The Annals of Statistics
1978, Vol. 6, No. 4, 740–754

EMPIRICAL BAYES ESTIMATION OF A DISTRIBUTION (SURVIVAL) FUNCTION FROM RIGHT CENSORED OBSERVATIONS[1]

BY V. SUSARLA[3] AND J. VAN RYZIN[2]

University of Wisconsin—Milwaukee and
University of California—Davis

This paper provides an empirical Bayes approach to the problem of nonparametric estimation of a distribution (or survival) function when the observations are censored on the right. The results use the notion of a Dirichlet process prior introduced by Ferguson.

The paper presents a generalization to the case of right censored observations of the rate result of an empirical Bayes nonparametric estimator of a distribution function of Korwar and Hollander in the uncensored case. The rate of asymptotic convergence of optimality is shown to be the best obtainable for the problem considered.

1. Introduction and summary. With $R(G)$ denoting the infimum Bayes risk in deciding about a parameter $\theta \sim G$, Robbins (1955) proposed decision rules, based on data gathered in n independent repetitions of the same decision problem, whose overall risk converges to $R(G)$ as $n \to \infty$. Since the appearance of this paper of Robbins (1955), much empirical Bayes work has evolved with most of the results obtained under the assumption that θ is a random vector. Some of the references where this assumption was relaxed are Johns (1957), Van Ryzin (1970), Korwar and Hollander (1976) and Susarla and Phadia (1976). The last two papers considered empirical Bayes squared error loss estimation and generalized empirical Bayes testing problems concerning a random distribution function F which is distributed according to the Dirichlet process prior D_α introduced by Ferguson (1973) (see also Antoniak (1974)).

This paper treats the generalized empirical Bayes estimation problem concerning F (described in detail in Section 2) when the observations in the sequence are independent but not identically distributed in that the distribution of the right censoring random variable associated with each decision problem in the sequence is not fixed.

Received February 1976; revised September 1977.

[1] Research sponsored in part by the United States Army under Contract No. DAAG29-75-C-0024 at Mathematics Research Center, Madison. Revision supported in part by DHEW, PHS, National Institute of Health under Grant No. 5 R01 GM23129-02.

[2] Research sponsored in part by National Institutes of Health Grant No. NIH-1-R01 CA18332-01.

[3] Revision supported in part by the National Science Foundation under Grant No. MCS76-05952.

AMS 1970 *subject classifications.* Primary 62C99; Secondary 62G05.

Key words and phrases. Empirical Bayes estimation, nonparametric estimation of a distribution (survival) function, Dirichlet process priors, right censored observations.

In Section 2, we describe the decision problem and its empirical Bayes analogue and an example where this empirical Bayes problem might be useful. Section 3 provides a necessary and a sufficient condition for the family of distributions on the real line to be identifiable with respect to Dirichlet process priors. (Identifiability is a necessary requirement if one hopes to solve the empirical Bayes problem.) Section 4 provides the best possible rate $O(n^{-1})$ uniformly in the parameter α for a risk convergent empirical Bayes estimator when the sequence of censoring distributions is known. Section 5 provides empirical Bayes estimators with rates $O(n^{-(1-\varepsilon)})$ $(0 < \varepsilon < 1)$ when the censoring distributions are unknown. The final section concludes with a few remarks concerning possible generalizations of the results of Sections 4 and 5.

For notational convenience, we let $[A]$ denote the indicator function of A and the arguments of functions will not be exhibited whenever they are clear from the context. Integrals are over $R = (-\infty, \infty)$ unless otherwise stated. Ratios of the form $\frac{0}{0}$ are taken to be zero (see (3), (15) and (28)). Familiarity with Ferguson's (1973) paper on Dirichlet process prior will be very helpful in reading this paper.

2. The problem and its empirical Bayes analogue. Let (F, X, Y) be a stochastic process where $1 - F$ is a random distribution function on $R = (-\infty, \infty)$ and distributed according to the Dirichlet process D_α (see Ferguson (1973)) with α denoting a finite measure on the Borel σ-field $\mathscr{B}$ of R; given F, $X \sim$ right sided distribution F (that is, $F(t) = P[X > t \mid F]$) and finally, Y is a univariate random variable independent of (F, X) and distributed according to the right sided distribution H. (Y may be defective, in which case $H \equiv 0$.) The decision problem is to estimate F using (δ, Z) where

$$(1) \qquad \delta = [X \leqq Y] \qquad \text{and} \qquad Z = \min\{X, Y\}$$

when the loss function is given by

$$(2) \qquad L(F, \hat{F}) = \int (F(u) - \hat{F}(u))^2\, dw(u)$$

where w is a known finite nonnegative measure on $\mathscr{B}$ and $\hat{F}$ is an estimate of F. Kaplan and Meier (1958) considered estimation of F in a nondecision theoretic setup and obtained two estimates of F, one of which is the maximum likelihood estimate.

A trivial extension to R for the case of unit sample size of (4.2) of Susarla and Van Ryzin (1975) shows that the Bayes estimate d_α of F in the above component decision problem is given by

$$(3) \qquad \begin{aligned} \{\alpha(R) + 1\}\, d_\alpha(u, (\delta, z)) &= 1 + \alpha(u) && \text{if} \quad u < z \\ &= \alpha(u) && \text{if} \quad \delta = 1 \quad \text{and} \quad u \geqq z \\ &= (1 + \alpha(z))\, \frac{(\alpha u)}{\alpha(z)} && \text{if} \quad \delta = 0 \quad \text{and} \quad u \geqq z\,, \end{aligned}$$

where $\alpha(\cdot)$ stands for $\alpha(\cdot, \infty)$. It is clear that one can use the above Bayes rule from (3) when α is known. By using (3) to estimate F one achieves the minimum possible risk $r(H, \alpha)$ for the family of loss functions in (2).

Suppose that α is unknown in the above component problem, but there is a sequence of independent stochastic processes $\{(F_n, \delta_n, Z_n)\}$ where for each n, (F_n, δ_n, Z_n) has the same probability structure as that for (F, δ, Z) of the component problem except that H is replaced by H_n. While the (δ_n, Z_n) are *independent* they are *not* necessarily *identically* distributed random vectors. In this setup, a generalization of the empirical Bayes approach of Robbins (1955) is applicable wherein one obtains an estimate S_n (called the nth component of the empirical Bayes procedure) of $d_\alpha(\cdot, (\delta_{n+1}, z_{n+1}))$ using the independent random vectors $(\delta_1, Z_1), \cdots, (\delta_n, Z_n)$ and shows that the overall risk less risk if α known approaches zero.

DEFINITION 1. The sequence of estimators $\{S_n\}$ is said to be asymptotically optimal with order $k(n)$ (a.o. $k(n)$) if

$$(4) \qquad 0 \leqq D_n = r_n(S_n, \{H_j\}_1^{n+1}, \alpha) - r(H_{n+1}, \alpha) \leqq ck(n)$$

where the constant c and $k(n)$ are independent of α and $\{H_n\}$ and $k(n) \to 0$ as $n \to \infty$.

We show that $k(n) = n^{-1}$ in Section 4 when $\{H_n\}$ is known and $k(n) = n^{-(1-\varepsilon)}(0 < \varepsilon < 1)$ when $\{H_n\}$ is unknown and under certain conditions on S_n, α, and $\{H_n\}$. The following lemma, whose proof is similar to the proof usually given in empirical Bayes parametric cases (see page 46, Maritz (1970)), is useful in obtaining the above rates of convergence.

LEMMA 1.

$$(5) \qquad 0 \leqq D_n = r_n(S_n, \{H_j\}_1^{n+1}, \alpha) - r(H_{n+1}, \alpha) = \int \{\int E_n[\{S_n(u, (\delta_{n+1}, z_{n+1})) - d_\alpha(u, (\delta_{n+1}, z_{n+1}))\}^2] \, dw(u)\} \, dG(\delta_{n+1}, z_{n+1})$$

where d_α is given by (3), *E_n stands for expectation operation with respect to the joint distribution of $(\delta_1, Z_1), \cdots, (\delta_n, Z_n)$ and G is the distribution function of (δ_{n+1}, Z_{n+1}).*

We will obtain rates of convergence results for D_n in Sections 4 and 5 by first obtaining rates of convergence results for the inside integral involved in the right-hand side of (5).

We briefly describe a possible application here in the context of survival times under a particular treatment T. Consider a cancer treatment study consisting of observing the length of survival time for the cancer patients under T. Due to dropping off from the study for some reason or other, some of these true survival times are censored on the right. Since the patients might respond differently for the same type of cancer under the same type of treatment T because of their health conditions, it is reasonable to assume that these patients have different survival distributions, but on the average they have the same survival distribution, say β. Thus, in statistical terminology, one has independent observations like $(F_1, X_1, Y_1), \cdots, (F_n, X_n, Y_n)$ where we do not observe the independent stochastic processes $F_1, \cdots, F_n$ having the same distribution, $X_1, \cdots, X_n$ are the unobservable random survival times, and $Y_1, \cdots, Y_n$ are the random censoring times, all corresponding to the first n patients. Notice that given $F_i = F$, $X_i \sim F$ for $i = 1, \cdots, n$. We observe only $\delta_i = [X_i \leqq Y_i]$ and $Z_i = \min\{X_i, Y_i\}$ for $i = 1, \cdots, n$ because of censoring. Looking at the data after the $(n+1)$th observation, one has not only (δ_{n+1}, Z_{n+1}) to infer about F_{n+1} (here we need a Bayes estimate of F_{n+1} with α unknown) but also $(\delta_1, Z_1), \cdots, (\delta_n, Z_n)$. These first n vector observations do contain information about α, and consequently should be useful in saying something about the survival distribution function F_{n+1} of the $(n+1)$th patient. Since Dirichlet process priors are reasonable to work with, we took in this paper $1 - F_1, \cdots, 1 - F_n$ to be i.i.d. Dirichlet processes with common parameter α with $\alpha(R)$ known. For further possible applications, see the books by Barlow and Proschan (1974) and Gross and Clark (1975).

We describe our procedures, their risks at the $(n+1)$th stage, and the Bayes risks at the $(n+1)$th stage with examples in which $H_n(u) = P[Y_n > u] = e^{-\beta u}$ and $\alpha(u)\ (= \alpha(u, \infty)) = ce^{-\theta u}$ for some constant $c > 0$ where $u > 0$. Of course, other distributions are possible. In the partially known censoring distribution case with the above choice for H_n with β unknown, we illustrate our procedure with cancer data. The data and the form of the empirical Bayes estimator are given in the Appendix. Generally, in life testing problems, the above choices for H_n and α are reasonable.

3. An identifiability problem. In this section, adhering to the notation of the decision problem of Section 2, we obtain a necessary condition and a sufficient condition for the identifiability of the class $\mathscr{F}$ of distribution functions on the real line R to be identifiable with respect to Dirichlet process priors. Such an identifiability result is needed to solve the empirical Bayes problem since in order to solve this problem, one must be able to estimate α in view of (3) from observations $(\delta_1, Z_1), \cdots, (\delta_n, Z_n)$ and the integrand in rhs equality (5). In short, identifiability here means that if two marginal distributions are equal, say for (δ, Z), in the decision problem described, then their prior distribution (here Dirichlet process priors) must be equal. So let the marginal distribution of (δ, Z) be G. Analogous to that of Teicher (1960), we have

DEFINITION 2. $\mathscr{F}$ is said to be identifiable with respect to Dirichlet process priors iff

$$G_1(i, x) = G_2(i, x) \quad \text{for} \quad (i, x) \quad \text{in} \quad \{0, 1\} \times R \Leftrightarrow \alpha_1 = c\alpha_2 \tag{6}$$

for some constant c where $(\delta, Z) \sim G_i$ when F involved in the definition of Z is $\sim \alpha_i$, $i = 1, 2$.

THEOREM 1 (Sufficiency). *If $H(x) = 0$ for all x (that is, X is uncensored) or the support of H is (a, ∞) for some $a < \infty$, then $\mathscr{F}$ is identifiable with respect to*

Dirichlet process priors. (*That is*, $G_1 = G_2 \Rightarrow \alpha_1(A)/\alpha_1(R) = \alpha_2(A)/\alpha_2(R)$ *for all measurable* A.)

PROOF. The first case follows from Proposition 4 of Section 3 of Ferguson (1973). So let the support of H be (a, ∞) and $G_1 = G_2$. We show that $\alpha_1 = c\alpha_2$ for some constant c. By the hypothesis and the above quoted result of Ferguson,

$$G_1(0, x) = \int_{-\infty}^{x} \frac{\alpha_1(y, \infty)}{\alpha_1(R)}\, dH(y) = \int_{-\infty}^{x} \frac{\alpha_2(y, \infty)}{\alpha_2(R)}\, dH(y) = G_2(0, x) \tag{7}$$

for x in R. If $\mu_i((-\infty, \cdot]) = G_i(0, \cdot)$ for $i = 1, 2$, then the corollary on page 28 of Chung (1974) shows that measures μ_1 and μ_2 are equal, or equivalently,

$$\int_B \left\{\frac{\alpha_1(y, \infty)}{\alpha_1(R)} - \frac{\alpha_2(y, \infty)}{\alpha_2(R)}\right\} dH(y) = 0 \qquad \text{for } B \text{ in } \mathscr{B}. \tag{8}$$

If there exists a y_0 such that

$$\alpha_1(y_0, \infty)/\alpha_1(R) \neq \alpha_2(y_0, \infty)/\alpha_2(R)\,, \tag{9}$$

then $y_0 \geqq a$ since the support of H is (a, ∞). Therefore, $G_1(0, x) = G_2(0, x)$ for all x implies that

$$\alpha_1(x, \infty)/\alpha_1(\alpha) = \alpha_2(x, \infty)/\alpha_2(R) \qquad \text{for} \quad x < a\,. \tag{10}$$

Moreover, since $\alpha_i(\cdot, \infty)/\alpha_i(R)$ is right continuous for $i = 1, 2$,

$$|\{\alpha_1(y, \infty)/\alpha_1(R)\} - \{\alpha_2(y, \infty)/\alpha_2(R)\}| > \delta > 0 \qquad \text{for} \quad y_0 \leqq y \leqq y_0 + \varepsilon$$

for some ε and $\delta > 0$. This inequality shows that $H(y_0 + \varepsilon) - H(y_0) = 0$ in view of (8) and hence $y_0 + 2^{-1}\varepsilon$ ($> a$) is not in the support H. This contradiction together with (10) gives that $\alpha_1(y, \infty)/\alpha_1(R) = \alpha_2(y, \infty)/\alpha_2(R)$ for all y. Now the above quoted result of Chung gives the theorem. □

THEOREM 2 (Necessity). *If* $H(t_0) = 1$ *for some* $t_0 < \infty$, *then* $\mathscr{F}$ *is not identifiable with respect to Dirichlet process priors.* (*If* Y *is bounded above by* M, *then we cannot distinguish between the measures* $\alpha_1(\cdot)/\alpha_1(R)$ *and* $\alpha_2(\cdot)/\alpha_2(R)$ *which agree up to* M *but* not *afterwards.*)

PROOF. Let α_1 be a fixed measure so that $\alpha_1(t_0, \infty) > 0$ and let $\alpha_2(B) = \alpha_1(B \cap (-\infty, t_0]) + \alpha_1(t_0, \infty)[t_1 \in B]$ for B in $\mathscr{B}$ where $t_1 > t_0$ and $[t_1 \in B] = 1$ if $t_1 \in B$ and $= 0$ if $t_1 \notin B$. Now it is a matter of simple verification that the measures α_1 and α_2 give the same marginal distribution for (δ, Z) and clearly $\alpha_1(R) = \alpha_2(R)$ and $\alpha_1 \neq \alpha_2$ if $\alpha_1(\{t_1\}) \neq \alpha_1(t_0, \infty)$. □

4. Empirical Bayes estimation with $\{H_n\}$ known. With a view towards defining an empirical Bayes estimator and to apply Lemma 1 of Section 2 to bound its risk difference from $r(H_{n+1}, \alpha)$, we note that $Z_1, \cdots, Z_n$ are independent and that

$$\alpha(R)P[Z_j > x] = \alpha(x)H_j(x) \qquad \text{for} \quad j = 1, \cdots, n \tag{11}$$

where, here and elsewhere, $\alpha(\cdot)$ abbreviates $\alpha(\cdot, \infty)$. Consequently,

$$n\{\hat{\alpha}_n(x)\} = \alpha(R) \textstyle\sum_{j=1}^{n} (H_j(x))^{-1}[Z_j > x] \tag{12}$$

is such that for each x in R,

$$E[\hat{\alpha}_n(x)] = \alpha(x) = \alpha(x, \infty) \tag{13}$$

$$n^2 \operatorname{Var}(\hat{\alpha}_n(x)) = \alpha(x) \textstyle\sum_{j=1}^{n} (H_j(x))^{-1}\{\alpha(R) - H_j(x)\alpha(x)\} = \gamma_n(x)\,. \tag{14}$$

We use (12), (13) and (14) to find the rate of convergence of D_n in (5) for the following empirical Bayes estimate S_n. S_n is motivated by the fact that in view of the rhs of the equality (5), we need a good estimate of $d_\alpha(u, (\delta_{n+1}, Z_{n+1}))$, and a good estimate of d_α (see (3)) is simply obtained by replacing α of d_α by $\hat{\alpha}_n$ of (12) with $\hat{\alpha}_n$ retracted to $[0, \alpha(R)]$ since $\alpha(\cdot) \leqq \alpha(R)$. Thus $\hat{S}_n$ is defined by

$$\begin{aligned}
(\alpha(R) + 1)\hat{S}_n(u, (\delta_{n+1}, Z_{n+1}))
&= 1 + \min\{\hat{\alpha}_n(u), \alpha(R)\} && \text{if } u < z_{n+1} \\
&= \min\{\hat{\alpha}_n(u), \alpha(R)\} && \text{if } \delta_{n+1} = 1 \text{ and } u \geqq z_{n+1} \\
&= \{1 + \min\{\hat{\alpha}_n(z_{n+1}), \alpha(R)\}\} \min\left\{\frac{\hat{\alpha}_n(u)}{\hat{\alpha}_n(z_{n+1})}, 1\right\} && \\
& && \text{if } \delta_{n+1} = 0 \text{ and } u \geqq z_{n+1}\,,
\end{aligned} \tag{15}$$

where, according to our convention, ratios of the form $\frac{0}{0}$ are to be taken as zero. To obtain the main results of this and the next section, the following lemma of Singh (1974) with $\gamma = 2$ will be useful. Its proof can be found in O'Bryan and Susarla (1975).

LEMMA 2. *Let a, b and l be in R with $b, l > 0$. If A and B are real valued random variables, then for all $\gamma > 0$,*

$$b^\gamma E[(\min\{|(A/B) - (a/b)|, l\})^\gamma] \leqq c\{E[|A - a|^\gamma] + (|a/b|^\gamma + l^\gamma)E[|B - b|^\gamma]\}$$

for some constant c depending on γ alone.

THEOREM 3. *For each n, let H_n be known with support $= (a_n, \infty]$.*

$$0 \leqq n^2 D_n - \textstyle\int \gamma_n(u)\, dw(u) \leqq c_1 E[[\delta_{n+1} = 0] \int_{Z_{n+1}}^{\infty} \{\gamma_n(u) + \gamma_n(Z_{n+1})\}\, dw(u)]$$

for some absolute constant c_1 where γ_n is defined by (14).

PROOF. In view of (5) and the definitions of d_α and $\hat{S}_n$ in (3) and (15) respectively, we consider the inner expectation of the rhs of (5) in two cases, namely, (1) $u < z_{n+1}$ or $\delta_{n+1} = 1$ and $u \geqq z_{n+1}$ and (2) $\delta_{n+1} = 0$ and $u \geqq z_{n+1}$.

In case (1), with $\gamma_n(\cdot)$ as in (14),

$$4n^2 E_n[\{\hat{S}_n(u, (\delta_{n+1}, z_{n+1})) - d_\alpha(u, (\delta_{n+1}, z_{n+1}))\}^2] \leqq \gamma_n(u)\,. \tag{16}$$

In case (2), the inequality $(a + b)^2 \leqq 2(a^2 + b^2)$ along with the intermediate term $(1 + \min\{\hat{\alpha}_n(z_{n+1}), 1\})\alpha(u)/\alpha(z_{n+1})$, the definitions of d_α and $\hat{S}_n$ in (3) and

(15) respectively imply that for $\delta_{n+1} = 0$ and $u \geqq z_{n+1}$,

$$\begin{aligned} &4E_n[\{\hat{S}_n(u, (\delta_{n+1}, z_{n+1})) - d_\alpha(u, (\delta_{n+1}, z_{n+1}))\}^2] \\ (17) \qquad &\leqq 2E_n[\{\hat{\alpha}_n(z_{n+1}) - \alpha(z_{n+1})\}^2] \\ &\quad + 2(\alpha(R)+1)^2 E_n\left[\left\{\min\left\{1, \frac{\hat{\alpha}_n(u)}{\hat{\alpha}_n(z_{n+1})}\right\} - \frac{\alpha(u)}{\alpha(z_{n+1})}\right\}^2\right] \end{aligned}$$

where we used the facts that $\min\{1, (\hat{\alpha}_n(u))/\hat{\alpha}_n(z_{n+1})\}$ and $\alpha(z_{n+1})$ are $\leqq \alpha(R) + 1$. The first term on the rhs of (17) is $2\gamma_n(z_{n+1})$. The second expectation is at most

$$(18) \qquad E_n\left[\left\{\min\left\{\left|\frac{\hat{\alpha}_n(u)}{\hat{\alpha}_n(z_{n+1})} - \frac{\alpha(u)}{\alpha(z_{n+1})}\right|, 1\right\}\right\}^2\right]$$

since $\alpha(u)\alpha(z_{n+1}) \leqq 1$.

Now we apply Lemma 2 to (18) with $A = 1 - \hat{\alpha}_n(u)$, $B = 1 - \hat{\alpha}_n(z_{n+1})$, $a = 1 - \alpha(u)$, $b = 1 - \alpha(z_{n+1}) > 0$, ($\alpha(z_{n+1}) = 1$ will be considered below), $l = 1$ and $\gamma = 2$. This application results in (18) $\leqq c_1\{\gamma_n(u) + \gamma_n(z_{n+1})\}$ for some absolute constant c_1 for $\alpha(z_{n+1}, \infty) > 0$.

If $\alpha(z_{n+1}) = 1$, then $\alpha(u) = 0$ since $u \geqq z_{n+1}$ and hence $\alpha(u)/\alpha(z_{n+1}) = 0$ by our convention. Moreover, $\alpha(R)P[X_j > z_{n+1}] = \alpha(z_{n+1}) = 0$ and hence $P[Z_j > z_{n+1}] = 0$ for $j = 1, \cdots, n$. Therefore, with probability one, $\{\hat{\alpha}_n(u)/\hat{\alpha}_n(z_{n+1})\} = 0$ for $j = 1, \cdots, n$ by our convention since $u \geqq z_{n+1}$. Hence (18) is equal to zero if $\alpha(z_{n+1}) = 1$.

Combining the results of the two last paragraphs, (16), (17) and (18), gives the desired bound. □

Three corollaries to Theorem 3 are given below. The first corollary shows that the rate of convergence is n^{-1} for S_n of (15), the second one shows that rate n^{-1} cannot be improved upon while the third corollary obtains the rate n^{-1} obtained by Korwar and Hollander (1976) in the uncensored case.

COROLLARY 1. *Let the conditions of Theorem* 3 *hold. Then* S_n *is* a.o. n^{-1} *uniformly in* α *provided* $n^{-1}\sum_{j=1}^n \int (H_j(u))^{-1}\,dw(u)$ *is bounded in* n.

PROOF. The result follows from the bound in the theorem upon noticing that the assumed condition implies that $(1/n)\int \gamma_n(u)\,dw(u)$ and $(1/n)\int_{Z_{n+1}}^\infty \gamma_n(Z_{n+1})\,dw(u) \leqq \int_{Z_{n+1}}^\infty n^{-1}\sum_{j=1}^n (H_j(u))^{-1}\,dw(u)$ are bounded in n. □

COROLLARY 2. *Under conditions of Theorem* 3, *the rate in Corollary* 1 *cannot be improved uniformly in all* α *and all sequences* $\{H_n\}$ *if the support of* $w = R$.

PROOF. The result will be proved by showing that $D_n = r_n(\hat{S}_n, \{H_j\}_1^{n+1}, \alpha) - r(H_{n+1}, \alpha) \geqq n^{-1}c_2$ for some positive constant c_2 for certain choices of α and $\{H_n\}$.

Let $\{H_n\}$ be any sequence satisfying the conditions of Theorem 3 and additionally, be such that $H_n(a) = 0$ for a point a in R and $\alpha(R) = 1$. Then $n\{\hat{\alpha}_n(x)\} = \sum_{j=1}^n [Z_j > x]$ for $x \leqq a$ by (12). Therefore using the definition of $\hat{S}_n$ for $u \geqq z_{n+1}$ and $\delta_{n+1} = 1$ as in case (1) of the proof of Theorem 3 and the equality obtained in Lemma 1 together with (14) show that

$$4nD_n \geqq \int_{[\delta_{n+1}=1, z_{n+1}\leqq a]} \{\int_{z_{n+1}}^{a} \alpha(u)(1-\alpha(u))\,dw(u)\}\,dG(\delta_{n+1}, z_{n+1}) \tag{19}$$

where G is the distribution function of (δ_{n+1}, Z_{n+1}).

Now let the measure α, z_0, and ε be such that $0 < 1-\alpha(z_0) < 1-\alpha(z_0+\varepsilon) < 1$ for $z_0 < a$ and ε in $(0, a-z_0)$. Now the rhs of (19) is positive since $\int_{z_{n+1}}^{a} a(u)(1-\alpha(u))\,dw(u) > 0$ for z_{n+1} in $[z_0, z_0+\varepsilon)$ and since $P[\delta_{n+1}=1, z_0 < z_{n+1} \leqq z_0+\varepsilon] = \alpha[z_0, z_0+\varepsilon) > 0$ by the choice of z_0 and ε. □

COROLLARY 3. *If X_n is uncensored for all n, then $\hat{S}_n$ is* a.o. (n^{-1}) *uniformly in* α.

PROOF. The result is immediate from the bound in the theorem since $P[\delta_{n+1} = 0] = 0$ and the integral $n^{-1}\int \gamma_n(u)\,dw(u) = \int \alpha(u)(\alpha(R)-\alpha(u))\,dw(u) < \infty$. □

EXAMPLE 1. Here we show how the theorem and its corollaries apply in the situation where $H_n(u) = e^{-\beta u}$ on $(0, \infty)$ and $\alpha(u) = ce^{-\theta u}$ for $u > 0, \theta > 0$ and $c > 0$. In this situation, all the conditions of the theorem and its corollaries are satisfied if $dw/dx = (2/\pi)^{\frac{1}{2}}e^{-x^2/2}[x > 0]$. Below, we provide expressions for $r(H_{n+1}, \alpha)$, and the bound in Theorem 3, and the form for S_n.

To obtain $r(H_{n+1}, \alpha)$, notice that

$$r(H_{n+1}, \alpha) = \int \{E[F^2(u)] - E[d_\alpha^2(u, (\delta_{n+1}, Z_{n+1}))]\}\,dw(u) \tag{20}$$

where d_α is given by (3). Obviously,

$$E[F^2(u)] = \alpha(u)(\alpha(u)+1)/c(c+1)\,, \tag{21}$$

since $1-F$ is a Dirichlet process with parameter α. Also, in view of (3),

$$\begin{aligned}(c+1)^2&E[d_\alpha^2(u, (\delta_{n+1}, Z_{n+1}))]\\ &= (1+ce^{-\theta u})^2e^{-(\beta+\theta)u} + c^2e^{-2\theta u}P[\delta_{n+1}=1, Z_{n+1}\leqq u]\\ &\quad + e^{-2\theta u}E[[\delta_{n+1}=0, Z_{n+1}\leqq u]e^{2\theta Z_{n+1}}(1+ce^{-\theta Z_{n+1}})^2]\\ &= (1+ce^{-\theta u})^2e^{-(\beta+\theta)u} + c^2e^{-2\theta u}(\theta/(\theta+\beta))(1-e^{-(\beta+\theta)u})\\ &\quad + \beta e^{-2\theta u}\{(\theta-\beta)^{-1}(e^{u(\theta-\beta)}-1) + 2c\beta^{-1}(1-e^{-\beta u})\\ &\quad + c^2(\beta+\theta)^{-1}(1-e^{-(\beta+\theta)u})\}\,.\end{aligned} \tag{22}$$

Combining (20), (21), and (22) gives the desired expression for $r(H_{n+1}, \alpha)$. Since $r(H_{n+1}, \alpha) = \int E[F^2(u)]\,dw(u) \leqq (c+1)^{-1}\int e^{-\theta u}(1+ce^{-\theta u})\,dw(u)$, it easily follows that $r(H_{n+1}, \alpha) \leqq \{ce^{-2\theta^2} + e^{-\theta^2/2}\}/(1+c) \leqq 1$.

In the situation described above, $\hat{S}_n$ takes the following simple form.

$$\begin{aligned}(c+1)&\hat{S}_n(u, (\delta_{n+1}, Z_{n+1}))\\ &= 1 + c\min\{e^{\beta u}\hat{G}_n(u), 1\} \qquad \text{if } Z_{n+1} > u\\ &= c\min\{e^{\beta u}\hat{G}_n(u), 1\} \qquad \text{if } \delta_{n+1}=1\,, \text{ and } u \geqq Z_{n+1}\\ &= \{1 + c\min\{e^{\beta Z_{n+1}}G_n(Z_{n+1}), 1\}\}\min\left\{\frac{e^{\beta(u-Z_{n+1})}\hat{G}_n(u)}{\hat{G}_n(Z_{n+1})}\,,\ 1\right\}\\ &\qquad \text{if } \delta_{n+1}=0\,, \text{ and } u \geqq Z_{n+1}\,,\end{aligned} \tag{23}$$

where $n\hat{G}_n(\cdot) = \sum_{j=1}^n [Z_j > \cdot]$. This estimator is easily programmable because the only quantity that changes with n is $\hat{G}_n$ which can be recursively obtained since $n\hat{G}_n(\cdot) = (n-1)\hat{G}_{n-1}(\cdot) + [Z_n > \cdot]$ for $n = 1, 2, \cdots$. The question of estimating β can be argued as follows. Since $\bar{Z} (= n^{-1} \sum_{j=1}^n Z_j)$ and $1 - \bar{\delta}$ $(\bar{\delta} = n^{-1} \sum_{j=1}^n \delta_j)$ are consistent moment estimators of $(\theta + \beta)^{-1}$ and $\beta(\theta + \beta)^{-1}$ respectively, it follows that $\hat{\beta} = (1 - \bar{\delta})/\bar{Z}$ is a consistent estimator of β. If β is assumed to be unknown, then one can use (23) with β replaced by $\hat{\beta}$.

By bounding the integrals in the rhs of the bound in Theorem 3 by $2\gamma_n(u)$, we obtain that this bound can simply be bounded by $nc\beta/(2\theta + \beta)$, and consequently, it can be shown that $\hat{S}_n$ is such that

$$D_n \leqq c_1 c n^{-1}[\{\beta/(2\theta + \beta)\} + e^{-(\beta-\theta)^2/2} + e^{-2\theta^2}] , \tag{24}$$

where c_1 is an absolute constant.

The estimator $\hat{S}_n$ of (23) was applied to a practical example involving survival times of melanoma patients. We obtain the survival curve estimator $\hat{S}_n$ of (23) using this data. This curve along with the actual survival data are given in the Appendix.

5. Empirical Bayes estimation with $\{H_n\}$ unknown. In this section, we obtain a rate of convergence of D_n, defined in (5), to zero corresponding to $\hat{\hat{S}}_n$ (see around (27)) when $\{H_n\}$ is unknown under some mild restrictions on $\{H_n\}$ and α. To startoff, the following assumption is made throughout this section:

$$\alpha(R) \text{ is known}, \quad \alpha(0) = \alpha(R) \quad \text{and support of} \quad H_n = (a_n, \infty) \quad \text{where} \quad \{a_n\} \subset R^\infty . \tag{25}$$

Throughout $c_1, \cdots, c_n$ denote constants. For ease of notation, let $\alpha(0) = \alpha(R) = 1$.

The plan of this section is as follows: we first obtain estimates of $(\alpha(x))(\bar{H}(x))$ and $-\alpha'(x)(\bar{H}(x))$ where $n\bar{H}(x) = \sum_{j=1}^n H_j(x)$ and $\alpha' = d\alpha/dx$. By taking the integral of the ratio of these two estimators over $(0, t]$, we obtain an estimate for $-\int_0^t \{\alpha'(x)/\alpha(x)\}\, dx = -\ln(\alpha(t)\cdot)$ where the equality follows by the assumption $\alpha(0) = 1$. From this estimate of $\ln(\alpha(t))$, we obtain an estimate $\hat{\alpha}_n(t)$ of $\alpha(t)\,(= e^{\ln(\alpha(t))})$. $\hat{\alpha}(\cdot)$ when substituted in $d_\alpha(\cdot, (\delta_{n+1}, z_{n+1}))$ of (3) gives the $(n+1)$th component $\hat{\hat{S}}_n$ of the empirical Bayes estimator whose rate of risk convergence is the main result of this section.

For first reading, one can look at the form of $\hat{\hat{S}}_n$ given below and the main result of this section (Theorem 4) omitting the details in between. For defining $\hat{\hat{S}}_n$, assume that

$$-\alpha'(x)/\alpha(x) \leqq r(x) , \quad \text{a known function,} \tag{26}$$

and that K is a known real valued bounded function on R vanishing off $(0, u_1)$, $u_1 < \infty$, such that $\int u^j K(u)\, du = 0$ for $j = 1, \cdots, l-1$ with l a fixed positive integer and $\int K(u)\, du = 1$. It will be shown below that $(n\varepsilon_n)^{-1}\alpha(R) \sum_{j=1}^n [\delta_j = 1]$ $K((Z_j \cdots x)/\varepsilon_n)$ is a good estimator of $-\alpha'(x)\bar{H}(x)$ whenever $\varepsilon_n \downarrow 0$, and obviously, $n^{-1}\alpha(R) \sum_{j=1}^{n} [Z_j > x]$ is a good estimate of $\alpha(x)\bar{H}(x)$. Consequently, the integral of the ratio of these estimators over $(0, t]$ will be a good estimator of $-\int_0^t \{\alpha'(x)\bar{H}(x)/\alpha(x)\bar{H}(x)\}\, dx = \ln \alpha(t)$. In view of this discussion, the estimator $\hat{\hat{S}}_n$ is defined by (15) with $\hat{\alpha}_n$ there replaced by $\hat{\hat{\alpha}}_n$ where

$$\hat{\hat{\alpha}}_n(t) = \exp(-\textstyle\int_0^t \hat{\varphi}_n(x)\, dx) \tag{27}$$

with

$$\hat{\varphi}_n(x) = \max\left\{\min\left\{\frac{\sum_{j=1}^{n} [\delta_j = 1]K((Z_j - x)/\varepsilon_n)}{\varepsilon_n \sum_{j=1}^{n} [Z_j > x]}\,,\ r(x)\right\},\ 0\right\}. \tag{28}$$

The next three results are designed to lead to a bound on the mean square error of $\hat{\hat{\alpha}}_n$ as an estimate of α. Since $Z_1, \cdots, Z_n$ are independent with $P[Z_j > x] = \alpha(x)H_j(x)$ for $j = 1, \cdots, n$, it readily follows that

$$\begin{aligned} n^2 E[\{n^{-1} \textstyle\sum_{j=1}^{n} [Z_j > x] - (\bar{H}(x))\alpha(x)\}^2] \\ = (\alpha(x))\{n(\bar{H}(x)) - (\alpha(x)) \textstyle\sum_{j=1}^{n} (H_j(x))^2\}. \end{aligned} \tag{29}$$

LEMMA 3. *If α' and the lth derivatives of α', $H_1, \cdots, H_n$ are uniformly bounded, then*

$$\begin{aligned} E[\{(n\varepsilon_n)^{-1} \textstyle\sum_{j=1}^{n} [\delta_j = 1]K((Z_j - x)/\varepsilon_n) - (\bar{H}(x))\alpha'(x)\}^2] \\ \leqq c_1(\varepsilon_n^{2l} + (n\varepsilon_n)^{-1}). \end{aligned} \tag{30}$$

PROOF. The distribution of (δ_j, Z_j) for $j = 1, \cdots, n$ and a change of variable imply that

$$\begin{aligned} E[(n\varepsilon_n)^{-1}\alpha(R) \textstyle\sum_{j=1}^{n} [\delta_j = 1]K((Z_j - x)/\varepsilon_n) + \alpha'(x)\bar{H}(x)] \\ = -[\textstyle\int K(v)\{\alpha'(x + \varepsilon_n v)(\bar{H}(x + \varepsilon_n v)) - \alpha'(x)(\bar{H}(x))\}\, dv] \end{aligned} \tag{31}$$

since $\int K(v)\, dv = 1$. Expanding the expression in the curly brackets in (31), using the lth order Taylor expansion, and then using the orthogonality conditions on K, one obtains that

$$|\text{lhs of (28)}| \leqq c_2 \varepsilon_n^{\,l}. \tag{32}$$

Moreover, by independence of $(\delta_1, Z_1), \cdots, (\delta_n, Z_n)$ and a change of variable,

$$\begin{aligned} (n\varepsilon_n)^2 \operatorname{Var}\{(n\varepsilon_n)^{-1} \textstyle\sum_{j=1}^{n} [\delta_j = 1]K((Z_j - x)/\varepsilon_n)\} \\ \leqq \textstyle\sum_{j=1}^{n} E[[\delta_j = 1]K^2((Z_j - x)/\varepsilon_n)] \\ \leqq n\varepsilon_n \textstyle\int K^2(v)(\bar{H}(x + \varepsilon_n v))\alpha'(x + \varepsilon_n v)\, dv \leqq c_3 n\varepsilon_n, \end{aligned} \tag{33}$$

where the last inequality follows since α' is bounded by assumption. Adding the square of (32) to (33) gives (30). □

A needed corollary to (29) and (30) is

COROLLARY 4. *Let $0 < \varepsilon_n < 1$. Under conditions of Lemma 3,*

$$E[|\hat{\varphi}_n(x) - (\alpha'(x)/(\alpha(x)))|^2] \leqq c_3(1 + r^2(x))(\alpha(x))^{-2}(\bar{H}(x))^{-2}\{\varepsilon_n^{2l} + (n\varepsilon_n)^{-1}\}.$$

PROOF. By the definition of $\hat{\varphi}_n$ in (28), $|\hat{\varphi}_n(x) - \{\alpha'(x)/(\alpha(x))\}|^2$ is exceeded by

$$\left\{\min\left\{\left|\frac{\sum_{j=1}^{n} [\delta_j = 1]K((Z_j - x)/\varepsilon_n)}{\varepsilon_n \sum_{j=1}^{n} [Z_j > x]} + \frac{\alpha'(x)}{\alpha(x)}\right|,\ r(x)\right\}\right\}^2 \tag{34}$$

since $-\alpha'(x)/(\alpha(x)) \leqq r(x)$. We now apply Lemma 2 of Section 4 to the expectation of (34). Applying Lemma 2 with

$$A = \sum_{j=1}^{n} \alpha(R)[\delta_j = 1]K((Z_j - x)/\varepsilon_n)/n\varepsilon_n\,, \qquad B = n^{-1}\alpha(R)\sum_{j=1}^{n}[Z_j > x]\,,$$

$a = -\alpha'(x)\bar{H}(x)$ and $b = \alpha(x)\bar{H}(x)$ (> 0 if $\alpha(x) > 0$ since $\bar{H}(x) > 0$ by assumption), $l = r(x)$ and $\gamma = 2$ results in

$$E[(34)] \leqq c_3(1 + r^2(x))(\alpha(x))^{-2}(\bar{H}(x))^{-2}\{\varepsilon_n^{2l} + (n\varepsilon_n)^{-1}\}\,, \tag{35}$$

after some simplification for $\alpha(x) > 0$ due to Lemma 3.

If $\alpha(x) = 0$, then $\alpha'(x)/(\alpha(x)) = 0$ by our convention. Also, $P[Z_j > x] = 0$ for $j = 1, \cdots, n$. Therefore, $[Z_j > x]$ and $K((Z_j - x)/\varepsilon_n)$ are both zero with probability one for $j = 1, \cdots, n$ since K vanishes on $(-\infty, 0)$. Thus both ratios in the modulus in (34) are zero with probability one by our convention and hence, if $1 - \alpha(x) = 0$, so is (34). This together with (34) and (35) completes the proof. □

Recalling that $\hat{\hat{\alpha}}_n(t) = \exp\{\int_0^t \hat{\varphi}_n(x)\,dx\}$ is the proposed estimator for $\alpha(t)$, its mean square error can be bounded as in the following subcorollary.

SUBCOROLLARY 1. *Under the conditions of Corollary* 4,

$$E[|\hat{\hat{\alpha}}_n(t) - \alpha(t)|^2] \leqq c_3 t\{\textstyle\int_0^t (1 + r^2(x))(\alpha(x))^{-2}(\bar{H}(x))^{-2}\,dx\}\{\varepsilon_n^{2l} + (n\varepsilon_n)^{-1}\}\,.$$

PROOF. The result follows immediately from Corollary 4 upon rewriting the lhs of the result as $E[|\exp\{-\int_0^t \hat{\varphi}_n(x)\,dx\} - \exp\{-\int_0^t (\alpha'(x)/(\alpha(x)))\,dx\}|^2]$ and observing that this expectation is $\leqq$

$$E[|\textstyle\int_0^t \{\hat{\varphi}_n(x) + \{\alpha'(x)/(\alpha(x))\}\,dx\}|^2] \leqq t\int_0^t E[\{\hat{\varphi}_n(x) + \{\alpha'(x)/(\alpha(x))\}\}^2]\,dx$$

where the first inequality follows by applying the mean value theorem and noting that $\int_0^t \hat{\varphi}_n(x)\,dx$ and $\int_0^t \{\alpha'(x)/(\alpha(x))\}\,dx \geqq 0$ and the second inequality follows from the inequality between first and second moments. □

The following main result can be proved using Lemma 2 of Section 4, the above subcorollary and the method given for the proof of theorem except for the following change in the case when $\alpha(Z_{n+1}) = 0$: If $a^* = \inf\{a \mid \alpha(a, \infty) = 0\}$, then $Z_{n+1} \leqq a^*$ with probability one so $\alpha(z_{n+1}) > 0$ for $z_{n+1} < a^*$ and $P[Z_{n+1} \leqq a^*] = 1$. Since Z_{n+1} has a density wrt Lebesgue under the conditions of Lemma 3, $P[Z_{n+1} = a^*] = 0$. So we essentially have to consider only those z_{n+1} for which $\alpha(z_{n+1}) > 0$ which is necessary for the application of Lemma 2 for our situation here as in the proof of Theorem 3.

Pulling together all the assumptions made so far, we state the main result of this section as

THEOREM 4. *Let $\alpha(R)$ be known and support of $H_n = (a_n, \infty)$. If α' and the* lth *derivatives of α', $H_1, \cdots, H_n$ are uniformly bounded and $-\alpha'(x)/(\alpha(x)) \leqq r(x)$, a known function, then $\hat{\hat{S}}_n$ defined by* (15) *with $\hat{\alpha}_n$ replaced by $\hat{\hat{\alpha}}_n$ of* (27) *with*

$\varepsilon_n^{2l+1} = n^{-1}$ *is such that*

$$n^{2l/(2l+1)} D_n = n^{2l/(2l+1)}\{r_n(\hat{\hat{S}}_n, \{H_j\}_1^{n+1}, \alpha) - r(H_{n+1}, \alpha)\}$$

$$\leq c_4 \int_0^\infty u\{\int_0^u (1 + r^2(x))(\alpha(x))^{-2}(\bar{H}(x))^{-2}\, dx\}\, dw(u)\,. \tag{36}$$

If the rhs *of* (36) *is uniformly bounded in* n, *then* $\hat{\hat{S}}_n$ *is* a.o. $(n^{-2l/(2l+1)})$.

REMARK 2. The above rate result is close to the best possible exact rate $O(n^{-1})$ in view of Corollary 2 of Section 4 for large l.

EXAMPLE 2. Here we explain how the above theorem can be applied to the situation in which $H_n(u) = e^{-\beta u}$ and $\alpha(u) = ce^{-\theta u}$ for $u, c > 0$ where θ and β are unknown. However, we assume that $\theta \leq \theta^*$, a known quantity, and hence $-\alpha'(x)/\alpha(x) \leq \theta \leq \theta^* = r(x)$ for $x > 0$. We take the kernel function K in our estimator (27) to be

$$K(u) = 4 - 6u \quad \text{if} \quad 0 \leq u \leq 1\,, \qquad = 0 \quad \text{otherwise.} \tag{37}$$

Then $\int K(u)\, du = 1 = 1 - \int uK(u)\, du$. Thus K satisfies the conditions imposed on it in the above theorem with $u_1 = l = 1$. Take $\varepsilon_n^3 = n^{-1}$ in (27) and (28). If $dw/dx = (2/\pi)^{\frac{1}{2}} e^{-x^2/2}[x > 0]$, then the bound (36) is finite. Consequently, the above theorem gives a.o. empirical Bayes distribution function estimators with rate $O(n^{-\frac{2}{3}})$ corresponding to $l = 1$. Also, observe that the expression for $r(H_{n+1}, \alpha)$ is given by (20), (21) and (22). The bound in (36) is given by

$$c_4(1 + \theta^{*2})\pi^{-1}e^{2(\theta+\beta)^2}\,. \tag{38}$$

If H_n and α are gamma distributions with parameters bounded above instead of the exponential distributions as described above, then $r(x)$ can be taken to be a polynomial, and hence, the bound (36) will still be finite. Therefore, the conclusion of the theorem follows. If K is as given in (37), then the rate is again $O(n^{-\frac{2}{3}})$.

REMARK 3. As in the known $\{H_n\}$ case, the boundedness of (33) is really a tail condition on $\bar{H}$, α and w.

6. Some possible allied results, generalizations, and concluding remarks. If α is known to be such that $\alpha(R) = \alpha(-\infty, a]$, then the weight function w in the loss function L of (2) should assign all its mass to $(-\infty, a]$ since $P[F_n \mid F_n(a) = 1] = 1$. If a is such that $H_n(a) = 1 > H_n(b)$ for all $b < a$ and n, then we should restrict the range of the integral in the loss function (2) to $(-\infty, a]$ since there is no way of estimating $\alpha(-\infty, c]$ (or $\alpha(c, \infty)$) for $c > a$ in view of Theorem 2 of Section 3 on identifiability, the form of d_α in (2) and Lemma 1 of Section 2.

The results of this paper can be used to obtain a.o. procedures in empirical Bayes monotone multiple testing problems, thus generalizing Theorem 3.1 of Susarla and Phadia (1976) in the uncensored case to the case when $\{H_n\}$ is known and a slightly weaker result when $\{H_n\}$ is unknown. The results given in this paper can be extended to the situation when one has m_n (in $\{1, \cdots, \bar{m}\}$) observations in the nth component problem with tedious manipulations. If F is distributed according to a process neutral to right (see Doksum (1974)), it should be possible to obtain results similar to Theorems 3 and 4, thus giving the empirical Bayes results with the component problem as that described in Ferguson and Phadia (1975) in a more general situation than that considered by them in their Bayes estimation problem.

Acknowledgments. The authors are grateful to an associate editor for his detailed comments and suggestions which improved the clarity of the paper. Also, the suggestion of both the associate editor and the editor leading to the addition of the numerical example is gratefully acknowledged.

APPENDIX

Listed below are the survival times (in weeks) of 81 participants from a melanoma study conducted by the Central Oncology Group with headquarters office at the University of Wisconsin–Madison. The survival times which are censored are indicated with a + sign. They are listed sequentially in order of entry into the study.

136,	58,	55+,	181+,	21,	23,	190+,	65,	234
194+,	14,	90,	20,	130,	213+,	215+,	124,	108+
54,	98,	193+,	138,	141,	110,	67+,	50,	26
103,	59,	134+,	147+,	152+,	65,	40,	34,	57
81+,	152+,	125+,	151+,	34,	158+,	27,	148+,	27
132+,	140+,	32,	130+,	38,	85,	129+,	100+,	19
118,	53,	120+,	66,	46,	37,	50+,	114+,	124+
26,	102,	93+,	80+,	60,	86+,	21+,	44+,	23
70,	73+,	19,	38,	31,	25,	76+,	13,	16+

Using these data, we compute the empirical Bayes estimator given by (23) with $\delta_{n+1} = 0$, $Z_{n+1} = 16$, $n = 80$ and $\beta = \hat{\beta} = (1 - \bar{\delta})/\bar{Z}$ as the empirically determined value of β as described immediately after equation (23). For the abvoe data, $1 - \bar{\delta} = 1 - \frac{46}{80} = .425$, $\bar{Z} = \frac{7055}{80} = 88.1875$ and $\hat{\beta} = (1 - \bar{\delta})/\bar{Z} = .00482$. Since $\exp(\hat{\beta} Z_{n+1})\hat{G}_n(16) = \exp(16\hat{\beta})(\frac{78}{80}) = 1.053 > 1$, we see that for this example the estimator (23) with $\beta = \hat{\beta}$ becomes,

$$\hat{S}_n(u) = \frac{1}{c+1}\{1 + c \min(\exp(\hat{\beta} u)\hat{G}_n(u), 1)\} \qquad \text{if} \quad u < 16$$

$$= \min\left(\frac{\exp(\hat{\beta}(u-16))\hat{G}_n(u)}{(\frac{78}{80})}, 1\right) \qquad \text{if} \quad u \geqq 16 ,$$

or more simply since $\exp(\hat{\beta}u)G_n(u) > 1$ for $u < 16$,

$$\hat{S}_n(u) = 1 \quad \text{if} \quad u < 16$$
$$= \min\left(\frac{\exp(\hat{\beta}(u-16))\hat{G}_n(u)}{(\frac{78}{80})}, 1\right) \quad \text{if} \quad u \geqq 16 .$$

The table below gives the values of this estimator for various values of u based on the above data.

	$\leqq 16$	20	40	60	80	100	120	140	160	180	200	220	233	$\geqq 234$
$\hat{S}_n(u)$	1	.980	.835	.745	.681	.615	.529	.326	.180	.198	.093	.034	.036	0

Note that $\hat{S}_n(180) = .198 > .180 = \hat{S}_n(160)$. This peculiarity of an increasing survival estimator over a short range is brought about by the product of 2 factors $\exp(\hat{\beta}(u-16))$ and $\hat{G}_n(u)$, the first $\exp(\hat{\beta}(u-16))$ increasing continuously in u and the second $\hat{G}_n(u)$ decreasing by discrete jumps. Thus over an interval in which $\hat{G}_n(u)$ stays constant, $\hat{S}_n(u)$ increases. This appears to be an undesirable feature of the estimator caused by use of $(H_j(x))^{-1}$ in equation (12). We have not investigated alternative estimators $\hat{\alpha}_n(x)$ defined by (12) which avoid this problem. This would be an interesting area for future work.

REFERENCES

[1] ANTONIAK, C. (1974). Mixtures of Dirichlet processes with applications to Bayesian nonparametric problems. *Ann. Statist.* **2** 1152–1174.

[2] BARLOW, R. and PROSCHAN, F. (1974). *Statistical Theory of Reliability and Life Testing-Probability Models.* Holt, Rinehart, and Winston, New York.

[3] CHUNG, KAI LAI (1974). *A Course in Probability Theory*, 2nd ed. Academic Press, New York.

[4] DOKSUM, K. A. (1974). Tailfree and neutral random probabilities and their posterior distributions. *Ann. Probability* **2** 183–201.

[5] EFRON, B. (1967). The two sample problem with censored data. *Proc. Fifth Berkeley Symp. Math. Statist. Prob.* **4** 831–854, Univ. of California Press.

[6] FERGUSON, T. S. (1973). A Bayesian analysis of some nonparametric problems. *Ann. Statist.* **1** 209–230.

[7] FERGUSON, T. S. and PHADIA, E. G. (1976). Private communications.

[8] GROSS, A. J. and CLARK, V. (1975). *Survival distributions: Reliability Applications in Biomedical Sciences.* Wiley, New York.

[9] JOHNS, M. V., JR. (1957). Non-parametric empirical Bayes procedures. *Ann. Math. Statist.* **28** 649–669.

[10] KAPLAN, E. L. and MEIER, P. (1958). Nonparametric estimation from incomplete observations. *J. Amer. Statist. Assoc.* **53** 457–481.

[11] KORWAR, R. and HOLLANDER, M. (1976). Empirical Bayes estimation of a distribution function. *Ann. Statist.* **4** 581–588.

[12] MARITZ, J. S. (1970). *Empirical Bayes Methods.* Methuen, London.

[13] O'BRYAN, T. and SUSARLA, V. (1975). An empirical Bayes estimation problem with nonidentical components involving normal distributions. *Comm. Statist.* **4** 1033–1042.

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

[15] SINGH, R. S. (1974). Estimation of derivatives of average of μ-densities and sequence-compound estimation in exponential families. RM-318, Dept. of Statist. and Prob., Michigan State Univ.

[16] SUSARLA, V. and PHADIA, E. G. (1976). Empirical Bayes testing of a distribution function with Dirichlet process priors. *Comm. Statist.* **5** 455–469.

[17] SUSARLA, V. and VAN RYZIN, J. (1976). Nonparametric Bayesian estimation of survival curves from incomplete observations. *J. Amer. Statist. Assoc.* **71** 897–902.

[18] TEICHER, H. (1960). On the mixtures of distributions. *Ann. Math. Statist.* **31** 55–73.

[19] VAN RYZIN, J. (1970). On some nonparametric empirical Bayes multiple decision problems. *Proc. First Internat. Symp. Nonparametric Techniques in Statist. Inference*, 585–601.

DEPARTMENT OF MATHEMATICS
UNIVERSITY OF WISCONSIN
MILWAUKEE, WISCONSIN 53201

DEPARTMENT OF MATHEMATICS
UNIVERSITY OF CALIFORNIA
DAVIS, CALIFORNIA 95616