The Annals of Mathematical Statistics
1972, Vol. 43, No. 3, 934-947

CONVERGENCE RATES FOR EMPIRICAL BAYES TWO-ACTION PROBLEMS II. CONTINUOUS CASE[1]

BY M. V. JOHNS, JR.[2] AND J. VAN RYZIN[3]

Stanford University and University of Wisconsin

1. Introduction and summary. For a general discussion of empirical Bayes problems and motivation of the present paper see Section 1 of the previous paper [1]. In that paper we studied the convergence to Bayes optimality and its rate properties for empirical Bayes two-action problems in certain discrete exponential families. This paper continues that investigation for the continuous case. Under appropriate conditions, Theorems 3 and 4 yield convergence rates to Bayes risk of $O(n^{-\beta})$ for $0 < \beta < 1$, for the $(n + 1)$st stage risk of the continuous case empirical Bayes procedures of Section 2. These theorems provide, for the continuous case, convergence rate results for the empirical Bayes procedures of the general type considered by Robbins [5] and Samuel [6] for two different parameterizations of a model. The rate results given here in the continuous case involve upper bounds and are weaker than the discrete case results in [1] wherein exact rates are reported.

Specifically, in Section 2 we present the two cases to be considered and define the appropriate empirical Bayes procedures for each. Section 3 gives some technical lemmas and Section 4 establishes the asymptotic optimality (the asymptotic Bayes property) of the procedures introduced. The main results on rates, Theorems 3 and 4, are given in Section 5. Section 6 examines in detail two specific examples—the negative exponential and the normal distributions—and gives corollaries to Theorems 3 and 4 which state convergence rates depending on moment properties of the unknown prior distribution of the parameters. Section 7 gives an example with β arbitrarily close to 1 in the rate $O(n^{-\beta})$.

The model we consider is the following. Let $f_\lambda(x)$ be a family of Lebesgue densities indexed by a parameter λ in an interval of the real line. As in [1], we wish to test the hypothesis $H_1: \lambda \leqq c$ vs. $H_2: \lambda > c$ with the loss function being

$$\begin{aligned} L_1(\lambda) &= 0 && \text{if } \lambda \leqq c \\ &= b(\lambda - c) && \text{if } \lambda > c \\ L_2(\lambda) &= b(c - \lambda) && \text{if } \lambda \leqq c \\ &= 0 && \text{if } \lambda > c \end{aligned}$$

where $L_i(\lambda)$ indicates the loss when action i (deciding in favor of H_i) is taken, $i = 1, 2$ and b is a positive constant.

Received October 23, 1967; revised September 9, 1971.

[1] Prepared under Contract Nonr-225(52) (NR-342-022) for Office of Naval Research.

[2] Supported by Public Health Service Grant USPHS-GM-14554-01.

[3] Supported in part by National Science Foundation Grant NSF GP-9324.

Let

$$\delta(x) = \Pr\{\text{accepting } H_1 \mid X = x\}$$

be a randomized decision rule for the above two-action problem. If $G = G(\lambda)$ is a prior distribution on λ, then the risk of the (randomized) decision procedure δ under prior distribution G is given as in [1] by,

$$(1) \qquad \begin{aligned} r(\delta, G) &= \int\int \{L_1(\lambda)f_\lambda(x)\delta(x) + L_2(\lambda)f_\lambda(x)(1 - \delta(x))\}\, dx\, dG(\lambda) \\ &= b \int \alpha(x)\delta(x)\, dx + C_G \end{aligned}$$

where $C_G = \int L_2(\lambda)\, dG(\lambda)$ and

$$(2) \qquad \alpha(x) = \int \lambda f_\lambda(x)\, dG(\lambda) - cf(x)$$

with

$$(3) \qquad f(x) = \int f_\lambda(x)\, dG(\lambda)\ .$$

From (1) it is clear that a Bayes rule (the minimizer of (1) given G) is

$$(4) \qquad \begin{aligned} \delta_G(x) &= 1 \qquad \text{if} \quad \alpha(x) \leqq 0 \\ &= 0 \qquad \text{if} \quad \alpha(x) > 0\ . \end{aligned}$$

Hence, the minimal attainable risk knowing G (the Bayes risk) is

$$(5) \qquad r^*(G) = \inf_\delta r(\delta, G) = r(\delta_G, G)\ .$$

2. The empirical Bayes approach where G is unknown. An empirical Bayes procedure for the $(n + 1)$st decision problem based on a previously seen sequence of independent, identically distributed random variables $X_1, \cdots, X_n$, each with probability density $f(x)$ is obtained by forming a function

$$\alpha_n(x) = \alpha_n(X_1, \cdots, X_n; x)$$

which estimates $\alpha(x)$ for each x and adopting the empirical Bayes procedure

$$(6) \qquad \begin{aligned} \delta_n(x) &= 1 \qquad \text{if} \quad \alpha_n(x) \leqq 0 \\ &= 0 \qquad \text{if} \quad \alpha_n(x) > 0\ . \end{aligned}$$

Here $\alpha_n(x)$ and hence $\delta_n(x)$ do not depend on the unknown prior distribution G.

Letting r_n denote the risk of δ_n, we seek procedures (6) and conditions under which $\lim r_n = r^*(G)$ and we investigate the speed of this convergence to Bayes optimality. Note that

$$(7) \qquad \begin{aligned} r_n &= E\{r(\delta_n, G)\} \\ &= bE\{\int \alpha(x)\delta_n(x)\, dx\} + C_G \end{aligned}$$

where E denotes expectation on $X_1, \cdots, X_n$.

The empirical Bayes approach for this problem has been previously treated in the continuous case in two papers under two different cases or parameterizations of a model for which $\alpha(x)$ may be estimated.

Case I (Robbins [5]).

$$f_\lambda(x) = e^{-\lambda x}\beta(\lambda)h(x) \quad \text{for} \quad x > a\,, \quad \lambda \text{ in an interval}$$
$$= 0 \quad \text{otherwise}\,,$$

where a may be finite or infinite and $h(x) > 0$ for $x > a$.

For this model, under certain regularity conditions, (see Lemma 2)

$$\int \lambda f_\lambda(x)\, dG(\lambda) = -f^{(1)}(x) + \frac{h^{(1)}(x)}{h(x)} f(x)$$

where $f^{(1)}(x)$ and $h^{(1)}(x)$ are the first derivatives of $f(x)$ and $h(x)$. Then by (2) we have

$$\alpha(x) = (v(x) - c)f(x) - f^{(1)}(x) \tag{8}$$

where

$$v(x) = \frac{h^{(1)}(x)}{h(x)}\,. \tag{9}$$

We shall estimate $\alpha(x)$ by estimating $f(x)$ and $f^{(1)}(x)$. Let $f_n(x) = f_n(X_1, \cdots, X_n; x)$ be an estimate of $f(x)$ given by (see Parzen [3])

$$f_n(x) = \frac{1}{nh_n} \sum_{j=1}^n K_1\left(\frac{X_j - x}{h_n}\right) \tag{10}$$

or by

$$f_n(x) = \frac{1}{nh_n} \sum_{j=1}^n \tfrac{1}{2}\left\{K_1\left(\frac{X_j - x}{h_n}\right) + K_1\left(\frac{x - X_j}{h_n}\right)\right\}, \tag{10'}$$

where $\{h_n\}$ is a sequence of positive numbers such that

$$h_n \downarrow 0\,, \qquad nh_n^3 \to \infty \tag{11}$$

and $K_1(u)$ by a real-valued measurable function on the real line such that

$$K_1(u) = 0 \quad \text{if} \quad u \leqq 0 \quad \text{or} \quad u \geqq u_1 \tag{12}$$

$$\int_0^{u_1} K_1(u)\, du = 1 \tag{13}$$

$$\sup_u |K_1(u)| < \infty\,. \tag{14}$$

Denote by $\mathscr{K}_1 = \mathscr{K}_1(u_1)$ the class of all real-valued measurable functions on the real line satisfying (12), (13) and (14).

To estimate $f^{(1)}(x)$, let $f_n^{(1)}(x) = f_n^{(1)}(X_1, \cdots, X_n; x)$ be given by

$$f_n^{(1)}(x) = \frac{1}{nh_n} \sum_{j=1}^n \left\{\frac{1}{2h_n} K_2\left(\frac{X_j - x}{2h_n}\right) - \frac{1}{h_n} K_2\left(\frac{X_j - x}{h_n}\right)\right\} \tag{15}$$

where $\{h_n\}$ is the sequence of (11) and $K_2(u)$ be a real-valued measurable function on the real line such that

$$K_2(u) = 0 \quad \text{if} \quad u \leqq 0 \quad \text{or} \quad u \geqq u_2 \tag{16}$$

$$\int_0^{u_2} uK_2(u)\, du = 1 \tag{17}$$

$$\sup_u |K_2(u)| < \infty\,. \tag{18}$$

Denote by $\mathscr{K}_2 = \mathscr{K}_2(u_2)$ the class of all real-valued measurable functions on the real line satisfying (16), (17) and (18).

As a motivation for (15), note that by the mean value theorem,

$$
\begin{aligned}
Ef_n^{(1)}(x) &= \frac{1}{h_n} E\left\{\frac{1}{2h_n} K_2\left(\frac{X-x}{2h_n}\right) - \frac{1}{h_n} K_2\left(\frac{X-x}{h_n}\right)\right\} \\
&= \frac{1}{h_n} \int_0^{u_2} K_2(u)\{f(x+2h_n u) - f(x+h_n u)\}\, du \qquad (19) \\
&= \int_0^{u_2} uK_2(u) f^{(1)}(x + h_n u + \xi_n(x,u))\, du
\end{aligned}
$$

where $0 < \xi_n(x, u) < h_n u$, provided that $f(x)$ has continuous first derivative for $x > a$. Hence, we see that as $n \to \infty$

$$Ef_n^{(1)}(x) \to f^{(1)}(x) \qquad \text{for all} \quad x > a . \tag{20}$$

We note at this point that the kernel K_1 used in the density estimate in (10) is not the usual symmetric kernel suggested by Parzen [4]. This asymmetry is necessary since counter-examples to the rate Theorems 3 and 4 can be constructed when symmetric kernels are used and the observations are bounded below. However, for convergence alone (Theorems 1 and 2) either (10) or its symmetrized analog (10′) will suffice.

Also, note that the estimate $f_n^{(1)}(x)$ given by (15) is not of the form originally suggested by Robbins ([5] page 204, (76)). The more sophisticated version used here is required in our proofs to achieve the rates of Theorems 3 and 4. For convergence alone (Theorem 1), the estimate of Robbins would suffice. (Subsequent to the submission of this paper, a detailed analysis of estimating $f^{(r)}(x)$ appeared in Schuster [7]. Use of his estimates of $f^{(1)}(x)$ and appropriate modification thereof could also be used in developing the theorems below.)

For Case I, we define $\delta_n(x)$ by (6) with

$$\alpha_n(x) = (v(x) - c)f_n(x) - f_n^{(1)}(x) \tag{21}$$

where $f_n(x)$ and $f_n^{(1)}(x)$ are given by (10) and (15) respectively.

Case II (Samuel [6]).

$$
\begin{aligned}
f_\lambda(x) &= \lambda^x \beta(\lambda) h(x) \qquad \text{for} \quad x > a , \qquad \lambda \quad \text{in a subinterval of} \quad (0, \infty) \\
&= 0 \qquad \text{otherwise} ,
\end{aligned}
$$

where a may be finite or infinite and $h(x) > 0$ for $x > a$.

Note that

$$\int \lambda f_\lambda(x)\, dG(\lambda) = w(x) f(x+1)$$

where

$$w(x) = \frac{h(x)}{h(x+1)} . \tag{22}$$

Then, by (2), we have

$$\alpha(x) = w(x) f(x+1) - cf(x) . \tag{23}$$

To obtain the empirical Bayes rule $\delta_n(x)$ in (6) define

$$\alpha_n(x) = w(x)f_n(x+1) - cf_n(x) \tag{24}$$

where $f_n(x)$ is given by (10) or (10′) with $\{h_n\}$ satisfying (11) and $K_1 \in \mathscr{K}_1$.

Next we shall develop some lemmas useful for later theorems on asymptotic optimality and convergence rates for our empirical Bayes rules for Cases I and II.

3. Some useful lemmas. Before giving results on asymptotic optimality and convergence rates, we give the following lemmas, the first of which is a restatement of Lemma 1 of [1] for the continuous case.

LEMMA 1. *With $r^*(G)$ and r_n given by* (5) *and* (7), *we have*

$$0 \leqq r_n - r^*(G) \leqq b \int |\alpha(x)| \Pr\{|\alpha_n(x) - \alpha(x)| \geqq |\alpha(x)|\}\, dx\,. \tag{25}$$

LEMMA 2. *In Case* I, *if G is any prior on the natural parameter space and if $h^{(r)}(x)$ exists and is continuous for all $x > a$, r a positive integer, then $f^{(r)}(x)$ exists and is continuous for all $x > a$. Furthermore, $f^{(1)}(x) = v(x)f(x) - \int \lambda f_\lambda(x)\, dG(\lambda)$, $v(x) = h^{(1)}(x)/h(x)$.*

PROOF. Observe that $f(x) = h(x)\xi(x)$, where $\xi(x) = \int e^{-\lambda x}\beta(\lambda)\, dG(\lambda)$. But the integral $\xi(x)$ is infinitely differentiable with repeated differentiation under the integral sign permissable by Theorem 2.9 in [2]. Hence, the differentiability of rth order of $f(x)$ follows from the differentiability of rth order of $h(x)$.

For Case II, the following lemma is also easily verified.

LEMMA 3. *In Case* II, *if $h(x)$ is continuous for all $x > a$, then $f(x)$ is continuous for all $x > a$. In Case* II, *if $h^{(r)}(x)$ exists and is continuous for all $x > a$, r a positive integer, then $f^{(r)}(x)$ exists and is continuous for all $x > a$.*

4. Asymptotic optimality of δ_n. The following theorem generalizes the examples and ideas incorporated in Section 5 of Robbins [5].

THEOREM 1. *In Case* I, *let δ_n be defined by* (6), (10) *or* (10′), (15) *and* (21) *with $K_1 \in \mathscr{K}_1$, $K_2 \in \mathscr{K}_2$ and $\{h_n\}$ satisfying* (11). *If $E|\Lambda| < \infty$ and $h^{(1)}(x)$ exists and is continuous for $x > a$, then $\lim_n r_n = r^*(G)$, that is, the empirical Bayes rule δ_n is asymptotically optimal* (see Robbins [5] page 198).

PROOF. Let $\beta_n(x) = |\alpha(x)| \Pr\{|\alpha_n(x) - \alpha(x)| \geqq |\alpha(x)|\}$. Since $\beta_n(x) \leqq |\alpha(x)|$ and $|\alpha(x)|$ is integrable by the assumption $E|\Lambda| < \infty$, it suffices by Lemma 1 and the bounded convergence theorem to show that $\lim_n \beta_n(x) = 0$ for all $x > a$. Now, by the Markov inequality,

$$\begin{aligned} \beta_n(x) &\leqq E|\alpha_n(x) - \alpha(x)| \\ &\leqq |v(x) - c|E|f_n(x) - f(x)| + E|f_n^{(1)}(x) - f^{(1)}(x)|\,. \end{aligned} \tag{26}$$

From Parzen [4], (Corollary 1 A and Theorem 2 A) for $K_1 \in \mathscr{K}_1$ and $\{h_n\}$ satisfying (11), the quadratic mean $E|f_n(x) - f(x)|^2 \to 0$ as $n \to \infty$ for all $x > a$

(continuity points of $f(x)$). Hence

$$\lim_n E|f_n(x) - f(x)| = 0 \qquad \text{for all} \quad x > a. \tag{27}$$

Next, by computing $\operatorname{Var}\{f_n^{(1)}(x)\}$, it is easy to show that

$$nh_n^3 \operatorname{Var}\{f_n^{(1)}(x)\} \to \{\tfrac{3}{2} \textstyle\int_0^{u_2} K_2^2(u)\,du - 2 \int_0^{u_2} K_2(u)K_2(2u)\,du\} f(x)$$

for all $x > a$. Also, since Lemma 2 implies $f^{(1)}(x)$ exists and is continuous for all $x > a$, we have by (20), $|Ef_n^{(1)}(x) - f^{(1)}(x)|^2 \to 0$ for all $x > a$. These two results imply, for $K_2 \in \mathscr{K}_2$ and $\{h_n\}$ satisfying (11), that $\lim_n E|f_n^{(1)}(x) - f^{(1)}(x)|^2 = 0$ for all $x > a$ and hence $\lim_n E|f_n^{(1)}(x) - f^{(1)}(x)| = 0$ for all $x > a$. This fact together with (27) and inequality (26) completes the proof of the desired result.

The following theorem for Case II is very similar to the results of Section 6 of Samuel [6] and is stated here without proof.

THEOREM 2. *In Case* II, *let* δ_n *be defined by* (6), (10) *or* (10′) *and* (24) *with* $K_1 \in \mathscr{K}_1$ *and the sequence* $\{h_n\}$ *satisfying* $h_n \downarrow 0$ *and* $nh_n \to \infty$. *If* $E\Lambda < \infty$ *and* $h(x)$ *is continuous for all* $x > a$, *then* $\lim_n r_n = r^*(G)$, *that is, the empirical Bayes rule* δ_n, *is asymptotically optimal.*

Examples of Theorems 1 and 2 will be given in Section 6 along with examples of the theorems proved in the next section.

5. Convergence rate theorems. The following two theorems give the main results of this paper concerning rates of convergence of r_n to the optimal Bayes risk. These theorems furnish rate results for slightly generalized versions of procedures first introduced by Robbins [5], Section 5 and Samuel [6], Section 6 for the two-action decision problem with continuous exponential families given by Cases I and II respectively. The conditions under which the theorems are stated are perhaps somewhat unintuitive and their content is more fully explored in the examples of Section 6.

THEOREM 3. *In Case* I, *let* δ_n *be defined by* (6), (10), (15) *and* (21) *with* $K_1 \in \mathscr{K}_1$, $K_2 \in \mathscr{K}_2$. *If* $E|\Lambda| < \infty$ *and* $h^{(r)}(x)$ *exists and is continuous for some integer* $r \geqq 2$, *and if for some* δ, $0 < \delta < 2$ *and some* $\varepsilon > 0$,

$$\int |\alpha(x)|^{1-\delta}\{f_\varepsilon^*(x)\}^{\delta/2}\,dx < \infty, \qquad \textit{where} \quad f_\varepsilon^*(x) = \sup_{0 \leqq t \leqq \varepsilon} f(x+t) \tag{3.1}$$

$$\int |\alpha(x)|^{1-\delta}[|v(x)|\{f_\varepsilon^*(x)\}^{\frac{1}{2}}]^\delta\,dx < \infty, \tag{3.2}$$

$$\int |\alpha(x)|^{1-\delta}\{q_\varepsilon^{(r)}(x)\}^\delta\,dx < \infty, \qquad \textit{where} \quad q_\varepsilon^{(r)}(x) = \sup_{0 < t \leqq \varepsilon} |f^{(r)}(x+t)| \tag{3.3}$$

$$\int |\alpha(x)|^{1-\delta}\{|v(x)|q_\varepsilon^{(r)}(x)\}^\delta\,dx < \infty, \tag{3.4}$$

then, choosing $h_n = O(n^{-(2r+1)^{-1}})$ *and* $K_i \in \mathscr{K}_i$ *such that* $\int u^{j+i-1}K_i(u)\,du = 0$ *for* $j = 1, \cdots, r-1$, $i = 1, 2$ *yields*

$$r_n - r^*(G) = O(n^{-(r-1)\delta/(2r+1)}).$$

PROOF. Let ε and δ be given. Applying the Markov inequality in (25), recalling (8) and (21), and using the c_r-inequality (Loève [2] page 155) with $r = \delta$, we have

$$r_n - r^*(G) \leqq b \int |\alpha(x)|^{1-\delta} E|\alpha_n(x) - \alpha(x)|^\delta \, dx \leqq A_n + B_n + C_n , \tag{28}$$

where

$$A_n = c_\delta^2 b|c| \int |\alpha(x)|^{1-\delta} E|f_n(x) - f(x)|^\delta \, dx ;$$
$$B_n = c_\delta^2 b \int |\alpha(x)|^{1-\delta} |v(x)|^\delta E|f_n(x) - f(x)|^\delta \, dx ;$$
$$C_n = c_\delta b \int |\alpha(x)|^{1-\delta} E|f_n^{(1)}(x) - f^{(1)}(x)|^\delta \, dx .$$

We shall find bounds for A_n, B_n and C_n. First, note that

$$E|f_n(x) - f(x)|^\delta \leqq [\mathrm{Var}\{f_n(x)\}]^{\delta/2} + |Ef_n(x) - f(x)|^\delta . \tag{29}$$

But, recalling (10)

$$\begin{aligned} \mathrm{Var}\{f_n(x)\} &= (nh_n^2)^{-1} \mathrm{Var}\left\{K_1\left(\frac{X-x}{h_n}\right)\right\} \\ &\leqq (nh_n)^{-1} \int K_1^2(u) f(x + h_n u) \, du \\ &\leqq (nh_n)^{-1} \{\int K_1^2(u) \, du\} f_\varepsilon^*(x) \qquad \text{if} \quad h_n u_1 \leqq \varepsilon . \end{aligned} \tag{30}$$

Also, by Lemma 2 since $f^{(r)}(x)$ exists and is continuous, a Taylor's series expansion of $f(x + h_n u)$ about x yields

$$\begin{aligned} |Ef_n(x) - f(x)| &= |\int \{f(x + h_n u) - f(x)\} K_1(u) \, du| \\ &\leqq \left| \sum_{j=1}^{r-1} \left\{ \frac{h_n^j \int u^j K_1(u) \, du}{j!} \right\} f^{(j)}(x) \right| \\ &\quad + \frac{h_n^r}{r!} |\int f^{(r)}(x + \zeta_n(x, u)) u^r K_1(u) \, du| , \end{aligned} \tag{31}$$

where $0 < \zeta(x, u) < h_n u \leqq h_n u_1$. Hence if $h_n u_1 \leqq \varepsilon$ our choice of K_1 making $\int u^j K_1(u) \, du = 0$ for $j = 1, \cdots, r-1$, yields

$$|Ef_n(x) - f(x)| \leqq q_\varepsilon^{(r)}(x) \frac{h_n^r \int_0^{u_1} u^r |K_1(u)| \, du}{r!} . \tag{32}$$

Thus, (30) and (32) substituted in (29) together with conditions (3.1) and (3.3) imply

$$\begin{aligned} A_n &= O((nh_n)^{-\delta/2}) + O(h_n^{r\delta}) \\ &= O(n^{-r\delta/(2r+1)}) \end{aligned} \tag{33}$$

where the second equality follows from our choice of h_n. Similarly, conditions (3.2) and (3.4) together with (28)—(30) imply

$$B_n = O(n^{-r\delta/(2r+1)}) . \tag{34}$$

To bound C_n, observe that as in (29),

$$E|f_n^{(1)}(x) - f^{(1)}(x)|^\delta \leqq [\mathrm{Var}\{f_n^{(1)}(x)\}]^{\delta/2} + |Ef_n^{(1)}(x) - f^{(1)}(x)|^\delta . \tag{35}$$

Recalling (15) and the c_2-inequality, we have

$$\operatorname{Var}\{f_n^{(1)}(x)\} = (nh_n^2)^{-1}\operatorname{Var}\left\{\frac{1}{2h_n}K_2\left(\frac{X-x}{2h_n}\right) - \frac{1}{h_n}K_2\left(\frac{X-x}{h_n}\right)\right\}$$

$$\leqq 2(nh_n^2)^{-1}\left\{E\left|\frac{1}{2h_n}K_2\left(\frac{X-x}{2h_n}\right)\right|^2 + E\left|\frac{1}{h_n}K_2\left(\frac{X-x}{h_n}\right)\right|^2\right\} \tag{36}$$

$$= 2(nh_n^3)^{-1}\int K_2^2(u)\{\tfrac{1}{2}f(x+2h_nu) + f(x+h_nu)\}\,du$$

$$\leqq 3(nh_n^3)^{-1}\{\int K_2^2(u)\,du\}f_\varepsilon^*(x) \qquad \text{if} \quad 2h_nu_2 \leqq \varepsilon\,.$$

Also, since $f^{(r)}(x)$ exists and is continuous, by our choice of K_2 we have as in (31),

$$|Ef_n^{(1)}(x) - f^{(1)}(x)| = \left|\frac{1}{h_n}\int K_2(u)\{f(x+2h_nu) - f(x+h_nu)\}\,du - f^{(1)}(x)\right|$$

$$\leqq h_n^{r-1}\frac{2^r+1}{r!}\int_0^{u_2} u^r|K_2(u)|\{|f^{(r)}(x+\zeta_n'(x,u))|$$

$$+ |f^{(r)}(x+\zeta_n''(x,u))|\}\,du$$

where $0 < \zeta_n'(x,u) < 2h_nu \leqq 2h_nu_2$ and $0 < \zeta_n''(x,u) < h_nu \leqq h_nu_2$. Therefore, if $2h_nu_2 \leqq \varepsilon$ with $a_r = 2(r!)^{-1}(2^r+1)\int u^r|K_2(u)|\,du$, we have

$$E|f_n^{(1)}(x) - f^{(1)}(x)| \leqq a_r h_n^{r-1} q_\varepsilon^{(r)}(x)\,. \tag{37}$$

Thus, from (35), (36), (37) and conditions (3.1) and (3.3) of the theorem we have

$$C_n = O((nh_n^3)^{-\delta/2}) + O(h_n^{(r-1)\delta}) = O(n^{-\delta(r-1)/(2r+1)}) \tag{38}$$

where again the choice of h_n furnishes the second equality.

From (33), (34) and (38) we see that C_n is the dominating term of the bound in (28) and since $r_n - r^*(G) \geqq 0$, the proof of the theorem is complete.

By essentially the same arguments as were used for the A_n and B_n terms of Theorem 3, it can be shown with the aid of Lemma 3 that a similar result holds for Case II for all $r \geqq 1$. We state this result without proof.

THEOREM 4. *In Case* II, *let* δ_n *be defined by* (6), (10) *and* (24) *with* $K_1 \in \mathscr{K}_1$. *If* $E\Lambda < \infty$ *and* $h^{(r)}(x)$ *exists and is continuous for some integer* $r \geqq 1$, *and if for some* δ, $0 < \delta < 2$, *and some* $\varepsilon > 0$.

$$\int |\alpha(x)|^{1-\delta}\{f_\varepsilon^*(x)\}^{\delta/2}\,dx < \infty\,, \qquad f_\varepsilon^*(x) = \sup_{0\leqq t\leqq\varepsilon} f(x+t) \tag{4.1}$$

$$\int |\alpha(x)|^{1-\delta}[w(x)\{f_\varepsilon^*(x+1)\}^{\frac{1}{2}}]^\delta\,dx < \infty \tag{4.2}$$

$$\int |\alpha(x)|^{1-\delta}\{q_\varepsilon^{(r)}(x)\}^\delta\,dx < \infty\,, \qquad q_\varepsilon^{(r)}(x) = \sup_{0\leqq t\leqq\varepsilon}|f^{(r)}(x+t)| \tag{4.3}$$

$$\int |\alpha(x)|^{1-\delta}\{w(x)q_\varepsilon^{(r)}(x+1)\}^\delta\,dx < \infty \tag{4.4}$$

then, choosing $h_n = O(n^{-(2r+1)^{-1}})$ *and* $K_1 \in \mathscr{K}_1$ *such that* $\int u^jK_1(u)\,du = 0, j = 1,\cdots,$ $r-1$ *yields,*

$$r_n - r^*(G) = O(n^{-\delta r/(2r+1)})\,.$$

6. Examples and moment condition corollaries when $\delta \leqq 1$. In this section we give two examples which illustrate quite clearly the meaning and usefulness of

Theorems 3 and 4 when $\delta \leqq 1$. For other examples to which our results could be applied see the papers of Robbins [5] and Samuel [6]. For an example when $\delta > 1$ see Section 7.

EXAMPLE 1. (Negative Exponential). Consider the family of negative exponential densities given in the Case I parameterization by

$$f_\lambda(x) = \lambda e^{-\lambda x} \qquad \lambda > 0\,, \quad x > 0 \tag{39}$$
$$= 0 \qquad \text{otherwise}\,,$$

where $\beta(\lambda) = \lambda$ and $h(x) = 1$ or 0 as $x > 0$ or $x \leqq 0$.

Observe the following facts about $f(x) = \int f_\lambda(x)\, dG(\lambda)$ when $f_\lambda(x)$ is given by (39).

$$f_\varepsilon^*(x) = \sup_{0 \leqq t \leqq \varepsilon} f(x+t) = f(x) \qquad \text{for all} \quad x \quad \text{and} \quad \varepsilon > 0\,. \tag{40}$$

$$f^{(r)}(x) \quad \text{exists for all} \quad x > 0 \quad \text{by Lemma 2}\,. \tag{41}$$

Furthermore, $|f^{(r)}(x)| = \int \lambda^{r+1} \exp\{-\lambda x\}\, dG(\lambda)$ and $|q_\varepsilon^{(r)}(x)| = \sup_{0 \leqq t \leqq \varepsilon} |f^{(r)}(x + t)| = |f^{(r)}(x)|$ for all x and $\varepsilon > 0$.

$$|f^{(r)}(x)| \leqq r!\, f(x-1) \qquad \text{for} \quad x > 1 \quad \text{since} \quad \exp\{-\lambda\}\lambda^r \leqq r!\,. \tag{42}$$

Also, since $v(x) = h^{(1)}(x)/h(x) = 0$, we have by (8) that $\alpha(x) = -cf(x) - f^{(1)}(x)$ and by (44) for $x > 1$

$$|\alpha(x)| \leqq cf(x) + f(x-1) \leqq (c+1)f(x-1)\,. \tag{43}$$

We now combine these facts to give the following consequence of Theorem 3.

COROLLARY 3.1. *For the negative exponential family* (39), *Theorem* 3 *holds for integer* $r \geqq 2$ *and* $0 < \delta \leqq 1$ *if*

$$E\Lambda^{r+1} < \infty \tag{3.1.1}$$

$$E\Lambda^{-(1+t)\delta/(2-\delta)} < \infty \qquad \text{for some} \quad t > 0\,. \tag{3.1.2}$$

PROOF. Since $v(x) \equiv 0$ Conditions (3.2) and (3.4) of Theorem 3 are trivially satisfied. To verify Conditions (3.3) note that by (41), (42) and (43), we have

$$\int_{x>1} |\alpha(x)|^{1-\delta}\{q_\varepsilon^{(r)}(x)\}^\delta\, dx \leqq r!(c+1) \int_1^\infty f(x-1)\, dx = r!(c+1)\,. \tag{44}$$

But, since $|\alpha(x)| \leqq cf(x) + f^{(1)}(x) \leqq cE\Lambda + E\Lambda^2 < \infty$ and by (41), $|q_\varepsilon^{(r)}(x)| \leqq |f^{(r)}(x)| < E\Lambda^{r+1} < \infty$, the integrand in Condition (3.3) of Theorem 3 is bounded for $0 < x \leqq 1$ and zero for $x \leqq 0$. Thus, (44) completes the verification of (3.3).

Finally, to verify (3.1) of Theorem 3 it suffices by (40) and (43) and the fact that the integrand of (3.1) is bounded by $(cE\Lambda + E\Lambda^2)^{1-\delta}(E\Lambda)^{\delta/2}$, for $0 < x \leqq 2$, to show

$$\int_1^\infty \{f(x)\}^{1-\delta/2}\, dx < \infty\,. \tag{45}$$

But with $\eta = (1+t)\delta/(2-\delta)$, the Hölder inequality implies

$$\int_1^\infty \{f(x)\}^{1-\delta/2}\, dx \leqq \{\int_1^\infty x^{-(1+t)}\, dx\}^{\delta/2}\{\int_0^\infty x^\eta f(x)\, dx\}^{1-\delta/2}$$
$$= \{\int_1^\infty x^{-(1+t)}\, dx\}^{\delta/2}\{E\Lambda^{-\eta}\Gamma(\eta+1)\}^{1-\delta/2}$$

where $\Gamma(\cdot)$ is the gamma function and the equality follows by Fubini's theorem applied to the second factor. Thus by (3.1.2), (45) is verified completing the proof of the corollary.

Corollary 3.1 illustrates that the conditions of Theorem 3 correspond to existence of certain moments on the prior distribution $G(\lambda)$. For example, if $G(\lambda)$ is a prior distribution having a Lebesgue density for some $t > 0$ and $\alpha > 0$

$$\begin{aligned} \xi(\lambda) &= \lambda^t e^{-\alpha\lambda} q(\lambda) \qquad && \lambda > 0 \\ &= 0 && \text{otherwise} \end{aligned}$$

where $q(\lambda)$ is bounded on $(0, \infty)$, then $\delta = 1$ and r may be taken arbitrarily large in Corollary 3.1. Thus, the rate for this case may be made arbitrarily close algebraically to $O(n^{-\frac{1}{2}})$, that is, the rate $O(n^{-\frac{1}{2}+\beta})$ is attainable for any $\beta > 0$ by a suitable choice of K_1 and K_2. For further analysis with $\delta > 1$ for a prior in the above class see Section 7.

By a quite similar analysis, here omitted, we give a corollary to Theorem 4 which is attainable for the negative exponential family under the Case II parameterization. Under Case II, $\beta(\lambda) = -\log \lambda$ for $0 < \lambda < 1$ and $h(x) = 1$ for $x > 0$, so that

$$\begin{aligned} f_\lambda(x) &= \lambda^x(-\log \lambda) \qquad && 0 < \lambda < 1 , \quad x > 0 \\ &= 0 && \text{otherwise} . \end{aligned} \tag{46}$$

Under this representation, we have

COROLLARY 4.1. *For the family* (46), *Theorem* 4 *holds with integer* $r \geqq 1$ *and* $0 < \delta \leqq 1$ *provided*

$$E(-\log \Lambda)^{r+1} < \infty , \qquad \textit{and} \tag{4.1.1}$$

$$E(-\log \Lambda)^{-(1+t)\delta/(2-\delta)} < \infty \qquad \textit{for some} \quad t > 0 . \tag{4.1.2}$$

EXAMPLE 2. (Normal, Unknown Mean). We shall now apply Theorems 3 and 4 to the family of normal densities given by

$$f_\theta(x) = (2\pi)^{-\frac{1}{2}} e^{-\frac{1}{2}(x-\theta)^2} , \qquad -\infty < x < \infty , \quad -\infty < \theta < \infty . \tag{47}$$

Under Case I, with $\lambda = -\theta$, $\beta(\lambda) = (2\pi)^{-\frac{1}{2}} \exp\{-\frac{1}{2}\lambda^2\}$, $h(x) = \exp\{-\frac{1}{2}x^2\}$, we have

$$f_\lambda(x) = e^{-\lambda x}\beta(\lambda)h(x) , \qquad -\infty < x < \infty , \quad -\infty < \lambda < \infty . \tag{48}$$

Observe the following facts about $f_\lambda(x)$ and $f_\varepsilon^*(x)$. If $0 \leqq t \leqq \varepsilon$, then $f_\lambda(x + t) \leqq f_\lambda(x) + \exp\{\frac{1}{2}\varepsilon^2\} f_\lambda(x + \varepsilon)$. Therefore, if $\exp\{\frac{1}{2}\varepsilon^2\} \leqq 2$, we see that

$$f_\lambda(x + t) \leqq f_\lambda(x) + 2f_\lambda(x + \varepsilon) \qquad \text{if} \quad 0 \leqq t \leqq \varepsilon \tag{49}$$

and,

$$\begin{aligned} f_\varepsilon^*(x) &\leqq \sup\nolimits_{0 \leqq t \leqq \varepsilon} f(x + t) \\ &\leqq f(x) + 2f(x + \varepsilon) . \end{aligned} \tag{50}$$

By Lemma 2, $f^{(r)}(x)$ can be computed by repeated differentiation under the integral sign in $f(x) = \int f(x)\, dG(\lambda)$ with the result that

$$f^{(r)}(x) = (-1)^r \int H_r(x+\lambda) f_\lambda(x)\, dG(\lambda)\,, \tag{51}$$

where $H_r(\cdot)$ is the rth Hermite polynomial.

Another useful fact to note is that since $v(x) = h^{(1)}(x)/h(x) = -x$, we have by (8) and (51) that $\alpha(x) = -(x+c)f(x) - \int (x+\lambda) f_\lambda(x)\, dG(\lambda)$ and hence

$$|\alpha(x)| \leqq (|c| + |x|)f(x) + \int |x+\lambda| f_\lambda(x)\, dG(\lambda)\,. \tag{52}$$

We are now in a position to verify the following consequence of Theorem 3.

COROLLARY 3.2. *For the normal family* (47), *Theorem* 3 *holds with integer* $r \geqq 2$, *and* $0 < \delta \leqq 1$ *provided*

$$E|\Lambda|^r < \infty\,, \qquad \textit{and} \tag{3.2.1}$$

$$E|\Lambda|^{1+\delta(3+t)/(2-\delta)} < \infty \qquad \textit{for some} \quad t > 0\,. \tag{3.2.2}$$

PROOF. Since $|v(x)| = |x|$, the integrability conditions (3.1) and (3.3) are implied by (3.2) and (3.4) respectively. To verify (3.2), let $0 < \varepsilon \leqq 1$ and $\delta > 0$ be given where $\exp\{\frac{1}{2}\varepsilon^2\} \leqq 2$. Then by defining

$$y(x) = (c' + |x|)f(x) + \int |x+\lambda| f_\lambda(x)\, dG(\lambda) + 2f(x+\varepsilon)$$

with $c' = \max\{1, |c|\}$ and noting that $f_\varepsilon^*(x) \leqq y(x)$ by (52) and $|\alpha(x)| \leqq y(x)$ by (54), we have (3.2) holding if $\int \{y(x)\}^{1-\delta/2}|x|^\delta\, dx < \infty$. Equivalently, (3.2) holds if

$$\int_{|x|\geqq 1} \{y(x)\}^{1-\delta/2}|x|^\delta\, dx < \infty\,. \tag{53}$$

To verify (53), note that by the Hölder inequality,

$$\int_{|x|\geqq 1} \{y(x)\}^{1-\delta/2}|x|^\delta\, dx \leqq c^*\{\int y(x)|x|^\beta\, dx\}^{1-\delta/2} \tag{54}$$

where $\beta = \beta(t) = \delta(3+t)/(2-\delta)$ and $c^* = c^*(t) = \int_{|x|\geqq 1} |x|^{-(1+t)}\, dx$ for some $t > 0$.

Now let $\mu_u = (2\pi)^{-\frac{1}{2}} \int |x|^u \exp\{-\frac{1}{2}x^2\}\, dx$ for $u > 0$. Note that by the c_r-inequality with $r = \beta$, $2\int |x|^\beta f(x+\varepsilon)\, dx = 2^\beta \int |x|^\beta f(x)\, dx + (2\varepsilon)^\beta$ and by Fubini's theorem we have

$$\begin{aligned} \int y(x)|x|^\beta\, dx &\leqq \int\int \{(c' + 2^\beta)|x|^\beta + |x|^{1+\beta} + |x+\lambda|\,|x|^\beta + (2\varepsilon)^\beta\} f_\lambda(x)\, dx dG(\lambda) \\ &\leqq 2^{\beta-1}(c' + 2^\beta)(\mu_\beta + E|\Lambda|^\beta) + 2^\beta(\mu_{1+\beta} + E|\Lambda|^{1+\beta}) \\ &\quad + 2^{\beta-1}(\mu_{1+\beta} + E|\Lambda|^\beta) + (2\varepsilon)^\beta\,, \end{aligned} \tag{55}$$

where the second inequality is by the c_r-inequality with $r = \beta$ and $1 + \beta$. From this inequality we see that (3.2.2) implies $\int y(x)|x|^\beta dx < \infty$ from which (54) implies (53), thus verifying Condition (3.2) of Theorem 3 for this example.

Next, we need to verify (3.4) of Theorem 3. With a_j^r as the jth coefficient of the rth Hermite polynomial, we have by (51) that

$$\begin{aligned} |f^{(r)}(x+t)| &\leqq \int |H_r(x+\lambda+t)| f_\lambda(x+t)\, dG(\lambda) \\ &\leqq \textstyle\sum_{j=0}^r |a_j^r| \int |x+\lambda+t|^j f_\lambda(x+t)\, dG(\lambda) \\ &\leqq \textstyle\sum_{j=0}^r |a_j^r| 2^{j-1}\{\int |x+\lambda|^j f_\lambda(x+t)\, dG(\lambda) + t^j f(x+t)\} \end{aligned} \tag{56}$$

where the last inequality is by the c_r-inequality. Note that by (49) and (50) and basic moment inequalities the right-hand side of (56) is bounded by a linear combination of integrals of the form $\int |x + \lambda|^j (f_\lambda(x) + 2f_\lambda(x + \varepsilon))\, dG\ j = 0, \cdots, r$. Therefore, to verify (3.4) with $|v(x)| = |x|$ from (56) and the fact that $|x + \lambda|^j \leqq |x + \lambda|^r + 1, j = 0, \cdots, r$, it suffices to show that

$$\int \{\alpha(x)\}^{1-\delta} |\alpha^*(x)|^\delta |x|^\delta\, dx < \infty \tag{57}$$

where $\alpha^*(x) = \int (|x + \lambda|^r + 1)(f_\lambda(x) + 2f_\lambda(x + \varepsilon))\, dG(\lambda)$. But noting that $\alpha^*(x) \leqq y^*(x)$ and $|\alpha(x)| \leqq y^*(x)$ where

$$y^*(x) = \int \{(|c| + 1 + |x| + |x + \lambda|^r) f_\lambda(x) + 2(|x + \lambda|^r + 1) f_\lambda(x + \varepsilon)\}\, dG(\lambda)$$

we see that (57) holds if

$$\int |x|^\delta y^*(x)\, dx < \infty\ . \tag{58}$$

By using Fubini's theorem and the c_r-inequality repeatedly in (58) one can bound, as in (55), the integral in (58) by $c_1 + c_2 E|\Lambda|^\delta + c_3 E|\Lambda|^{1+\delta}$, where c_1, c_2 and c_3 are certain finite constants depending on absolute central moments of the standard normal density. Thus, (57) holds if $E|\Lambda|^{1+\delta} < \infty$, which is always true under (3.2.1). This completes verification of (57) and hence (3.4) of Theorem 3 for this normal distribution example.

As in the previous example, a similar analysis yields the following corollary to Theorem 4 which is attainable for the normal family under Case II. In Case II in (47) let $\lambda = \exp\{\theta\}$, $\beta(\lambda) = (2\pi)^{-\frac{1}{2}} \exp(-\log \lambda)^2$, $h(x) = \exp\{-\frac{1}{2}x^2\}$. Then,

$$f_\lambda(x) = \lambda^x \beta(\lambda) h(x)\ , \qquad 0 < \lambda < \infty\ , \quad -\infty < x < \infty\ . \tag{59}$$

Under this representation, we have

COROLLARY 4.2. *For the family* (59), *Theorem* 4 *holds with integer* $r \geqq 1$ *and* $0 < \delta \leqq 1$ *provided*

$$E|\log \Lambda|^r < \infty\ , \qquad E\Lambda < \infty\ , \qquad \textit{and} \tag{4.2.1}$$

$$E|\log \Lambda|^{1+\delta(3+t)/(2-\delta)} < \infty \qquad \textit{for some} \quad t > 0\ . \tag{4.2.2}$$

7. Examples with rate algebraically close to $O(n^{-1})$. We present now example in the negative exponential case and the normal case in which the δ of Theorem 3 satisfies $1 < \delta < 2$ and can be arbitrarily close to 2 depending on the prior involved while r in the theorem can also be taken arbitrarily large. This illustrates that the rate of convergence of the empirical Bayes procedures in some continuous cases can be arbitrarily close (algebraically) to $O(n^{-1})$, that is, can be $O(n^{-1+\varepsilon})$ for $\varepsilon > 0$ and arbitrarily small.

Specifically, let $f_\lambda(x)$ be the negative exponential density given by (39) and let $G(\lambda)$ be a prior distribution having a Lebesgue density,

$$\xi(\lambda) = \frac{\alpha^{1+t} \lambda^t e^{-\alpha\lambda}}{\Gamma(1 + t)}\ , \qquad \text{where} \quad t > 0\ , \quad \alpha > 0 \tag{60}$$

and $\Gamma(\cdot)$ is the gamma function. With this prior on λ we obtain

$$f(x) = \alpha^{1+t}(1+t)(x+\alpha)^{-(2+t)} \quad \text{and} \tag{61}$$

$$f^{(r)}(x) = \frac{(-1)^r \alpha^{1+t}\Gamma(r+2+t)}{\Gamma(1+t)(x+\alpha)^{r+2+t}}.$$

From (61), it can be shown that

$$\int |\alpha(x)|^{1-\delta}\{f(x)\}^{\delta/2}\,dx = a_1 \int_0^\infty |a_2 - cx|^{1-\delta}(x+\alpha)^{2\delta-3-t+(\delta t)/2}\,dx \tag{62}$$

where $a_1 = [\alpha^{(1+t)}(1+t)]^{1-\delta/2}$ and $a_2 = 2 + t - c\alpha$. By examining the behavior of the integrand in (62) as $x \to a_2/c$ and as $x \to \infty$, we see that (62) is finite if $\delta < 2$ and if $\delta - 1 - 2\delta + 3 + t - (\delta t)/2 > 1$, or equivalently if $\delta < 2(1 + t)/(2 + t)$. A similar analysis using (41) and (61) reveals that (3.3) of Theorem 3 is true for this example if $\delta < 2$. We summarize as follows.

COROLLARY 3.3. *For the negative exponential family* (39) *with prior distribution* (60) *on* $(0, \infty)$, *Theorem* 3 *holds for any integer* $r \geqq 2$ *and any* $\delta < 2(1+t)/(2+t)$, *where t is the parameter of the conjugate gamma prior* (60).

In the normal case with unknown mean, a similar detailed analysis (here omitted) yields

COROLLARY 3.4. *For the normal family* (47) *and* (48) *with conjugate normal prior distribution in* λ *having Lebesgue density on* $(-\infty, \infty)$ *given by*

$$\xi(\lambda) = (2\pi\sigma_0^2)^{-\frac{1}{2}} \exp\{-(\tfrac{1}{2}\sigma_0^2)(\lambda - \lambda_0)^2\},$$

Theorem 3 *holds for any integer* $r \geqq$ *and any* δ, $0 < \delta < 2$.

Corollaries 3.3 and 3.4 thus exhibit examples for which a rate algebraically close to $O(n^{-1})$ may be obtained by choosing K_1 and K_2 suitably for certain specific priors.

8. Concluding remarks. This paper has demonstrated that the difference between the actual $(n+1)$st stage risk of the empirical Bayes rule and the optimal Bayes risk is $O(n^{-\beta})$, where β, $0 < \beta < 1$, depends upon appropriate conditions in Theorems 3 and 4, for certain continuous exponential distributions in certain two-action decision problems. The true convergence rate depends on the nature of the unknown prior distribution $G(\lambda)$. Section 6 illustrates by examples that for rates up to $\beta \leqq \frac{1}{2}$, the conditions of the theorems reduce to the existence of certain moments on the prior distribution. Section 7 gave examples where β could be arbitrarily close to 1. Other examples of this type can be easily constructed with other exponential families and various priors. However, since the conditions of Theorems 3 and 4 with the factor $|\alpha(x)|^{1-\delta}$ when $\delta > 1$ are hard to verify, general corollaries as those in Section 6 are difficult to obtain for $\delta > 1$.

In this regard, we wish to stress that although the actual rate may depend on $G(\lambda)$ the *construction* of the procedure given by (6), (10), (15) and (21) for Case I and (6), (10) and (24) for Case II *only* depend on choosing $\{h_n\}$ and the kernels

$K_1(u)$ and/or $K_2(u)$. For the judicious choice of these dictated by Theorems 3 and 4 we need only know the differentiability properties of $h(x)$, a specified known function for a given exponential family, and the assumed moment properties $E\Lambda^r < \infty$ (Theorem 3) or $E|\log \Lambda|^r < \infty$ (Theorem 4). Hence, in general, depending on what moment assumptions one is willing to make on the prior $G(\lambda)$, Theorems 3 and 4 dictate the procedure to be chosen. It must be remarked however that the *actual* convergence rate truly depends on the unknown prior.

Clearly the methods of this paper are easily extendable to give results for two-tail testing problems and certain polynomial loss functions as examined by Samuel in [6].

It is our hope that this and the previous paper [1] have shed light on the more exact asymptotic behavior of certain empirical Bayes procedures in two-action problems. Specifically, we have examined extensively the question of rate of convergence to optimality raised by Robbins in [5].

Acknowledgment. The authors wish to thank the referee for the improved versions of Lemmas 2 and 3 which removed additional assumptions in an earlier version.

REFERENCES

[1] 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.

[2] LEHMAN, E. L. (1959). *Testing Statistical Hypotheses*. Wiley, New York.

[3] LOÈVE, M. (1963). *Probability Theory*, 3rd ed. Van Nostrand, Princeton.

[4] PARZEN, E. (1962). On estimation of a probability density function and mode. *Ann. Math. Statist.* **33** 1065–1076.

[5] ROBBINS, H. (1963). The empirical Bayes approach to testing statistical hypotheses. *Rev. Inst. Internat. Statist.* **31** 195–208.

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

[7] SCHUSTER, EUGENE F. (1969). Estimation of a probability density function and its derivatives. *Ann. Math. Statist.* **40** 1187–1195.