.1.

PROOF. Let $A_n = \{x \,|\, f(x) > 2\delta_n\}$, $B_n = \{x_1, \cdots, x_n; x \,|\, |f_n(x)| \leqq \delta_n\}$; then,

$$\begin{aligned} P(B_n) &\leqq P(B_n \cap A_n) + P(A_n^c) \\ &= \int_{A_n} P(B_n \,|\, x) f(x)\,dx + P(A_n^c) .\end{aligned}$$

Now $x \in A_n$ implies $B_n \subset \{|f(x) - f_n(x)| > \delta_n\}$, and applying Chebyshev's inequality conditional on $x$, one obtains the RHS of (3.19) with the integral restricted to $A_n$, which of course can be removed.

The following corollaries to Theorem 3.2 present the convergence rate results without Condition (v).

COROLLARY 3.2.1. *If Conditions* (iv) *and* (v) *of Theorem* 3.2 *are replaced by*

(vi) *for some* $\delta > 0$, $\int_a^\infty |d_G(x)|^{2+\delta} f(x)\,dx < \infty$, *and*
(vii) *for some* $\beta$, $0 < \beta \leqq 1$, $P\{f(x) \leqq \delta_n\} \leqq d\delta_n^\beta$, $0 < d < \infty$.

*Then the result* (3.16) *holds with q given by* (3.17) *and* $h = \beta\delta/(\delta + 2)$.

PROOF. It suffices to show that (vi) and (vii) imply (iv) and (v) with $h = \beta\delta/(\delta+2)$. Since (iv) is a direct consequence of (vi), it remains to show that (v) holds. If we partition the region of integration of the LHS of (3.15) by $A_n = \{|d_G(x)| \leqq n^\nu\}$ and $B_n = \{|d_G(x)| > n^\nu\}$ for some $\nu > 0$. Then, with the aid of Lemma 3.2, the LHS of (3.15) becomes

$$\begin{aligned}
&\int_a^\infty \{d_G(x)\}^2 P\{|f_n(x)| \leqq \delta_n \mid x\} f(x)\, dx \\
&\quad = \int_{A_n} + \int_{B_n} \\
&\quad \leqq n^{2\nu} P\{|f_n(x)| \leqq \delta_n\} + n^{-\delta\nu} \int_a^\infty |d_G(x)|^{2+\delta} f(x)\, dx \\
&\quad \leqq n^{2\nu} \left\{ \frac{1}{\delta_n^{\,2}} \int_a^\infty E[f_n(x) - f(x)]^2 f(x)\, dx + P[f(x) \leqq 2\delta_n] \right\} + O(n^{-\delta\nu}) \\
&\quad = O(n^{2\nu+2\gamma-r/(r+1)}) + O(n^{2\nu-\gamma\beta}) + O(n^{-\delta\nu}) \\
&\quad = O(n^{2\nu-r\beta/\{(r+1)(\beta+2)\}}) + O(n^{-\delta\nu}), \qquad \text{by choosing } \gamma = r/\{(r+1)(\beta+2)\} \\
&\quad = O(n^{-\gamma h}), \qquad \text{by choosing } \nu = r\beta/\{(r+1)(\beta+2)(\delta+2)\},
\end{aligned}$$

where $h = \beta\delta/(\delta + 2)$. This completes the proof of Corôllary 3.2.1.

It should be noted that in order to apply this corollary the sequence of real numbers $\{\delta_n\}$ used in constructing the kernel functions $K_t(u)$ must be of order $n^{-\gamma}$ with $\gamma = r/\{(r + 1)(\beta + 2)\}$. The rate of convergence obtained in Corollary 3.2.1 is $\delta/(\delta + 2)$ times that obtained in Theorem 3.2. If $|d_G(x)|$ is uniformly bounded for all $x > a$, all absolute moments of $d_G(x)$ exist and $h \to \beta$ as $\delta \to \infty$. This can be obtained directly from Theorem 3.2 as well.

COROLLARY 3.2.2. *If Conditions* (iv) *and* (v) *of Theorem* 3.2 *are replaced by*

(viii) $\sup_x |d_G(x)| < \infty$, *and*
(ix) $P\{f(x) \leqq \delta_n\} \leqq d\delta_n^{\,h}$, $0 < d < \infty$, $0 < h \leqq 1$.

*Then the result* (3.16) *holds with q given by* (3.17).

PROOF. It suffices to show that (viii) and (ix) imply (iv) and (v). The proof is easy and is omitted.

It is noted that, in Corollary 3.2.2, $q \leqq \frac{1}{3}$. The value of $h$ in Condition (ix) achieves its maximum for the density $f(x) = e^{-x}$. In this case $q = (r - m)/\{3(r + 1)\}$, which converges to $\frac{1}{3}$ as $r \to \infty$.

**4. An example.** Consider the family of negative exponential distributions with density

$$\begin{aligned} f(x \mid \theta) &= \theta e^{-\theta x} && \text{if } x > 0, \quad \theta > 0, \\ &= 0 && \text{otherwise.} \end{aligned}$$

We wish to estimate $\theta$ relative to the squared error loss. If the prior distribution is given by

$$G(\theta) = \frac{1}{\Gamma(\alpha)} \int_0^\theta e^{-u} u^{\alpha-1}\, du, \qquad 0 < \alpha < \infty.$$

Then

$$\begin{aligned} f(x) &= \alpha(x+1)^{-(\alpha+1)} && x > 0, \quad 0 < \alpha < \infty, \\ &= 0 && \text{otherwise.} \end{aligned} \tag{4.1}$$

By a direct computation we have, for each $t = 0, 1, \cdots,$

$$f^{(t)}(x) = (-1)^t \frac{\Gamma(\alpha + t + 1)}{\Gamma(\alpha)} (x + 1)^{-(\alpha+t+1)}, \tag{4.2}$$

which is continuous for all $x > 0$. From (4.2) it is clear that

$$f_\varepsilon^{(r)}(x) \leqq \frac{\Gamma(\alpha + r + 1)}{\Gamma(\alpha)} = M, \qquad 0 < M < \infty, r > t, \tag{4.3}$$

uniformly in $x$, verifying (3.12). The Bayes estimator of $\theta$ is

$$\begin{aligned} d_G(x) &= -f^{(1)}(x)/f(x) \\ &= \frac{\alpha(\alpha + 1)(x + 1)^{-(\alpha+2)}}{\alpha(x + 1)^{-(\alpha+1)}} \\ &= \frac{\alpha + 1}{x + 1}, \end{aligned} \tag{4.4}$$

that is, $w_t(x) = 1$ for $t = 1$ and 0 for $t \neq 1$. This implies that Condition (iii) of Theorem 3.2 holds for $m = 1$. Moreover, it is seen from (4.4) that $d_G(x)$ is uniformly bounded for all $x > 0$, Corollary 3.2.2 may be applied. To do so, it remains to find the value of $h$ such that Condition (ix) holds. Note that $f(x)$ is strictly decreasing in $x > 0$ and that

$$\begin{aligned} P\{f(x) \leqq \delta_n\} &= P\left\{x \geqq \left(\frac{\alpha}{\delta_n}\right)^{1/(\alpha+1)} - 1\right\} \\ &= \int_{\{(\alpha/\delta_n)^{1/(\alpha+1)}-1\}}^{\infty} \{\alpha(x + 1)^{-(\alpha+1)}\}\, dx \leqq \delta_n^{\alpha/(\alpha+1)}, \end{aligned}$$

for $1 \leqq \alpha < \infty$. That is, $h = \alpha/(\alpha + 1)$. Therefore, Corollary 3.2.2 holds with $q = \alpha(r - 1)/\{(3\alpha + 2)(r + 1)\}$ which is close to $\frac{1}{3}$ for sufficiently large $\alpha$ and $r$.

In concluding, we present a kernel function $K_t(u) \in \mathscr{K}_t$ satisfying conditions (2.1), (2.2), (2.3), (3.1) and (3.2) with an arbitrary positive integer $r > t$. Let $D$ be the determinant of the $(r - t)$-matrix with elements $d_{ij} = 1/(t + i + j - 1)$ and let $D^*$ be the determinant of the same matrix except with $d_{1j}$ replaced by $(u/u_t)^{j-1}$. The kernel function is then given by $K_t(u) = D^*/(Du_t^{t+1})$ for $0 < u < u_t$ and 0 otherwise.

**Acknowledgment.** The author is grateful to the referees for some useful comments and suggestions.

## REFERENCES

[1] Feller, W. (1957). *An Introduction to Probability Theory and Its Applications* **1**, (2nd ed.). Wiley, New York.

[2] Johns, M. V. Jr. (1957). Non-parametric empirical Bayes procedures. *Ann. Math. Statist.* **28** 649–669.

[3] Johns, M. V. Jr. and Van Ryzin, J. (1971). Convergence rates in empirical Bayes two-action problems I. Discrete case. *Ann. Math. Statist.* **42** 1521–1539.

[4] Johns, M. V. Jr. and Van Ryzin, J. (1972). Convergence rates in empirical Bayes two-action problems II. Continuous case. *Ann. Math. Statist.* **43** 934–947.

[5] Krutchkoff, R. G. (1967). A supplementary sample non-parametric empirical Bayes approach to some statistical decision problems. *Biometrika* **54** 451–458.
[6] Lin, P. E. (1972). Rates of convergence in empirical Bayes estimation problems: Discrete case. *Ann. Inst. Statist. Math.* **24** 319–325.
[7] Loève, M. (1960). *Probability Theory*, (3rd ed.). Van Nostrand, Princeton.
[8] Maritz, J. S. (1970). *Empirical Bayes Methods.* Methuen Monograph. Methuen, London.
[9] Robbins, H. (1955). An empirical Bayes approach to statistics. *Proc. Third Berkeley Symp. Math. Statist. Prob.* **1** 157–164. Univ. of California Press.
[10] Robbins, H. (1963). The empirical Bayes approach to testing statistical hypotheses. *Rev. Inst. Internat. Statist.* **31** 195–208.
[11] Robbins, H. (1964). The empirical Bayes approach to statistical decision problems. *Ann. Math. Statist.* **35** 1–20.
[12] Samuel, E. (1963). An empirical Bayes approach to the testing of certain parametric hypotheses. *Ann. Math. Statist.* **34** 1370–1385.
[13] Schuster, E. (1969). Estimation of probability density function and its derivatives. *Ann. Math. Statist.* **40** 1187–1196.

Department of Statistics
Florida State University
Tallahassee, Florida 32306