Ann. Inst. Statist. Math.
28 (1976), Part A, 389-397

# RATES IN THE EMPIRICAL BAYES ESTIMATION PROBLEM WITH NON-IDENTICAL COMPONENTS

## —CASE OF NORMAL DISTRIBUTIONS—

THOMAS E. O'BRYAN AND V. SUSARLA

(Received Aug. 4, 1973)


## Abstract

This paper considers empirical Bayes estimation of the mean $\theta$ of the univariate normal density $f_\theta$ with known variance where the sample sizes $m(n)$ may vary with the component problems but remain bounded by $\bar{m}<\infty$. Let $\{(\theta_n, X_n=(X_{n,1},\cdots, X_{n,m(n)}))\}$ be a sequence of independent random vectors where the $\theta_n$ are unobservable and iid $G$ and, given $\theta_n=\theta$, $X_n$ has density $f_\theta^{m(n)}$. The first part of the paper exhibits estimators for the density of $\sum_{j=1}^{m(n)} X_{n,j}$ and its derivative whose mean-squared errors go to zero with rates $O(n^{-1/\bar{m}}\log n)$ and $O(n^{-1/\bar{m}}(\log n)^2)$ respectively. Let $R^{m(n+1)}(G)$ denote the Bayes risk in the squared-error loss estimation of $\theta_{n+1}$ using $X_{n+1}$. For given $0<a<1$, we exhibit $t_n(X_1,\cdots, X_n; X_{n+1})$ such that $D(t_n, G)=\mathrm{E}\,[(t_n-\theta_{n+1})^2]-R^{m(n+1)}(G)\leqq c_1(a, \bar{m})(\log n)^2\cdot n^{-a/((2+a)\bar{m})}$ for $n>1$ under the assumption that the support of $G$ is in $[0, 1]$. Under the weaker condition that $\mathrm{E}\,[|\theta|^{2+\gamma}]<\infty$ for some $\gamma>0$, we exhibit $t_n^*(X_1,\cdots, X_n; X_{n+1})$ such that $D(t_n^*, G)\leqq c_2(\bar{m}, \gamma)(\log n)^{-\gamma/(2+\gamma)}$ for $n>1$.


## 1. Introduction

In this paper a variant of the standard empirical Bayes decision problem is considered where the sequence of component problems are not identical in that the sample sizes may vary with the component problems. For a general discussion on such problems and their motivation, see O'Bryan [3]. O'Bryan and Susarla [4] considered the above problem for squared-error loss estimation in certain continuous exponential families including normal distributions with known variance. The drawbacks associated with the results of the above papers are (1) the empirical Bayes procedures exhibited there are difficult to calculate explicitly since they involve inversions of null sequences with divergent sequences and (2), a priori, it seems difficult to obtain any rate of risk convergence results for the procedures. The purpose of this paper is to exhibit empirical Bayes procedures having some rate of risk convergence properties when the component problem is the squared-error loss estimation of $\theta$, where $\theta$, a random variable, has an unknown distribution $G$ on some subset $\Theta$ of the real line, based on (given $\theta$) conditionally independent observations $x_1, \cdots, x_m$ from $f_\theta$, the univariate normal density with mean $\theta$ and variance unity, where $m$ may vary with the component problems. In so doing, estimates for a density and its derivative are obtained having better rates of mean square error convergence than have been obtained by using either the kernel estimates of Parzen [5] or of Johns and Van Ryzin [1].

In the component problem with sample size $m$, let $R^m(t, G)$ denote the Bayes risk of the estimator $t$ against $G$ and $R^m(G) = \inf_t R^m(t, G)$. In the modified empirical Bayes estimation problem, there is a sequence $\{(\theta_j, \boldsymbol{X}_j = (X_{j,1}, \cdots, X_{j,m(j)}))\}$ of independent random vectors where the unobservable $\theta_j$ are iid $G$ and, for $\theta_j = \theta$, $X_{j,1}, \cdots, X_{j,m(j)}$ are iid with density $f_\theta$. With $t_n(\boldsymbol{X}_{n+1}) = t_n(\boldsymbol{X}_1, \cdots, \boldsymbol{X}_n; \boldsymbol{X}_{n+1})$ an estimator for use in the $(n+1)$st problem, its risk conditional on $\boldsymbol{X}_1, \cdots, \boldsymbol{X}_n$ is given by

$$R^{m(n+1)}(t_n, G) \equiv \int\int\int (\theta - a)^2 t_n(\boldsymbol{x}, da) f_\theta^{m(n+1)}(\boldsymbol{x}) d\boldsymbol{x} dG(\theta) \tag{1.1}$$

where, for each $\boldsymbol{x}$, $t_n(\boldsymbol{x}, \cdot)$ is a probability measure on a $\sigma$-field of subsets of $\Theta$ containing all singleton sets of $\Theta$. With $R_n(t_n, G)$ denoting the overall expectation, we have

$$R_n(t_n, G) \equiv \mathrm{E}\,[R^{m(n+1)}(t_n, G)] \geqq R^{m(n+1)}(G) \tag{1.2}$$

for any $t_n$ which motivates the following definition.

DEFINITION. A sequence of rules $\{t_n\}$ is said to be asymptotically optimal with order $g(n)$ (denoted hereafter by a.o. $(g(n))$) for some function $g(n) \to 0$ as $n \to \infty$ if, for some constant $c > 0$,

$$0 \leqq R_n(t_n, G) - R^{m(n+1)}(G) \leqq c g(n)\ . \tag{1.3}$$

In the component problem of size $m$, a non-randomized Bayes estimator based on the sum $y$ of the $m$ observations is given by

$$t_G^m(y) = m^{-1} y + (p'(y)/p(y)) \tag{1.4}$$

where $(2\pi m)^{1/2} p(y) = \int \exp\{-(2m)^{-1}(y - m\theta)^2\} dG(\theta)$. It is well-known that if $\theta$ and $t_n$ are $L_2$ random variables, then

$$R_n(t_n, G) - R^{m(n+1)}(G) = \mathrm{E}\,[(t_n - t_G^{m(n+1)})^2] \tag{1.5}$$

so that in order to define $\{t_n\}$ satisfying (1.3), it seems reasonable to define $t_n$ to be a natural estimate of $t_G^{m(n+1)}$. In view of (1.4), such a $t_n$ can be defined by taking the ratio of estimators of $p$ and $p'$. The definition of these estimators and some of their properties are the subject matters of the next section. Sections 3 and 4 deal with the estimation problem under the assumptions that $G$ has compact support and $\mathrm{E}\,[(|\theta|^{2+\gamma})]<\infty$ for some $\gamma>0$ respectively and exhibit procedures $\{t_n\}$ satisfying (1.3) for some $g(n)$.

Throughout this paper, all integrals without limits are taken to be over $(-\infty, \infty)$ and all limits are as $n\to\infty$. P and E denote probability and expectation taken over all random variables on which they operate.

## 2. Estimation of a density and its derivative

For $j=1,\cdots,n+1$, let

$$Y_j=\sum_{l=1}^{m(j)} X_{j,l}\ , \tag{2.1}$$

$$ {}_jp(y)=(2\pi m(j))^{-1/2}\int \exp\{-(2m(j))^{-1}(y-m(j)\theta)^2\}dG(\theta)\ , \tag{2.2}$$

and, therefore, by Fubini's theorem,

$$\varphi_j(t)=\mathrm{E}\,[\exp\{itY_j\}]=\exp\{-t^2m(j)/2\}\int \exp\{itm(j)\theta\}dG(\theta)\ . \tag{2.3}$$

Abbreviating $m(n+1)$ by $m$, ${}_{n+1}p$ by $p$, and $\varphi_{n+1}$ by $\varphi$ for simplicity, we consider the problem of estimation of $p$ and its derivative $p'$ using $Y_1,\cdots,Y_n$. Since $\int|\varphi(t)|dt<\infty$, we obtain (Loéve [2], p. 188) that

$$p(y)=(2\pi)^{-1}\int \exp\{-ity\}\varphi(t)dt\ . \tag{2.4}$$

It immediately follows from (2.3) and (2.4) that, for any $0<M<\infty$,

$$\begin{aligned}&\left|2\pi p(y)-\int_{-M}^{M}\exp\{-ity\}\varphi(t)dt\right|\\ &\qquad\leqq 2\int_M^\infty|\varphi(t)|dt=2(2\pi/m)^{1/2}\Phi(-m^{1/2}M)\end{aligned} \tag{2.5}$$

where $\Phi$ is the standard normal distribution function. Since, for large $M$, the rhs of (2.5) is small, one can consider estimating $\int_{-M}^{M}\exp\{-ity\}\cdot\varphi(t)dt$ with large $M$ in order to estimate $p(y)$. Since

$$\varphi(t)=\exp\{mt^2((m/m(j))-1)/2\}\varphi_j(mt/m(j)) \tag{2.6}$$

for $j=1,\cdots,n$, a simple unbiased estimate of $\int_{-M}^{M}\exp\{-ity\}\varphi(t)dt$ based on $Y_1,\cdots,Y_n$ is

$$\hat{p}_{n,M}(y)=(n\pi)^{-1}\sum_{j=1}^{n}\int_0^M \exp\{mt^2((m/m(j))-1)/2\} \cdot \cos t((mY_j/m(j))-y)dt\ . \tag{2.7}$$

Since the $j$th summand on the rhs of (2.7) is bounded (in absolute value) by $\int_0^M \exp\{mt^2((m/m(j))-1)/2\}dt$, the independence of $Y_1,\cdots,Y_n$ leads to

$$\mathrm{Var}\,(\hat{p}_{n,M}(y))\leqq(n\pi)^{-2}\sum_{j=1}^{n}\left\{\int_0^M \exp\{mt^2((m/m(j))-1)/2\}dt\right\}^2 . \tag{2.8}$$

Combining (2.5) and (2.8) yields that

$$\sup_y \mathrm{E}\,[(\hat{p}_{n,M}(y)-p(y))^2]\leqq\text{rhs of }(2.8)+2(m\pi)^{-1}\Phi^2(-m^{-1/2}M)\ . \tag{2.9}$$

For estimating $p'$, observe that since $\int|t\varphi(t)|dt<\infty$,

$$p'(y)=(2\pi)^{-1}\int(-it)\exp\{-ity\}\varphi(t)dt\ . \tag{2.10}$$

Following the same procedure as in the case of estimation of $p$, one can show that, for any $0<N<\infty$,

$$\hat{p}'_{n,N}(y)=(n\pi)^{-1}\sum_{j=1}^{n}\int_0^N t\exp\{mt^2((m/m(j))-1)/2\} \cdot \sin t((mY_j/m(j))-y)dt \tag{2.11}$$

is an unbiased estimate of the integral in (2.10) truncated to $(-N, N)$ and that

$$\sup_y \mathrm{E}\,[(\hat{p}'_{n,N}(y)-p'(y))^2] \leqq(n\pi)^{-2}\sum_{j=1}^{n}\left\{\int_0^N t\exp\{mt^2((m/m(j))-1)/2\}dt\right\}^2 +(\pi m)^{-2}\exp\{-mN^2\}\ . \tag{2.12}$$

(The estimator (2.11) can also be obtained by differentiating the rhs of (2.7) wrt $y$.)

*Remark* 2.1. In case the sample sizes are the same, (2.7) becomes $\hat{p}_{n,M}(y)=(n\pi)^{-1}\sum_{j=1}^{n}(Y_j-y)^{-1}\sin M(Y_j-y)$.

Inequalities (2.9) and $\Phi(-x)\leqq((2\pi)^{1/2}x)^{-1}\exp\{-x^2/2\}$ for $x>0$ yield

(2.13) $$\sup_y \mathrm{E}\,[(\hat{p}_{n,M}(y)-p(y))^2] \leqq (\pi^2 n)^{-1}M^2+(m\pi)^{-2}\exp\{-mM^2\}$$

for $M^2>1$. Similarly,

(2.14) $$\sup_y \mathrm{E}\,[(\hat{p}'_{n,N}(y)-p'(y))^2] \leqq (4\pi^2 n)^{-1}N^4+(m\pi)^{-2}\exp\{-mN^2\}$$

where $\hat{p}'_{n,N}(y)=(n\pi)^{-1}\sum_{j=1}^{n}(Y_j-y)^{-2}\{\sin N(Y_j-y)-N(Y_j-y)\cos N(Y_j-y)\}$. As a result of inequalities (2.13) and (2.14) and their intent, we obtain the following theorem.

THEOREM 2.1. *If $m(j)=m$ for all $j$ and $M^2(n)=N^2(n)=m^{-1}\log n$, then $\hat{p}_{n,M(n)}$ and $\hat{p}'_{n,N(n)}$ are such that*

(2.15) $$\sup_y \mathrm{E}\,[(\hat{p}_{n,M(n)}(y)-p(y))^2]=O(n^{-1}\log n)$$

*and*

(2.16) $$\sup_y \mathrm{E}\,[(\hat{p}'_{n,N(n)}(y)-p'(y))^2]=O(n^{-1}(\log n)^2)\ .$$

The rates mentioned above are better than those which have been obtained by using the kernel estimates of Parzen [5] (in the case of estimation of $p$) or of Johns and Van Ryzin [1] (in the case of estimation $p$ and $p'$). The reason that better rates are attained is that we have exploited the fact that the density $p$ to be estimated is a mixture of normal densities.

*Remark* 2.2. In the more general case of varying sample sizes, (2.9) and (2.12) can be weakened to obtain for $M^2>1$,

(2.17) $$\sup_y \mathrm{E}\,[(\hat{p}_{n,M}(y)-p(y))^2] \leqq (\pi^2 n)^{-1}M^2\exp\{m(m-1)M^2\}+(m\pi)^{-2}\exp\{-mM^2\}$$

and

(2.18) $$\sup_y \mathrm{E}\,[(\hat{p}'_{n,N}(y)-p'(y))^2] \leqq (4n\pi^2)^{-1}N^4\exp\{m(m-1)N^2\}+(m\pi)^{-2}\exp\{-mN^2\}\ .$$

The following theorem concerning the case where the sample sizes are bounded follows immediately from inequalities (2.17) and (2.18)

THEOREM 2.2. *If $m(j)\leqq\bar{m}<\infty$ for all $j$ and $M^2(n)=N^2(n)=(m(n+1))^{-2}\log n$, then $\hat{p}_{n,M(n)}$ and $\hat{p}'_{n,N(n)}$ are such that*

(2.19) $$\sup_y \mathrm{E}\,[(\hat{p}_{n,M(n)}(y)-p(y))^2]=O(n^{-1/\bar{m}}\log n)$$

*and*

(2.20) $$\sup_y \mathrm{E}\,[(\hat{p}_{n,N(n)}(y)-p'(y))^2]=O(n^{-1/\bar{m}}(\log n)^2)\ .$$

The estimates $\hat{p}_{n,M(n)}$ and $\hat{p}'_{n,N(n)}$ of Theorem 2.2 are better than those exhibited in O'Bryan and Susarla [4] in that they are more explicit, can readily be evaluated by computer methods, and yield rates for mean square error convergence which makes it possible to obtain rate of convergence results in the non-identical empirical Bayes estimation problem discussed in Section 1.

*Remark* 2.3. Slight improvements in the definition of the estimators $\hat{p}_{n,M(n)}$ and $\hat{p}'_{n,N(n)}$ can be achieved by retracting these estimators to the intervals $[0,(2\pi m(n+1))^{-1/2}]$ and to $[-(2\pi m(n+1))^{-1/2},(2\pi m(n+1))^{-1/2}]$ respectively.

## 3. Empirical Bayes estimation under the assumption $\Theta=[0,1]$

In this section we assume that $G$ is a distribution on $[0,1]$ and as in Section 2 we will denote $m(n+1)$ by $m$ and ${}_{n+1}p$ by $p$. Recalling from (1.4) that a nonrandomized rule which is Bayes with respect to $G$ in the $(n+1)$st problem is given by

(3.1) $$t_G^m(y)=m^{-1}y+(p'(y)/p(y))\ ,$$

it is immediate that

(3.2) $$|p'(y)/p(y)|\leqq 1+m^{-1}|y|\ .$$

In view of the discussion following (1.5), as a candidate for an empirical Bayes estimator $\{t_n\}$ satisfying (1.3), let $0<M(n)\to\infty$, $0<N(n)\to\infty$, and $0<\delta(n)\to 0$ be sequences to be specified later on and define

(3.3) $$t_n(y)=\mathrm{tr}\,(m^{-1}y+(\hat{p}'_n(y)/\max\{\hat{p}_n(y),\delta(n)\}))$$

where tr stands for retraction to $[0,1]$ and

(3.4) $$\hat{p}_n=\hat{p}_{n,M(n)} \quad\text{and}\quad \hat{p}'_n=\hat{p}'_{n,N(n)}$$

with $\hat{p}_{n,M}$ and $\hat{p}'_{n,N}$ defined respectively by (2.7) and (2.11).

In light of (1.5), we examine $(t_G^m-t_n)^2$ and obtain from (3.1), (3.2), (3.3), and the inequality $(ab^{-1}-c(\max\{d,\delta\})^{-1})^2\leqq 2\delta^{-2}\{(a-c)^2+a^2b^{-2}((b-d)^2+\delta^2[b<\delta])\}$ for all $a,b>0$, $c$, $d$, and $\delta>0$,

(3.5) $$(t_G^m(y)-t_n(y))^2\leqq 2(\delta(n))^{-2}\{(\hat{p}'_n(y)-p'(y))^2+(1+m^{-1}|y|)^2 \cdot(\hat{p}_n(y)-p(y))^2\}+2(1+m^{-1}|y|)^2[p(y)<\delta(n)]\ ,$$

where [ ] has been used to denote the indicator function of the set inside the brackets. The following lemma is needed to bound the expectation of the last term in the rhs of (3.5) and can be obtained in the same manner as Corollary 4.2 of Susarla [7].

LEMMA 3.1. *For* $0<a<1$ *and* $\delta>0$,

$$\mathrm{E}\,[(1+m^{-1}|Y_{n+1}|)^2[p(Y_{n+1})<\delta]]\leqq c(a,m)\delta^a \tag{3.6}$$

*for some finite constant* $c(a, m)$ *depending on* $a$ *and* $m$.

As a consequence of (1.5), (3.5), and Lemma 3.1, we have that for $0<a<1$,

$$\begin{aligned}0&\leqq R_n(t_n, G)-R^m(G)\\ &\leqq 2(\delta(n))^{-2}\{\mathrm{E}\,[(\hat{p}_n'(Y_{n+1})-p'(Y_{n+1}))^2]+\mathrm{E}\,[(1+m^{-1}|Y_{n+1}|)^2\\ &\quad\cdot(\hat{p}_n(Y_{n+1})-p(Y_{n+1}))^2]\}+c(a,m)(\delta(n))^a\,.\end{aligned} \tag{3.7}$$

Using the inequality $\mathrm{E}\,[(1+m^{-1}|Y_{n+1}|)^2]\leqq(1+\sqrt{2}\,)^2$, and (3.7), the following two theorems follow directly from Theorems 2.1 and 2.2.

THEOREM 3.1. *Let* $m(j)=m$ *for all* $j$ *and* $0<a<1$. *Then with* $M^2(n)=N^2(n)=m^{-1}\log n$ *and* $\delta(n)=n^{-1/(2+a)}$, *the sequence of rules* $\{t_n\}$ *with* $t_n$ *defined by* (3.3) *and* (3.4) *is a.o.* $(n^{-a/(2+a)}(\log n)^2)$.

THEOREM 3.2. *Let* $m(j)\leqq\overline{m}<\infty$ *for all* $j$ *and* $0<a<1$. *Then with* $M^2(n)=N^2(n)=(m(n+1))^{-2}\log n$ *and* $\delta(n)=n^{-1/((a+2)\overline{m})}$, *the sequence of rules* $\{t_n\}$ *with* $t_n$ *defined by* (3.3) *and* (3.4) *is a.o.* $(n^{-a/(\overline{m}(2+a))}(\log n)^2)$.

## 4. Empirical Bayes estimation under the assumption $\mathrm{E}\,[|\theta|^{2+\gamma}]<\infty$

Since the results in Section 2 did not depend on $G$ having a compact support, it should be possible to obtain a rate for risk convergence of the sequence $\{t_n\}$, with $t_n$ defined in (3.3) less the truncation to [0, 1], under the assumption that $m(j)\leqq\overline{m}$ for all $j$ and a weaker condition than that the support of $G$ is compact. To this end we prove a lemma which is analogous to Lemma 3.1 and is interesting in its own right. Let $Y$ abbreviate $Y_{n+1}$, and recall that $p$ is the density of $Y$, and $m$ abbreviates $m(n+1)$.

LEMMA 4.1. *If* $\mathrm{E}\,[|\theta|^{2+\gamma}]=B_\gamma<\infty$ *for some* $\gamma>0$, *then*

$$\mathrm{E}\,[(p'(Y)/p(Y))^2[p(Y)<\delta]]\leqq c(m, B_\gamma, \gamma)(-\log(\delta(2\pi m)^{1/2}))^{-\gamma/(2+\gamma)} \tag{4.1}$$

*for* $0<\delta<1$ *and some finite constant* $c(m, B_\gamma, \gamma)$ *depending on* $m$, $B_\gamma$, *and* $\gamma$.

PROOF. Since $G$ is a probability measure, parts (a) and (d) of problem 5, pages 70–71 of Rudin [6] yield the set inequality

$$\{p(y)<\delta\}\subseteq\left\{\int(y-m\theta)^2dG(\theta)>-2m\log(\delta(2\pi m)^{1/2})\right\}. \tag{4.2}$$

Since $\int (y-m\theta)^2 dG(\theta) \leqq (y+mB)^2$ where $B^2 = \mathrm{E}[\theta^2] < \infty$ since $B_\gamma < \infty$, a Markov inequality and (4.2) yield

$$(4.3) \qquad \begin{aligned} \mathrm{P}[p(Y)<\delta] &\leqq (-2m \log(\delta(2\pi m)^{1/2}))^{-1} \mathrm{E}[(Y-mB)^2] \\ &= (-\log(\delta(2\pi m)^{1/2}))^{-1} c_1(m, B) \end{aligned}$$

where $c_1(m, B)$ is a finite constant depending on $m$ and $B$. By Hölder's inequality,

$$(4.4) \qquad \text{lhs of } (4.1) \leqq (\mathrm{E}[|p'(Y)/p(Y)|^{2+\gamma}])^{2/(2+\gamma)} (\mathrm{P}[p(Y)<\delta])^{\gamma/(2+\gamma)} .$$

Since $p'(Y) = p(Y)(\mathrm{E}[\theta|Y] - m^{-1}Y)$, the finiteness of $B_\gamma$ and the $c_r$-inequality imply that the first term of the rhs of (4.4) is bounded by a finite constant $c_2(m, B_\gamma, \gamma)$ depending on $m$, $B_\gamma$ and $\gamma$. Consequently, (4.3) and (4.4) complete the proof of the lemma.

With $t_n$ defined by

$$(4.5) \qquad t_n(y) = \hat{p}_n'(y) / \max\{\hat{p}_n(y), \delta(n)\}$$

where $\hat{p}_n$ and $\hat{p}_n'$ defined by (3.4), one obtains using Lemma 4.1 an inequality similar to (3.7) under the assumption that, $\mathrm{E}[|\theta|^{2+\gamma}] < \infty$ for some $\gamma > 0$,

$$(4.6) \qquad \begin{aligned} 0 &\leqq R_n(t_n, G) - R^m(G) \\ &\leqq 2(\delta(n))^{-2} \{\mathrm{E}[(\hat{p}_n'(Y) - p'(Y))^2] + \mathrm{E}[(p'(Y)/p(Y))^2 (\hat{p}_n(Y) - p(Y))^2]\} \\ &\quad + c(m, B_\gamma, \gamma)(-\log(\delta(n)(2\pi m)^{1/2}))^{-\gamma/(2+\gamma)} . \end{aligned}$$

Since the finiteness of $B_\gamma$ implies that $\mathrm{E}[(p'(Y)/p(Y))^2] < \infty$ as in the proof of Lemma 4.1, Theorems 2.1 and 2.2 together with (4.6) yield the following theorem.

THEOREM 4.1. *Let* $\mathrm{E}[|\theta|^{2+\gamma}] < \infty$ *for some* $\gamma > 0$. *If* $m(j) = m$ *for all* $j$, *let* $M^2(n) = N^2(n) = m^{-1} \log n$ *and* $(2\pi m)^{1/2} \delta(n) = n^{-\beta}$ *for some* $0 < \beta < 1/2$. *If* $m(j) \leqq \bar{m} < \infty$ *for all* $j$, *let* $M^2(n) = N^2(n) = (m(n+1))^{-2} \log n$ *and* $(2\pi m (n+1))^{1/2} \delta(n) = n^{-\beta}$ *for some* $0 < \beta < (2\bar{m})^{-1}$. *In either case the sequence of rules* $\{t_n\}$ *with* $t_n$ *defined by* (4.5) *is a.o.* $((\log n)^{-\gamma/(2+\gamma)})$.

## 5. Concluding remarks

Under the assumption that the sample sizes are bounded, Theorem 3.2 is a distinct improvement over Theorem 3.1 (in the case of normal distributions with known variance) of O'Bryan and Susarla [4] in two directions; namely, (1) the simplicity of the estimator and (2) the rate of risk convergence of the estimator. Again under the assumption that the sample sizes are bounded, Theorem 4.1 is an improvement over

Section 3 of this paper and Theorem 3.1 of the above paper in that the weaker condition $E[|\theta|^{2+\gamma}]<\infty$ for some $\gamma>0$ was sufficient for the results as compared to the assumption that $G$ has compact support.

As a final remark, we point out that perhaps the techniques of this paper can be used analogously to obtain similar results for empirical Bayes estimation of the location parameter of some exponential families.

## Acknowledgement

The authors wish to thank Professors James Hannan and Dennis Gilliland for suggestions which led to the improvement of the form of the estimates introduced in Section 2.

UNIVERSITY OE WISCONSIN-MILWAUKEE

## REFERENCES

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

[2] Loéve, Michel (1963). *Probability Theory*, Van Nostrand, New Jersey.

[3] O'Bryan, Thomas E. (1976). Some empirical Bayes results in the case of component problems with varying sample sizes for discrete exponential families, *Ann. Statist.*, **4**, to appear.

[4] O'Bryan, Thomas E. and Susarla, V. (1973). Empirical Bayes estimation with non-identical components, Continuous case, Submitted for publication.

[5] Parzen, Emanuel (1962). On estimation of the probability density and mode, *Ann. Math. Statist.*, **33**, 1065-1076.

[6] Rudin, Walter (1966). *Real and Complex Analysis*, McGraw-Hill, New York.

[7] Susarla, V. (1974). Rate of convergence in the sequence-compound squared-distance loss estimation problem for a family of *m*-variate normal distributions, *Ann. Statist.*, **2**, 118-133.