The Annals of Statistics
1978, Vol. 6, No. 4, 846–853

SOLUTIONS TO EMPIRICAL BAYES SQUARED ERROR LOSS ESTIMATION PROBLEMS

BY RICHARD J. FOX[1]

Michigan State University

Asymptotically optimal empirical Bayes squared error loss estimation procedures are developed for three families of continuous distributions, uniform $(0, \theta)$, $\theta > 0$, uniform $[\theta, \theta + 1)$, θ arbitrary, and a location parameter family of gamma distributions. The approach taken is to estimate the Bayes estimator directly. However, for the $[\theta, \theta + 1)$ case, it is shown that the indirect approach of applying the Bayes estimator, versus an almost sure weakly convergent estimator of the prior, also yields an asymptotically optimal procedure.

1. Introduction. Consider a sequence of independent and identically structured Bayes statistical decision problems for which the common prior probability distribution over the state space is unknown. The empirical Bayes problem consists of constructing a procedure which makes use of historical data from previous problems in the sequence, as well as the current observation, and whose risk for the current problem converges to the generic Bayes optimal risk. A procedure having this convergence property is said to be asymptotically optimal (a.o.).

In this paper, a.o. procedures are developed for the empirical Bayes squared error loss estimation problem (see Robbins (1955)) for three parametric families for which the Bayes estimator is easily estimated from the historical data. Although these examples are not of major methodological interest, they have played a traditionally important role in developing statistical theory. None of the existing literature establishing general solutions to empirical Bayes problems gives sufficient conditions which are satisfied in these particular examples. One obvious possibility is Theorem 1 of Robbins (1964). However, his Assumption [C] regarding the generic problem, which requires that the sup of the loss function over the action space be an integrable function of the parameter or "state of nature," is not satisfied.

The following notational devices will be used. A distribution function F will also be used to denote the associated measure. Occasionally, the argument of a function will not be displayed and operator notation will be used extensively to represent integrals, e.g., $\int f(t)\, d\lambda(t)$ might be written as $\lambda(f(t))$ or $\lambda(f)$. All intervals of integration will be open on the left and closed on the right. A

Received May 1974; revised June 1977.

[1] Research supported by National Science Foundation, Grant No. GP-13484; now at Procter & Gamble.

AMS 1970 *subject classifications*. Primary 62C25, 62C10; Secondary 62A15, 62F10, 62F15.

Kew words and phrases. Empirical Bayes squared error loss estimation, posterior mean, asymptotically optimal empirical Bayes procedure, Bayes optimal risk.

"+" or "−" superscript on a function will denote the positive or negative part of the function respectively. $[A]$ represents the indicator function of the event A.

Let $\{F_\theta \mid \theta \in \Omega\}$, Ω being a Borel subset of the real line, be a class of distributions, possessing densities, f_θ, with respect to Lebesgue measure μ. Let G be a prior distribution on Ω. Define $K(x) = \int F_\theta(x)\, dG(\theta)$ and $k(x) = \int f_\theta(x)\, dG(\theta)$, i.e., K and k are respectively the marginal distribution function and density of x of the pair (θ, x). Let $(x_1, x_2, \cdots)$ be a sequence of i.i.d., according to K, random variables. Let P be the product measure on the space of sequences $(x_1, x_2, \cdots, (\theta, x))$, resulting from K^∞ and the joint distribution of (θ, x). Define

$$K_n(x) = n^{-1} \textstyle\sum_{i=1}^n [x_i \leqq x] \,. \tag{1.1}$$

Let $\phi(x)$ denote the Bayes estimator versus G assuming squared error loss (the posterior mean), i.e.,

$$\phi(x) = \frac{\int \theta f_\theta(x)\, dG(\theta)}{\int f_\theta(x)\, dG(\theta)} \tag{1.2}$$

(undefined ratios are taken to be zero unless otherwise specified).

Let R be the Bayes optimal risk versus G, i.e., $R = P((\phi(x) - \theta)^2)$. Our objective is to find an estimator of ϕ, say ϕ_n, based on $x_1, x_2, \cdots, x_n$, whose risk, defined by $P((\phi_n(x) - \theta)^2)$ and denoted by R_n, converges to R as n increases, i.e., ϕ_n is a.o. Assuming that R_n and R are finite, we have

$$R_n - R = P((\phi_n - \phi)^2) \,. \tag{1.3}$$

Hence, when (1.3) holds, asymptotic optimality is equivalent to $P((\phi_n - \phi)^2) \to 0$ as $n \to \infty$. In Sections 2 and 4 of this paper, we assume

$$G(\theta^2) < \infty \,. \tag{A}$$

By Jensen's inequality, (A) implies $P(\phi^2) < \infty$ and thus guarantees $R < \infty$. For the family considered in Section 3, it is easily shown that $R \leqq 1$ without any restrictions on the prior G.

2. Uniform $(0, \theta)$ **case.** Let $f_\theta(x) = \theta^{-1}[0 < x < \theta]$ where $\theta \in \Omega = (0, \infty)$. For this family,

$$K(x) = G(x) + k(x) \tag{2.1}$$

and (1.2) becomes

$$\phi(x) = \frac{1 - G(x)}{k(x)} \,. \tag{2.2}$$

Note that $k(x) = 0$ implies that both $G(x) = 1$ and $K(x) = 1$ so that $\phi(x) = 0$ by convention. Hence, it follows from (2.1) and (2.2) that

$$\phi(x) = x[k(x) > 0] + \psi(x) \tag{2.3}$$

where $\psi(x)$ is defined to be $(1 - K(x))/k(x)$. With h positive and depending on n, recall (1.1) and define k_n, a one-sided version of the typical divided difference estimator of k, by

$$k_n(x) = h^{-1}(K_n(x) - K_n(x - h)) \,.$$

This modification enables us to obtain certain bounds, e.g., (2.7), which are essential to the development of this section.

For each n, let $a_n(x)$ be a bounded nonnegative function defined on the positive reals. Estimate $\psi(x)$ by

$$\psi_n(x) = \left(\min\left(a_n(x), \frac{1-K_n(x)}{k_n(x)}\right)\right)[x \geqq h].$$

From (2.3), estimate $\phi(x)$ by

$$\phi_n(x) = x + \psi_n(x). \tag{2.4}$$

If $P(\phi^2) < \infty$, then $P(x^2) < \infty$. Therefore, under Assumption (A), (1.3) holds and we are concerned with choosing h and a_n so that $P((\phi_n - \phi)^2) = P((\psi_n - \psi)^2) \to 0$ as $n \to \infty$.

LEMMA 2.1. *Under* (A), *if* $nh^2 \to \infty$, $h \to 0$ *and if* $a_n(x) \to \infty$ *for each* x, *then* $P(\{(\psi_n - \psi)^-\}^2) \to 0$ *as* $n \to \infty$.

PROOF. Let $x \in A = \{x \mid k(x) > 0, G(y) \to G(x)$ as $y \to x\}$. Since k is continuous at x, $k(x) = K'(x)$. Since $nh^2 \to \infty$, by the Tchebichev inequality, $k_n(x) - h^{-1}(K(x) - K(x-h)) \to 0(K^\infty)$ where, as is typical, the parenthesized K^∞ denotes "in K^∞-measure." Hence, since $h^{-1}(K(x) - K(x-h)) \to k(x)$ as $h \to 0$,

$$k_n(x) \to k(x)(K^\infty). \tag{2.5}$$

By the Glivenko–Cantelli theorem, page 20 of Loève (1963), the Slutsky theorem, page 174 of Loève (1963) and (2.5), $(1 - K_n(x))/k_n(x) \to \psi(x)(K^\infty)$. Therefore, since $a_n(x) \to \infty$, $h \to 0$ and $P([A]) = 1$, $\psi_n \to \psi(P)$ so that

$$(\psi_n - \psi)^- \to 0(P). \tag{2.6}$$

Since $(\psi_n - \psi)^- \leqq \psi$ and under (A), $P(\psi^2) < \infty$, by (2.6) and the dominated convergence theorem, $P(\{(\psi_n - \psi)^-\}^2) \to 0$.

LEMMA 2.2. *For each* $x > 0$,

$$K^\infty(\{(\psi_n - \psi)^+\}^2) < \frac{a_n}{(nk)^{\frac{1}{2}}}\left\{\frac{c(a_n + 2h)}{(h\{1 - K(h)\})^{\frac{1}{2}}} + \frac{2(h^{-1}a_n + 1)}{k^{\frac{1}{2}}}\right\}$$

where c is the Berry–Esseen constant.

PROOF. If $k(x) = 0$ or if $x < h$, the result is obvious. Hence, fix x such that $x \geqq h$ and $k(x) > 0$ and let $(a_n(x) - \psi(x))^+ = b$. Since for $v \geqq b$, $K^\infty([_n\psi - \psi > v]) = 0$, it follows that $K^\infty(\{(\psi_n - \psi)^+\}^2) = \int_0^b K^\infty([\psi_n - \psi > v])\,dv^2$. Fix v such that $0 < v < b$ and for $i = 1, 2, \cdots, n$, define w_i, i.i.d., depending on x and v, by $w_i = h^{-1}[x - h < x_i \leqq x](\psi + v) - [x_i > x]$ and let $\bar{w} = n^{-1}\sum_{i=1}^n w_i$. Note that $K^\infty([\psi_n - \psi > v]) = K^\infty([\bar{w} < 0])$. Let $\Delta = h^{-1}(K(x) - K(x-h))$ and note that $K(w_1) = \Delta(\psi + v) - (1 - K)$. Since k is decreasing, for $x \geqq h$, $\Delta \geqq k(x)$. Therefore, recalling the definition of ψ, for $x \geqq h$,

$$K(w_1) \geqq vk \; . \tag{2.7}$$

By the Berry–Esseen theorem, page 288 of Loève (1963), with σ^2 and r denoting respectively the variance and range of w_1,

$$K^\infty[\bar{w} < 0] \leqq \Phi(z) + cn^{-\frac{1}{2}}r\sigma^{-1} \tag{2.8}$$

where $z = -n^{\frac{1}{2}}\sigma^{-1}K(w_1)$, c is the Berry–Esseen constant and Φ is the standard normal distribution function. Noting that $\sigma^2 \geqq h^{-1}\{(\psi + v)^2\Delta(1 - K(h))\}$ and that $hr = (\psi + v) + h$, we obtain

$$\textstyle\int_0^b r\sigma^{-1}\, dv^2 \leqq \{h\Delta(1 - K(h))\}^{-\frac{1}{2}}(b^2 + 2hb) \, . \tag{2.9}$$

By weakening the tail bound of Φ, page 166 of Feller (1957), by (2.7) and by the fact that $\sigma \leqq r$, it follows that $\Phi(z) \leqq (n^{\frac{1}{2}}vk)^{-1}r$. Hence, since $b \leqq a_n$ and for $0 < v < b$, $hr < a_n + h$,

$$\textstyle\int_0^b \Phi(z)\, dv^2 < \dfrac{2(h^{-1}a_n{}^2 + a_n)}{n^{\frac{1}{2}}k} \, . \tag{2.10}$$

Substituting a_n for b and k for Δ in the right hand side of (2.9) and combining the result with the bounds displayed in (2.8) and (2.10), we obtain the bound of the lemma for $\int_0^b K^\infty([\bar{w} < 0])\, dv^2$ and the proof is complete.

Let $||a_n||_r$ denote the L_r-norm of the function a_n (see, e.g., pages 188 and 346 of Hewitt and Stromberg (1965)) with respect to Lebesgue measure restricted to the positive reals.

LEMMA 2.3.

$$P(\{(\psi_n - \psi)^+\}^2) \leqq n^{-\frac{1}{2}}\{c(h\{1 - K(h)\})^{-\frac{1}{2}}||a_n||_2(||a_n||_\infty + 2h) + 2(h^{-1}||a_n||_2{}^2 + ||a_n||_1)\} \, .$$

PROOF. Note that $P(\{(\psi_n - \psi)^+\}^2) = \int_h^\infty K^\infty(\{(\psi_n - \psi)^+\}^2)k\, du$. Extending the range of integration from (h, ∞) to $(0, \infty)$, introducing the bound of Lemma 2.2 in the integrand, noting that $\mu(k) = 1$ and $a_n{}^2 \leqq ||a_n||_\infty a_n$ and applying the Schwarz inequality, we obtain the desired bound.

THEOREM 2.1. *Under* (A), *if* $a_n(x) \to \infty$ *for each* $x > 0$, *if* $h \to 0$ *and if* $||a_n||_1 = o(n^{\frac{1}{2}})$, $||a_n||_2{}^2 = o(n^{\frac{1}{2}}h)$ *and* $||a_n||_\infty = 0(n^{\frac{1}{2}})$, *then* ϕ_n, *defined by* (2.4), *is* a.o.

PROOF. Noting that the hypotheses of this theorem imply that $nh^2 \to \infty$, the theorem follows directly from Lemmas 2.1 and 2.3.

REMARK. Define $h = n^{-\gamma}$ where $0 < \gamma < \frac{1}{2}$. Also, for $x > 0$, let $a_n(x) = n^\delta f(x)$ where $0 < 2\delta < \frac{1}{2} - \gamma$ and f is a positive-valued bounded function of $x > 0$ having the properties that $\int_0^\infty f(x)\, dx$ and $\int_0^\infty f^2(x)\, dx$ are both finite. It is easily seen that this selection of h and $a_n(x)$ satisfies the conditions of Theorem 2.1. In fact, $||a_n||_\infty = o(n^{\frac{1}{2}})$, which is a slightly stronger condition than required.

3. Uniform $[\theta, \theta + 1)$ **case.** Let $f_\theta(x) = [\theta \leqq x < \theta + 1]$ where $\theta \in \Omega = (-\infty, +\infty)$. For this family,

$$K(x) = G(x - 1) + xk(x) - \textstyle\int_{x-1}^x \theta\, dG \tag{3.1}$$

$$k(x) = G(x) - G(x - 1) \, . \tag{3.2}$$

By the right continuity of G, k is right continuous and it follows that k is the right hand derivative of K for all x. Thus, for convenience, estimate $k(x)$ by right divided differences instead of the typical divided differences, i.e.,

$$k_n(x) = h^{-1}(K_n(x+h) - K_n(x)) \tag{3.3}$$

where h is chosen so that $0 < h < 1$. By (3.2), for all x,

$$G(x) = \sum_{j=0}^{\infty} k(x-j) \tag{3.4}$$

and

$$K(x) = \int_{-\infty}^{x} k(y)\,dy = \int_{x-1}^{x} G(y)\,dy\,. \tag{3.5}$$

Using (3.4) and recalling (3.3), estimate $G(x)$, for all x, by

$$G_n^*(x) = \sum_{j=0}^{\infty} k_n(x-j)\,. \tag{3.6}$$

LEMMA 3.1. *If $h \to 0$ and $nh^2 \to \infty$, then for each x, $G_n^*(x) \to G(x)(K^\infty)$.*

PROOF. Let x be fixed. By (3.5),

$$\sum_{j=0}^{\infty} (K(x+h-j) - K(x-j)) = \sum_{j=0}^{\infty} \left(\int_{x-j}^{x+h-j} G(y)\,dy - \int_{x-j-1}^{x+h-j-1} G(y)\,dy\right).$$

Since the series on the right hand side of the above equality is telescopic and $\int_{x-j}^{x+h-j} G(y)\,dy \to 0$ as $j \to \infty$, we have

$$\sum_{j=0}^{\infty} (K(x+h-j) - K(x-j)) = \int_{x}^{x+h} G(y)\,dy\,. \tag{3.7}$$

By (3.7), $G_n^*(x)$, defined by (3.6), is the average of i.i.d. random variables having expectation $h^{-1}\int_{x}^{x+h} G(y)\,dy$ which converges to $G(x)$ as $h \to 0$. Therefore, since $nh^2 \to \infty$, by the Tchebichev inequality, $G_n^*(x) \to G(x)(K^\infty)$. By (1.2), (3.1), and (3.2), with $\psi(x) = (G(x) - K(x))/k(x)$,

$$\phi(x) = (x-1)[k(x) > 0] + \psi(x)\,. \tag{3.8}$$

Since the conditional distribution of θ given x is concentrated on $(x-1, x]$, it follows that $0 \leqq \psi \leqq 1$. Define the function ψ_n by

$$\psi_n = \min\left(1, \max\left(0, \frac{G_n^* - K_n}{k_n}\right)\right) \tag{3.9}$$

i.e., ψ_n is $((G_n^* - K_n)/k_n)^+$ truncated at 1. Recalling (3.8) and (3.9), define

$$\phi_n(x) = (x-1) + \psi_n(x)\,. \tag{3.10}$$

THEOREM 3.1. *If $nh^2 \to \infty$ and $h \to 0$, then ϕ_n, defined by* (3.10), *is* a.o.

PROOF. Since $P((\phi - \theta)^2) \leqq 1$ and $P((\phi_n - \theta)^2) \leqq 1$, it follows that (1.3) holds and it suffices to show that $P((\psi - \psi_n)^2) \to 0$. Let x be fixed so that $k(x) > 0$. Applying Lemma 3.1 and the same logic as used in the proof of Lemma 2.1, it can be shown that $\psi_n(x) \to \psi(x)(K^\infty)$. Hence, since $P([k(x) > 0]) = 1$, $\psi_n \to \psi(P)$. By the bounded convergence theorem, $P((\psi_n - \psi)^2) \to 0$ and the proof is complete.

REMARK. Note that (A) ($G(\theta^2) < \infty$), which implies that $R < \infty$, is not made in this section. For this family, $R \leqq 1$ for any prior G. However, ϕ can have an infinite second moment as is the case when G is discrete and attaches mass $Cj^{-\frac{3}{2}}$ at j, where $j = 1, 2, \cdots$ and C is a normalizing constant.

REMARK. Let $\hat{G}_n$, based on $(x_1, x_2, \cdots, x_n)$, be a sequence of distribution functions converging weakly to G a.s. K^∞ (see Robbins (1964) and Fox (1970)). Redefine $\phi_n(x)$ to be the Bayes estimator versus $\hat{G}_n$ with $\frac{0}{0}$ defined to be x so that $x - 1 < \phi_n(x) \leqq x$. It is easily established via the bounded convergence theorem that this ϕ_n is also a.o.

4. A location parameter family of gamma distributions. For $\theta \in \Omega = (-\infty, +\infty)$, let $f_\theta(x) = (\Gamma(\alpha))^{-1}(x-\theta)^{\alpha-1}e^{-(x-\theta)}[x \geqq \theta]$ where $\alpha \geqq 1$ and Γ represents the gamma function. For this family,

$$k(x) = (\Gamma(\alpha))^{-1} \int_{-\infty}^{x} (x-\theta)^{\alpha-1}e^{-(x-\theta)}\, dG(\theta)\,. \tag{4.1}$$

By (1.2) and the definition of $f_\theta(x)$ for this family,

$$\phi(x) = \frac{\int_{-\infty}^{x} \theta(x-\theta)^{\alpha-1}e^{-(x-\theta)}\, dG(\theta)}{\Gamma(\alpha)k(x)}\,.$$

LEMMA 4.1. *If $k(x) > 0$, then $\phi(x) = x - \alpha\psi(x)$, where*

$$\psi(x) = \frac{\int_{-\infty}^{t} e^{-(x-t)}\, dK(t)}{k(x)}\,.$$

PROOF. Note that $\alpha \int_{-\infty}^{x} e^{-(x-t)}\, dK(t) = \alpha \int_{-\infty}^{x} e^{-(x-t)}k(t)\, dt$. Replacing k by the expression of (4.1), inverting the order of integration in the resulting expression and performing the inner integration yields $k(x)(x - \phi(x))$ and by the definition of $\psi(x)$ the proof is complete.

Estimate $\psi(x)$ by

$$\psi_n(x) = \min\left(a_n(x), \frac{\int_{-\infty}^{x} e^{-(x-t)}\, dK_n(t)}{k_n(x)}\right), \tag{4.2}$$

where K_n is defined by (1.1), k_n is defined by (3.3) and a_n is a bounded non-negative function of x for each n. Besides being convenient, the use of (3.3) also enables certain crucial bounds, e.g. (4.5), to be obtained. Define

$$\hat{\psi}(x) = \frac{\int_{-\infty}^{x} e^{-(x-t)}\, dK(t)}{h^{-1}(K(x+h) - K(x))}\,.$$

This function, $\hat{\psi}$, is a tool for establishing the convergence in quadratic mean of ψ_n to ψ. Note that (4.1) and the fact that $\alpha \geqq 1$ imply that for any x and $a > 0$,

$$k(x+a) \geqq e^{-a}k(x)\,. \tag{4.3}$$

Further, it follows from (4.3) that

$$h^{-1}(K(x+h) - K(x)) = h^{-1} \int_{x}^{x+h} k(y)\, dy \geqq e^{-h}k(x)\,. \tag{4.4}$$

LEMMA 4.2. *Under* (A) $(G(\theta^2) < \infty)$, *if $h \to 0$, then $P((\hat{\psi} - \psi)^2) \to 0$.*

PROOF. If x is such that $k(x)$ is positive and if $h \to 0$, then $\hat{\psi}(x) \to \psi(x)$. Recall that (A) implies $P(\phi^2) < \infty$ and since $\int x^2\, dP = \alpha(\alpha + 1) + 2\alpha G(\theta) + G(\theta^2)$, (A) also implies $P(x^2) < \infty$. Since for $k(x) > 0$, $\alpha\psi(x) = x - \phi(x)$, it follows that $P(\psi^2) < \infty$. If $k(x) > 0$, by (4.4), $|\hat{\psi}(x) - \psi(x)| \leqq (e^h + 1)\psi(x)$. Hence, by the dominated convergence theorem, $P((\hat{\psi} - \psi)^2) \to 0$.

LEMMA 4.3. *Under* (A), *if* $h \to 0$, $nh^2 \to \infty$ *and* $a_n(x) \to \infty$ *for all* x, *then* $P(\{(\Psi_n - \hat{\Psi})^-\}^2) \to 0$.

PROOF. Fix x such that $k(x) > 0$. By the strong low of large numbers, $\int_{-\infty}^x e^{-(x-t)}\, dK_n(t) \to \int_{-\infty}^x e^{-(x-t)}\, dK(t)$ a.s. K^∞. Continuing as in the proof of Lemma 2.1, we see that $\psi_n(x) \to \psi(x)(P)$. Further, since $\hat{\psi}(x) \to \psi(x)$, $\psi_n(x) - \hat{\psi} \to 0(P)$. By (4.4), $\{\psi_n(x) - \hat{\psi}(x)\}^- \leqq e^h\psi(x)$ and the result follows by applying the dominated convergence theorem as in the proof of Lemma 4.2.

LEMMA 4.4. $P(\{(\psi_n - \hat{\psi})^+\}^2) \leqq 2e^h n^{-\frac{1}{2}}(c + 1)\mu(a_n + h^{-1}a_n^2)$, *where c is the Berry–Esseen constant.*

PROOF. Fix x such that $k(x) > 0$. Analogous to the proof of Lemma 2.2, we deal with $\int_0^b K^\infty([\psi_n - \hat{\psi} > v])\, dv^2 = \int_0^b K^\infty([\bar{w} > 0])\, dv^2$ where $b = (a_n(x) - \hat{\psi}(x))^+$ and $w_i = [x_i \leqq x]e^{-(x-x_i)} - (\hat{\psi} + v)h^{-1}[x < x_i \leqq x + h]$ for $0 < v < b$ and $i = 1, 2, \cdots, n$. Analogous to (2.7) and (2.8), we obtain from (4.4)

$$K(w_1) = -vh^{-1}(K(x + h) - K(x)) \leqq -ve^{-h}k(x) \tag{4.5}$$

and consequently

$$K^\infty([\bar{w} > 0]) \leqq \Phi(z) + cn^{-\frac{1}{2}}\sigma^{-1}(1 + h^{-1}a_n) \tag{4.6}$$

where σ is the standard deviation of w_1, $1 + h^{-1}a_n$ is a bound on the range of w_1 and $\sigma z = -n^{\frac{1}{2}}ve^{-h}k(x)$. Note that $ve^{-h}k(x) \leqq \sigma \leqq 1 + h^{-1}a_n$ (the lower bound being obtained via (4.4)). Using these bounds for σ and continuing from (4.6) as in the proof of Lemma 2.2, leads to

$$K^\infty(\{(\psi_n - \hat{\psi})^+\}^2) \leqq \frac{2a_n e^h(c + 1)(1 + a_n h^{-1})}{n^{\frac{1}{2}}k(x)}\,.$$

The proof is completed by continuing as in the proof of Lemma 2.3.

Using Lemma 4.1, with ψ_n defined by (4.2), estimate $\phi(x)$, for all x, by

$$\phi_n(x) = x - \alpha\psi_n(x)\,. \tag{4.7}$$

THEOREM 4.1. *Under* (A), *if* $||a_n||_1 = o(n^{\frac{1}{2}})$, $||a_n||_2^2 = o(n^{\frac{1}{2}}h)$, $h \to 0$ *and if* $a_n(x) \to \infty$ *for each* x, *then* ϕ_n, *defined by* (4.7), *is* a.o.

PROOF. Under (A), $R < \infty$ and $P(x^2) < \infty$ as seen in the proof of Lemma 4.2. Therefore, by (A) and the fact that a_n is bounded for each n, the conditions implying (1.3) hold and thus it suffices to show that $P((\phi_n - \phi)^2) \to 0$. Since $a_n(x) \to \infty$ for each x, $\mu(a_n^2) = o((nh^2)^{\frac{1}{2}})$ implies that $nh^2 \to \infty$. Hence, by Lemmas 4.3 and 4.4, $P((\psi_n - \hat{\psi})^2) \to 0$. By Lemma 4.2 and the triangle inequality for

L_2-norm, $P((\psi_n - \psi)^2) \to 0$ or equivalently $P((\phi_n - \phi)^2) \to 0$ and the proof is complete.

REMARK. The above development could easily be applied to the more general case of a two-parameter gamma distribution, i.e.,

$$f_\theta(x) = (\Gamma(\alpha))^{-1}\lambda^\alpha(x-\theta)^{\alpha-1}e^{-\lambda(x-\theta)}[x \geqq \theta]$$

where $\alpha \geqq 1$ and $\lambda > 0$. Of course, the expressions would have to be modified slightly to reflect the additional parameter.

REMARK. The conditions on h and $a_n(x)$ of the above theorem are a subset of those of Theorem 2.1. Hence, the sequences of values of h and functions $a_n(x)$ (with the domain of f extended to the entire real line) constructed in the remark following Theorem 2.1 satisfy the conditions of Theorem 4.1.

Acknowledgment. This paper forms part of a (Michigan State University) Ph. D. thesis and I wish to thank Professor James Hannan for suggesting these examples to me and for his guidance throughout the development of this work. Also, my thanks to the National Science Foundation whose grant financed this research and to the referees and editors of this journal for their extremely helpful comments and suggestions.

REFERENCES

FELLER, W. (1957). *An Introduction to Probability Theory and its Applications, Volume I* (2nd ed.). Wiley, New York.

FOX, R. (1970). Estimating the empiric distribution function of a parameter sequence. *Ann. Math. Statist.* **41** 1845–1852.

HEWITT, E. and STROMBERG, K. (1965). *Real and Abstract Analysis*. Springer-Verlag, New York.

LOÈVE, MICHEL (1963). *Probability Theory* (3rd ed.). Van Nostrand, Princeton.

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

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

B1S38 SHARON WOODS TECHNICAL CENTER
PROCTER & GAMBLE COMPANY
11520 REED HARTMAN HIGHWAY
CINCINNATI, OHIO 45241